

Teaching and Learning ^1H Nuclear Magnetic Resonance Spectroscopy

by

Megan C. Connor

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in the University of Michigan
2021

Doctoral Committee:

Assistant Professor Ginger V. Shultz, Chair
Professor Timothy A. McKay
Professor Anne J. McNeil
Associate Professor Kerri Pratt

Megan C. Connor

mcarole@umich.edu

ORCID iD: 0000-0003-3266-4162

© Megan C. Connor 2021

Dedication

To my Grandpa.

Thank you for being so incredibly proud of me.

Acknowledgements

I am grateful to a number of people for their mentorship, friendship, support, and encouragement as I completed this work over the past several years. Professor Ginger Shultz helped me develop into a confident chemistry education researcher while also looking after my emotional and physical wellbeing. She demonstrated on a daily basis that it is possible to be both excellent and kind, one of the most valuable lessons that I will carry with me throughout my career. Dr. Solaire Finkenstaedt-Quinn helped me think through innumerable research problems and shared many laughs with me; I am not sure which of these two acts proved more important for completing this work, but I am thankful for both. Eleni Zotos and I also shared many laughs as we learned to navigate graduate school and the field of chemistry education research together. Her kindness and positivity made each day that much brighter. Amanda Kubic and Olivia Hankinson helped make Ann Arbor my home away from home, and they filled my time outside of work with happiness and love; I was very lucky to live with such wonderful friends. Lastly, my family regularly expressed their love and encouragement through phone calls, texts, care packages, and visits. I attribute my accomplishments in no small part to their never-ending support.

Table of Contents

Dedication	ii
Acknowledgements	iii
List of Figures	xii
List of Schemes	xvi
List of Tables	xvii
Abstract	xxii
Chapter 1: Introduction.....	1
1.1 Document overview	1
1.2 Abstract.....	2
1.3 The role of spectroscopic analysis in organic synthesis	3
1.4 Research investigating the teaching and learning of spectroscopic structure elucidation.....	4
1.4.1 NMR and IR spectroscopy as the focus of research.....	5
1.4.2 Theoretical and conceptual frameworks guiding research on learning	8

1.4.3 Methodological trends: study populations, data collection, and analysis	11
1.5 Empirical insights into teaching and learning spectroscopic structure elucidation	13
1.5.1 Insights into problem-solving aspects of spectroscopic structure elucidation	14
1.5.2 Insights into reasoning aspects of spectroscopic structure elucidation.....	16
1.5.3 General insights provided by problem-solving and reasoning perspectives ..	22
1.5.4 Insights into teaching spectroscopic structure elucidation	22
1.6 Instructional innovations	24
1.6.1 Scaffolding strategies	25
1.6.2 Laboratory experiments and activities.....	28
1.7 Implications.....	29
1.7.1 Implications for research	30
1.7.2 Implications for instruction.....	32
1.8 Conclusions	35
1.9 References	35
Chapter 2: Constraints on Organic Chemistry Students' Reasoning During IR and ¹H NMR Spectral Interpretation	41
2.1 Initial remarks.....	41
2.2 Abstract.....	43

2.3 Introduction	43
2.3.1 Teaching and learning IR and NMR spectroscopy	44
2.4 Theoretical framework	45
2.4.1 Categorization of mental representations	45
2.4.2 Assumptions and heuristics as cognitive constraints	46
2.5 Research questions	48
2.6 Methods	49
2.6.1 Sample and setting	49
2.6.2 Data collection	50
2.6.3 Description of interpretation tasks	52
2.6.4 Qualitative analysis of RTA interviews	55
2.6.5 Quantitative analysis of assumptions and heuristics	56
2.7 Results and discussion	57
2.7.1 Qualitative findings: invalid chemical assumptions	58
2.7.2 Qualitative findings: heuristic reasoning strategies	65
2.7.3 Quantitative findings: identification of invalid chemical assumptions and heuristics that most severely constrained reasoning	73
2.8 Conclusions	78

2.9 Limitations.....	80
2.10 Implications for teaching and research	81
2.10.1 Implications for teaching	81
2.10.2 Implications for research	82
2.11 Supporting Information	83
2.12 References	85
Chapter 3: Developing Expertise in ¹H NMR Spectral Interpretation	92
3.1 Initial remarks.....	92
3.2 Abstract.....	94
3.3 Introduction	94
3.3.1 Mapping the development of expertise in ¹ H NMR spectral interpretation.....	95
3.3.2 Mapping the evolution of conceptual understanding	96
3.3.3 Mapping the evolution of information processing.....	97
3.4 Expertise analysis	98
3.5 Results and discussion	100
3.5.1 Chemical assumptions	100
3.5.2 Cognitive processes	103
3.6 Conclusions	108

3.7 Experimental section	109
3.7.1 Sample and setting	109
3.7.2 Data collection.....	110
3.7.3 Qualitative analysis of RTA interviews	114
3.7.4 Quantitative analysis of eye-tracking data	116
3.7.5 Limitations	117
3.8 Supporting Information	120
3.8.1 Additional description of areas of understanding, levels, and contributing assumptions	128
3.8.2 Description of transitions not described in the text.....	134
3.8.3 Instructional model for graduate and undergraduate-level instructors	135
3.9 Acknowledgements.....	137
3.10 References	137
Chapter 4: Teaching Assistants' Topic-Specific Pedagogical Content Knowledge in ¹H NMR Spectroscopy	142
4.1 Initial remarks.....	142
4.2 Abstract.....	144
4.3 Introduction	145
4.4 Theoretical Framework and Background.....	147

4.4.1 Pedagogical Content Knowledge	147
4.4.2 Content knowledge	148
4.4.3 Research on TS-PCK in chemistry topics	149
4.5 Methods	150
4.5.1 Participants	151
4.5.2 Questionnaire Design.....	152
4.5.3 Data analysis.....	155
4.6 Results	156
4.6.1 Questionnaire development and Rasch model validity	156
4.6.2 Relationship between CK and PCK person measures.....	157
4.6.3 Dependence of PCK on CK	160
4.6.4 Relationship between PCK and teaching experience	161
4.6.5 Identifying a leverage point for transfer of TS-PCK.....	162
4.6.6 Teaching strategies of TAs	163
4.6.7 Problem-specificity of PCK.....	164
4.7 Discussion	166
4.8 Conclusions	168
4.9 Limitations.....	168

4.10 Implications	169
4.11 Supporting information.....	170
4.11.1 Exemplars and second tier response frequencies	170
4.11.2 CK component of questionnaire	176
4.11.3 PCK component of questionnaire	178
4.12 References	180
Chapter 5: Development of the NMR Lexical Representational Competence (NMR-LRC) Instrument as a Formative Assessment of Lexical Ability in ¹H NMR Spectroscopy	185
5.1 Initial remarks.....	185
5.2 Abstract.....	186
5.3 Introduction	187
5.4 Conceptual framework	190
5.4.1 Representational competence	190
5.4.2 Lexical representational competence.....	191
5.4.3 Illusions of lexical representational competence.....	193
5.4.4 Methodological interactionalism	193
5.5 Research goals	195
5.6 Methods	196

5.6.1 Participants and instructional context.....	196
5.6.2 Instrument development.....	197
5.6.3 Instrument evaluation	199
5.7 Results and discussion	203
5.7.1 Descriptive statistics.....	203
5.7.2 Statistical evidence of validity and reliability from Rasch analysis	204
5.7.3 Supporting sources of validity evidence	211
5.7.4 Cluster analysis	213
5.8 Limitations.....	216
5.9 Conclusions and implications	217
5.9.1 Implications for research	218
5.9.2 Implications for teaching	219
5.10 Acknowledgements.....	221
5.11 Supporting Information	222
5.12 References	222
Chapter 6: Closing Remarks	229

List of Figures

- Figure 1.1.** A summary of the chapter's contents, including hallmarks of ability and typical characteristics of undergraduates' reasoning and approach identified through literature on the teaching and learning spectroscopic structural elucidation, as well as potential avenues for developing this ability.5
- Figure 1.2.** Spectral interpretation task in which undergraduates in the Connor et al. (2019) study assumed signals corresponding to hydrogen atoms on nitrogen and oxygen near 6.8 ppm and 4.2 ppm, respectively, should appear as triplets rather than singlets. 19
- Figure 1.3.** Spectral interpretation task in which undergraduates in the Connor et al. (2019) study assumed the broad IR peak characteristic of a hydroxyl group near 3000 cm^{-1} corresponded to a carbon-hydrogen bond.20
- Figure 2.1.** Spectral interpretation task (Synthesis 1) asking participants to determine if N-(2-hydroxyethyl)-propanamide was successfully synthesized using the provided IR spectrum and ^1H NMR spectrum.52
- Figure 2.2.** Compounds identified as desired products in each interpretation task. .53
- Figure 2.3.** Co-occurrence of assumptions and heuristics in incorrect responses. ..76
- Figure 2.4.** Spectral interpretation task (Synthesis 2) asking participants to determine if isochroman was successfully synthesized using the provided IR spectrum and ^1H NMR spectrum.83

Figure 2.5. Spectral interpretation task (Synthesis 3) asking participants to determine if 3-(allyloxy)propanal was successfully synthesized using the provided IR spectrum and ¹ H NMR spectrum.....	84
Figure 3.1. Spectral interpretation task (Synthesis 1) asking participants to determine if N-(2-hydroxyethyl)-propanamide was successfully synthesized using the provided IR spectrum and ¹ H NMR spectrum.....	99
Figure 3.2. Five areas of understanding and description of increasing levels of sophistication.....	101
Figure 3.3. Chord diagrams of significantly different transition counts between undergraduate and doctoral participants for Syntheses 1-3.....	107
Figure 3.4. Spectral interpretation task (Synthesis 1) for which participants were asked to determine if N-(2-hydroxyethyl)-propanamide was successfully synthesized using the provided IR and ¹ H NMR spectra.	122
Figure 3.5. Spectral interpretation task (Synthesis 2) for which participants were asked to determine if isochroman was successfully synthesized using the IR and ¹ H NMR spectra.....	123
Figure 3.6. Spectral interpretation task (Synthesis 3) for which participants were asked to determine if 3-(allyloxy)propanal was successfully synthesized using the IR and ¹ H NMR spectra.....	123
Figure 3.7. IR and NMR tables provided as the second page in each of the interpretation tasks.	124
Figure 3.8. Open-ended assessment that requires students to explain how specific structural features result in spectral peaks.	136

Figure 4.1. Item-person map of TA CK in ^1H NMR spectroscopy..... 158

Figure 4.2. Item-person map of TA PCK in ^1H NMR spectroscopy. 158

Figure 4.3. Scatterplot of participants' CK and PCK measures ($r=0.670$, $p<0.01$). Higher CK logit measures indicate a greater ability to correctly respond to CK questions, and higher PCK logit measures indicate a greater ability to adequately respond to PCK questions. 159

Figure 4.4. CKQ1: Determine the molecular structure corresponding to the given molecular formula and ^1H NMR spectrum. 176

Figure 4.5. CKQ2: Determine the molecular structure corresponding to the given molecular formula and ^1H NMR spectrum. 176

Figure 4.6. CKQ3: In which molecules are H_a and H_b equivalent? 177

Figure 4.7. CKQ4: Circle all molecule pairs that are distinguishable using ^1H NMR spectroscopy. 177

Figure 4.8. CKQ5: How many ^1H NMR signals are produced by the following compound? 177

Figure 4.9. PCKQ1, which assessed knowledge of what makes this topic difficult. 178

Figure 4.10. PCKQ2, which assessed what makes this topic difficult and teaching strategies. 178

Figure 4.11. PCKQ3, which assessed what makes this topic difficult and teaching strategies. 179

Figure 4.12. PCKQ4, which assessed what makes this topic difficult and teaching strategies. 179

Figure 5.1. NMR-LRC multiple-choice question and associated confidence tier.... 197

Figure 5.2. Descriptive statistics of NMR-LRC data with A) distribution of participants' total raw scores on the NMR-LRC, B) distribution of participants' average confidence ratings, and (C) a scatterplot of participants' total raw scores and average confidence ratings.....203

Figure 5.3. Wright map of NMR-LRC participant ability and item difficulty measures on a logit scale. Student ability measures are on the left, where the size of each bar corresponds to the number of participants with a given ability measure. Item difficulty measures are on the right.208

Figure 5.4. Cluster solution of Rasch ability measures and average confidence ratings.214

List of Schemes

Scheme 3.1. Empirical insights into developing expertise in ^1H NMR spectral interpretation95

Scheme 3.2. Compounds identified as targets of synthesis in each interpretation task.
..... 100

List of Tables

Table 1.1. Research studies investigating the teaching and learning of spectroscopic structural elucidation, including a description of study population, sample size, and the focus of investigation.	6
Table 2.1. Invalid chemical assumptions that constrained students' reasoning and themes among these assumptions. n-values in the third column correspond to the number of participants (n=18) who used specific assumptions at least once. n-values under each theme correspond to the number of participants with responses contributing to each theme.....	59
Table 2.2. Heuristics identified in at least 20% of RTA interview transcripts, corresponding definitions, the number of participants who used corresponding heuristics at least once (n = 18), and total number of responses containing each heuristic (N = 54).	66
Table 2.3. P-values and odds ratios corresponding to one-sided Fisher's exact tests for assumptions and heuristics associated with incorrect responses.	75
Table 2.4. Invalid chemical assumption themes, the number of participants who used corresponding assumptions at least once (n=18), and frequencies of responses containing corresponding assumptions.	84
Table 2.5. Heuristics identified in at least 20% of RTA interview transcripts, the number of participants who used corresponding heuristics at least once (n=18), and frequencies of responses containing corresponding heuristics.	85

Table 3.1. Number of undergraduate and doctoral participants with assumptions contributing to corresponding areas and levels..... 102

Table 3.2. Mann-Whitney U test comparisons of fixation counts between undergraduate (n=18) and doctoral (n=7) participants. 104

Table 3.3. Exemplar foundational assumptions, including levels of sophistication and the number of undergraduate and doctoral participants holding assumptions from given levels. 120

Table 3.4. Exemplar assumptions about deviations from the N+1 rule, including levels of sophistication and the number of undergraduate and doctoral participants holding assumptions from given levels. 120

Table 3.5. Exemplar assumptions about the variability of spectral data, including levels of sophistication and the number of undergraduate and doctoral participants holding assumptions from given levels. 121

Table 3.6. Exemplar visuospatial assumptions, including levels of sophistication and the number of undergraduate and doctoral participants holding assumptions from given levels. 121

Table 3.7. Exemplar practical assumptions, including levels of sophistication and the number of undergraduate and doctoral participants holding assumptions from given levels. 122

Table 3.8. Areas of interest in each synthesis and corresponding letter codes used to report Mann-Whitney U test comparisons in Tables 3.9-3.11. 124

Table 3.9. Mann-Whitney U test comparisons of transition counts between undergraduate (n=18) and doctoral (n=7) participants for Synthesis 1, including the U statistic, p-value, and effect size for significant transitions..... 125

Table 3.10. Mann-Whitney U test comparisons of transition counts between undergraduate (n=18) and doctoral (n=7) participants for Synthesis 2, including the U statistic, p-value, and effect size for significant transitions.....	126
Table 3.11. Mann-Whitney U test comparisons of transition counts between undergraduate (n=18) and doctoral (n=7) participants for Synthesis 3, including the U statistic, p-value, and effect size for significant transitions.....	127
Table 4.1. Questionnaire Blueprint.	154
Table 4.2. PCK scoring rubrics.	155
Table 4.3. Pearson correlations between CK, PCK, and TA characteristics. * indicates p (two-tailed) < 0.05, ** indicates p (two-tailed) < 0.01.....	160
Table 4.4. Frequencies of PCK components incorporated into “developing” or “exemplary” ‘what makes a topic difficult’ responses	162
Table 4.5. PCKQ1 exemplars: What aspects of NMR spectral interpretation are most challenging for introductory organic chemistry students?	170
Table 4.6. PCKQ2a exemplars: A student brings the ¹ H NMR spectrum and chemical shift table depicted below to your office hours. The student says that he or she is having trouble figuring out the corresponding molecular structure. Identify and describe particular features (if any) in the ¹ H NMR spectrum that may be creating difficulty for the student.	171
Table 4.7. PCKQ2b exemplars: As though you are talking with the student, explain how he or she should go about interpreting the spectrum below.....	172

Table 4.8. PCKQ3a exemplars: You are grading problem #1 on a student’s quiz. For problem #1, was this student able to provide the correct answer? What (if any) features contributed to difficulty?.....	173
Table 4.9. PCKQ3b exemplars: If incorrect, how would you help this student correctly interpret this spectrum?	174
Table 4.10. PCKQ4a exemplars: You are now grading problem #2 on a student’s quiz. For problem #2, was this student able to provide the correct answer? What (if any) features contributed to difficulty?.....	174
Table 4.11. PCKQ4b exemplars: If incorrect, how would you help this student determine the correct answer?	175
Table 4.12. Frequency of reported experiences contributing to TAs’ PCK responses	175
Table 5.1. Ten NMR-LRC items, including terminology in each multiple-choice question (in italics) and a question summary. Each question summary reflects the final question format following changes to ensure response process validity.	198
Table 5.2. Descriptive statistics of participants’ total scores and average confidence ratings.....	203
Table 5.3. Psychometric estimates of NMR-LRC items from Rasch analysis, including infit and outfit statistics and item difficulty measures. Items are ordered from easiest to most difficulty.....	207
Table 5.4. Description of clusters, including mean Rasch ability measure, median total score, mean average confidence rating, and cluster size. All values are from one of the seven five-cluster solutions. This solution was selected randomly, and all values across solutions were highly similar.....	215

Table 5.5. Principal component analysis factor loadings of NMR-LRC items on the second contrast.....222

Abstract

The present state of modern organic chemistry is in part due to the advent of nuclear magnetic resonance (NMR) spectroscopy. This analytical technique has transformed the field from one that previously relied on macroscopic properties when determining chemical identity to one that does so almost exclusively through the spectroscopic characterization of molecular structure. In addition to its transformative role, NMR spectroscopy is essential for future advancements in the field. Though its disciplinary value is undeniable, NMR spectroscopy is unfortunately difficult to both teach and learn. This difficulty is compounded by a general lack of chemistry education research on teaching and learning this technique, which in turn results in a lack of evidence-based instruction that cultivates relevant expertise. The work presented herein constitutes some of the first research on teaching and learning NMR spectroscopy. Specifically, this research focuses on ^1H NMR spectroscopy, an application of the technique widely used by organic chemists. This work, guided by a number of research questions, provides insight that will serve to transform undergraduate and graduate-level instruction to effectively foster expertise in this practice. Among these questions were: (1) How do undergraduate and doctoral chemistry students develop expertise in ^1H NMR spectral interpretation; (2) What knowledge do teaching assistants have for teaching ^1H NMR spectroscopy, and how does this knowledge develop; and (3) Can we develop an assessment in ^1H NMR spectroscopy that supports undergraduate instruction? Studies to investigate these questions drew from several theoretical and conceptual frameworks from the fields of education and cognitive psychology. Further, these studies incorporated a range of data collection methods, including eye tracking, interviewing, and surveys. Eye-tracking data were analyzed quantitatively, and interview and survey data were analyzed using a combination of qualitative and quantitative methods. Findings from the studies on developing expertise suggest that multiple areas of understanding are

necessary for interpreting ^1H NMR spectra, where progress in understanding corresponds to increasing knowledge of experimental and implicit chemical variables. More sophisticated understanding then facilitates an efficient and selective focus on features relevant for decision making. Further, less sophisticated understanding involving the overgeneralization of learned principles presents a significant barrier to learning when this inaccurate knowledge is used in combination with rule-based, shortcut reasoning strategies. These findings suggest that instruction should be designed to cultivate understanding across the identified areas, to promote the selective processing of relevant information, and to shift novices from rule-based, shortcut reasoning to analytical thinking that involves evaluating multiple underlying variables. Findings from the studies on instructors' knowledge and assessment development provide instructors with additional guidance for providing effective instruction. In our investigation, teaching assistants appeared to lack knowledge of evidence-based instructional strategies to shift undergraduates to analytical thinking, indicating that this population of instructors may require targeted training to adequately support undergraduates' learning. Further, the successful development of an assessment that measures students' ability to communicate about ^1H NMR spectra provides undergraduate instructors with a tool they can use to give feedback that promotes learning. Insight afforded by this body of foundational research has considerable, specific implications for reforming instruction on ^1H NMR spectroscopy, all of which are summarized in detail in the closing remarks.

Chapter 1

Introduction

1.1 Document overview

This document contains an ensemble of articles constituting the author's published and 'in preparation' body of work. The introductory chapter presents the larger body of chemistry education research on spectroscopic analysis in organic synthesis, with a focus on ^1H NMR spectroscopy as a tool for characterizing molecular structure. This chapter highlights the author's contributions to this larger body of work, providing a full scope of their research undertaken to meet requirements of the doctorate. The chapter was originally published by the *Royal Society of Chemistry* in the book *Problems and Problem Solving in Chemistry*, though it has been modified to incorporate the author's more recent research. Original publication and copyright information are provided below. Subsequent chapters correspond to the author's published and 'in preparation' research articles on studies investigating the teaching and learning of ^1H NMR spectroscopy. Each subsequent chapter contains initial remarks highlighting the significance of the study, relevant findings, insights into teaching and learning ^1H NMR spectroscopy provided by the findings, original publication and copyright information, and contributions by coauthors. These remarks also integrate findings from each study with those of the author's previous studies, providing a cohesive narrative of insights afforded by the author's research. Chapters 2 and 3 correspond to studies on learning ^1H NMR spectroscopy, and Chapters 4 and 5 correspond to studies on teaching ^1H NMR spectroscopy. The concluding chapter contains closing remarks.

Original publication and copyright information

Reproduced from “Connor, M. C.; Shultz, G. V. Problem Solving Using NMR and IR Spectroscopy for Structural Characterization in Organic Chemistry. In *Problems and Problem Solving in Chemistry*; Tsapalis, G., Ed.; Royal Society of Chemistry, 2021” with permission from the Royal Society of Chemistry.

1.2 Abstract

Determination of chemical identity is a fundamental chemistry practice that previously relied on the analysis of properties such as chemical composition, solubility, and reactivity. This practice now depends almost exclusively on the characterization of molecular structure through spectroscopic analysis. This analysis is a day-to-day task of organic chemists, and instruction in modern organic chemistry aims to cultivate such ability. This chapter provides an overview of literature on teaching and learning spectroscopic structure elucidation, with the aim of presenting the current state of research, empirical insights into teaching and learning this practice, and trends in instructional innovations. A number of studies have investigated reasoning and problem-solving approaches used to evaluate spectroscopic data for organic structural determination, and they provide a foundation for understanding how this ability develops and how instruction may facilitate such learning. These studies consistently suggest that individuals with different amounts of prior coursework and research experience interpret spectra differently. Specifically, individuals with less coursework and experience often possess limited chemical knowledge and restrict their decision making to fewer observations while ignoring relevant spectral information. Practice articles on spectroscopic structural elucidation focus on two general types of instructional innovations: instructional scaffolding and laboratory exercises. These articles notably incorporate few empirical insights, suggesting a gap between the substantial research conducted on the learning of spectroscopic structural elucidation and instructional innovation.

1.3 The role of spectroscopic analysis in organic synthesis

Prior to the 1950s, synthetic chemists established the chemical identity of organic compounds using properties such as elemental composition, melting and boiling point, visual characteristics such as crystal form and color, solubility in various solvents, and notable chemical reactivities.¹ Yet with advancements in instrumentation beginning in the 1950s and 60s, this characterization rapidly shifted to rely almost exclusively on molecular structure determination via spectroscopic analysis. Within the century, the average number of spectroscopic methods used for characterization increased from zero to 4.5. During this time, infrared (IR) spectroscopy, ¹H and ¹³C nuclear magnetic resonance (NMR) spectroscopy, and x-ray diffraction emerged as standard characterization methods that are still regularly used today.¹ This shift has allowed for significant advancements in synthetic chemistry, and spectroscopic structure elucidation is now a daily practice for synthetic chemists. A foundational, ethnographic study of synthetic chemists revealed that NMR spectra are among the most common visualizations used by these individuals in both academic and industrial settings, where they serve as both representations of submicroscopic entities and tools that support social discourse.² Further, a study of problem types in synthetic chemistry research revealed that spectroscopic analysis features prominently in day-to-day problems; synthetic chemists regularly evaluate the formation of byproducts and unwanted products via spectroscopic analysis, as well as analyze the purity of starting materials and reactants using this practice when unwanted products are obtained.³ They also utilize spectroscopy for mechanistic analysis, with NMR serving as common analytical tool in mechanistic studies.^{4,5} The disciplinary role and value of spectroscopic structure elucidation is thus remarkable, and instruction in modern organic chemistry aims to cultivate students' ability to engage in this essential practice, though with variable success. Research suggests that individuals eventually develop such ability, though this development is only accomplished in later years of graduate-level study and not within the undergraduate chemistry curriculum.⁶ Insight into the teaching and learning of spectroscopic structural elucidation is thus necessary to inform instruction and ensure that students complete this curriculum with requisite knowledge and skills. Such preparation will be particularly

important for individuals either pursuing graduate-level studies in organic chemistry or entering industry directly from their undergraduate institution. Moreover, this evidence-based instruction will also benefit individuals who pursue careers in fields other than organic chemistry, as the ability to analyze and interpret data is essential across science disciplines.

This chapter provides an overview of literature on teaching and learning organic structural determination via spectroscopic analysis, with the goals of highlighting the author's contributions to this larger body of research, as well as providing organic chemistry instructors with pedagogical insight and chemistry education researchers with the current state of research and directions for future study. The author sought to understand the teaching and learning of this practice from the perspective of both researchers and practitioners; the search for relevant literature was thus broad and included research and practice articles. The search was conducted using the key words *spectroscopy* and *organic chemistry* in journals most frequently used by the chemistry education community, including *Chemistry Education Research and Practice*, the *Journal of Chemical Education*, and the *Journal of Research in Science Teaching*. These search criteria resulted in nearly three thousand articles, a number which reflects the community's significant interest in the teaching and learning of this practice. The articles described herein were selected from this larger body of literature based on their potential to accomplish the two aforementioned goals. A general summary of the chapter's contents is provided in Figure 1.1.

1.4 Research investigating the teaching and learning of spectroscopic structure elucidation

Chemistry education researchers have allocated increased attention to the teaching and learning of spectroscopic structure elucidation over the past decade, with several studies investigating how ability in this practice develops and how it may be cultivated through instruction. Using a variety of frameworks and research methods, these studies predominantly investigated the chemical knowledge, reasoning, and problem-solving strategies involved in the interpretation of spectral data. Findings from these

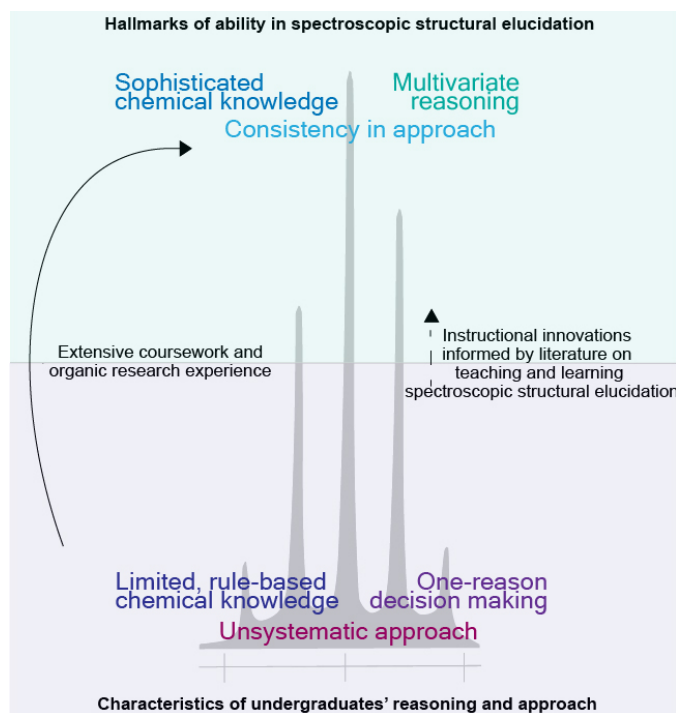


Figure 1.1. A summary of the chapter's contents, including hallmarks of ability and typical characteristics of undergraduates' reasoning and approach identified through literature on the teaching and learning spectroscopic structural elucidation, as well as potential avenues for developing this ability.

investigations collectively suggest that the development of such ability occurs through experience interpreting data and corresponds to increasingly sophisticated chemical knowledge, multivariate reasoning, and consistency in problem-solving approach. These studies also provide insight into how instruction and curricular materials may be designed to expedite this development and support learning. Studies investigating the teaching and learning of spectroscopic structure elucidation have used a range of theoretical and conceptual frameworks, data collection and analysis methods, and study populations (Table 1.1). Several trends emerged from these investigative approaches and are described herein to provide a comprehensive overview of the current state of research on the teaching and learning of this practice.

1.4.1 NMR and IR spectroscopy as the focus of research

NMR and IR spectroscopy serve as the focus of investigations on the teaching and learning of spectroscopic structure elucidation. The power of NMR spectroscopy lies in its potential for determining chemical identity through the nearly complete elucidation of

Table 1.1. Research studies investigating the teaching and learning of spectroscopic structural elucidation, including a description of study population, sample size, and the focus of investigation.

Studies on learning	Study population and sample size (N)	Investigative focus
<i>Problem-solving perspective</i>		
Topczewski <i>et al.</i> (2017) ⁷	Undergraduates (N=15) Graduate students (N=12)	Differences in approach for matching correct structure to ¹ H NMR spectrum (expert vs. novice participants)
Domin and Bodner (2012) ⁸	Graduate students (N=15)	Differences in external and internal representations constructed for 2D-NMR conceptual questions (successful vs. unsuccessful participants)
Cartrette and Bodner (2010) ⁹	Graduate students (N=13) Faculty (N=2)	Differences in knowledge and approach for complete structural elucidation using IR and ¹ H NMR spectra (successful vs. unsuccessful participants)
<i>Reasoning perspective</i>		
Wright and Oliver-Hoyo (2020) ⁹	Undergraduates (N=29)	Assumptions and mental models used when explaining how molecular structure results in IR peaks, including how they progress in sophistication
Connor <i>et al.</i> (2020) ¹⁰	Undergraduates (N=18) Graduate students (N=7)	Assumptions and cognitive processes when evaluating success of syntheses using IR and ¹ H NMR spectra
Connor <i>et al.</i> (2019) ¹¹	Undergraduates (N=18)	Assumptions and heuristics that constrain reasoning when evaluating success of syntheses using IR and ¹ H NMR spectra
Stowe and Cooper (2019) ¹²	Undergraduates (N=300)	Effect of scaffolding for complete structural elucidation using IR, ¹ H NMR, and ¹³ C NMR spectra
Cullipher and Sevan (2015) ¹³	Undergraduates (N=20) Graduate students (N=6)	Assumptions used when explaining how molecular structure results in IR peaks, including how assumptions progress in sophistication
Studies on teaching		
Connor <i>et al.</i> (<i>in preparation</i>) ¹⁴	Undergraduates (N=678)	Development of instrument to measure lexical ability in ¹ H NMR spectroscopy
Anderson <i>et al.</i> (2020) ¹⁵	n/a	Textbooks' potential for scaffolding interpretation of ¹ H NMR spectra
Connor and Shultz (2018) ¹⁶	Undergraduates (N=2) Graduate students (N=16) Postdoctoral fellows (N=2)	Teaching assistants' knowledge for teaching ¹ H NMR

organic molecular structure, and these spectra serve as some of the most common visualizations used by synthetic chemists.² Development of the ability to interpret NMR spectra therefore clearly merits investigation. The benefits of investigating the teaching and learning of IR spectroscopy, however, are more nuanced. IR spectroscopy previously served as a routine step when determining chemical identity during organic synthesis, though this technique is becoming increasingly obsolete as more powerful analytical tools, like high resolution mass spectrometry, emerge as commonplace. Many synthetic journals still require characterization via IR spectroscopy for publication, yet the technique no longer affords significant structural insight. Though while IR spectroscopy is no longer an essential tool to many organic chemists, investigating the learning of this technique affords significant insight into students' conceptual understanding of structure-property relationships and light-matter interactions, fundamental aspects underlying all modes of spectroscopic analysis. Further, it affords insight into students' reasoning during the analysis, interpretation, and argumentation from spectral data, as well as the triangulation of data across multiple representations when included with other spectral data, an essential component of chemists' representational competence.¹⁷ These findings are then partially transferable to other structural characterization techniques beyond spectroscopy (e.g., mass spectrometry).

The majority of studies have investigated the learning of spectroscopic structure elucidation, with three studies focusing on the teaching of this practice (Table 1.1). The author has contributed two studies on learning and two studies on teaching to this body of work (see studies authored by *Connor et al.* in Table 1.1). Among studies investigating the learning of this practice almost all involved the interpretation of spectra, though one involved structured, short-response conceptual questions about underlying physical principles.⁸ Further, nearly half of these studies combined multiple forms of spectral data in their investigations, providing insight into the conceptual understanding, reasoning abilities, and problem-solving approaches requisite to analysis, interpretation, and triangulation. Three of these investigations involved participants completing interpretation tasks that combined IR and ¹H NMR spectra; a foundational study involved participants' complete structural elucidation using provided spectra and a molecular formula, and two

studies by the author involved participants' confirmation of a synthetic product using authentic spectra.^{6,10,11} A fourth investigation involved participants' complete structural elucidation using molecular weight with IR, ¹³C NMR, and ¹H NMR spectra.¹² Further, two studies focused exclusively on the interpretation of IR spectra, with each requiring participants to explain how differences in molecular structure affect the appearance of corresponding spectra.^{9,13} Lastly, one study focused exclusively on the interpretation of ¹H NMR spectra and involved participants matching one of several structural formulae to a provided spectrum,⁷ while another study focused on advanced physical principles underlying two-dimensional NMR spectroscopy (2D-NMR) techniques and involved participants answering structured questions that probed their conceptual understanding.⁸ For studies focused on teaching, one study by the author investigated how teaching assistants develop knowledge for teaching ¹H NMR, another study by the author involved developing a formative assessment to measure lexical ability in ¹H NMR spectroscopy, and a final examined textbooks' potential to scaffold learning of ¹H NMR spectral interpretation.^{14–16}

Notably, the majority of studies on learning involved the interpretation of spectra in contexts that do not mirror their authentic use, as chemists typically no longer perform complete structural elucidation using spectra combined with only a molecular formula or weight. Nor do they match one of many molecules to a given spectrum. While these studies provide a useful foundation for understanding how the ability to engage in this practice develops, the degree to which findings transfer to more authentic contexts is thus uncertain.

1.4.2 Theoretical and conceptual frameworks guiding research on learning

Researchers used a variety of theoretical and conceptual frameworks to investigate the learning of this practice. Together, their approaches reflect two epistemological perspectives on what it means to cultivate relevant ability. Several studies have adopted a problem-solving perspective that aligns with a body of research in fields such as cognitive psychology and mathematics education (Table 1.1). This perspective focuses on the cognitive processing strategies and behaviors involved in problem solving and how

these change with increasing expertise.⁶⁻⁸ Studies adopting this perspective refer to spectral interpretation as a “problem-solving task.” This perspective is largely characteristic of earlier research studies. More recently, researchers shifted from viewing spectroscopic structure elucidation as a problem-solving task to considering it as a scientific practice that involves the analysis and interpretation of data, argumentation from evidence, and evaluation of structure-property relationships (Table 1.1).^{18,19} Studies adopting this second perspective focus on participants’ reasoning during spectral interpretation and draw upon conceptual lenses such as chemical thinking and argumentation to guide their investigations.^{12,13} Notably, the author has completed one study in which data is analyzed through the lens of both perspectives, demonstrating that these perspectives can be integrated to provide a more comprehensive understanding of learning.¹⁰

Studies adopting the problem-solving perspective are notably all concerned with identifying differences between either experts and novices or successful and unsuccessful problem-solvers. These studies predominantly investigated the strategies participants use to interpret spectral data, in particular the order and frequency with which they evaluate different aspects of spectra and the consistency of their approach.^{6,7} A subset of these studies also investigated the role of conceptual understanding involved in problem-solving; one study by Cartrette and Bodner focused in part on individuals’ declarative and procedural knowledge relating to coupling constants and the N+1 rule. Another by Domin and Bodner investigated participants’ construction of internal and external representations.^{6,8}

Conversely, studies adopting the reasoning perspective are not concerned with identifying differences between experts and novices or successful and unsuccessful problem-solvers. Instead, they focused on how participants’ reasoning during spectral interpretation progressed in sophistication. These investigations often drew upon the chemical thinking framework, a proposed learning progression that describes how thinking involved in the analysis, synthesis, and transformation of matter changes through secondary and post-secondary education.¹⁹ The framework identifies six crosscutting disciplinary concepts essential to such thinking, two of which directly underly

spectroscopic analysis: chemical identity and structure-property relationships. Spectroscopic analysis is an essential means of determining chemical identity, and molecular structure is a direct cause of spectroscopic properties. In this framework, conceptual sophistication and modes of reasoning are two essential progress variables involved in the development and application of chemical knowledge and practices. Progress variables refer to dimensions of understanding or ability along which individuals develop (i.e. progress).²⁰ Conceptual sophistication is reflected by individuals' underlying assumptions about the nature and behaviour of chemical entities and phenomena, and modes of reasoning encompass the complexity of thinking as it relates to the connection of ideas, decision making, and construction of explanations.¹⁹ Mapping the assumptions that guide individuals' thinking serves as a means of measuring progress in understanding, and characterizing changes in modes of reasoning serves as a means of assessing progress in the complexity of thinking. According to this framework, progress in the complexity of thinking corresponds to movement from the consideration of individual, explicit variables when explaining phenomena to the consideration of multiple implicit variables and their emergent properties.

Studies drawing upon the chemical thinking framework predominantly investigated the chemical assumptions that guide participants' reasoning during the interpretation of IR and ¹H NMR spectra.^{9,11,13} These assumptions relate to the nature and behaviour of molecular and spectral features, as well as the relationship between molecular structure and spectroscopic properties. One study also investigated participants' modes of reasoning by characterizing their use of heuristics when determining the success of syntheses using spectral data.¹¹ Further, these studies often incorporate additional cognitive theories into their respective frameworks while also drawing upon chemical thinking, thus contributing to the field's growing understanding of chemistry students' cognition. Among these are theories relating to mental models, dual processing, and categorization.^{9,11} Moreover, one study adopting the reasoning perspective did not draw from the chemical thinking framework, but instead investigated participants' thinking using a lens of argumentation and a resources-based view of procedural knowledge.¹²

1.4.3 Methodological trends: study populations, data collection, and analysis

Investigations incorporated a range of study populations that represent different levels of knowledge and ability in spectroscopic analysis, including undergraduate students from introductory and upper-level chemistry courses, graduate students, and faculty members (Table 1.1). Both studies adopting the problem-solving perspective and those adopting the reasoning perspective included individuals from multiple levels as study participants, though with different objectives for their inclusion. Those adopting the problem-solving perspective did so to identify differences between experts and novices or successful and unsuccessful problem-solvers. Problem-solving studies that chose to categorize their participants as successful or unsuccessful rather than experts or novices argued that this categorization is more appropriate given that experts may not necessarily be successful at solving problems.⁶ Those adopting the reasoning perspective included individuals from different educational and experience levels to investigate how conceptual understanding and reasoning abilities evolve as individuals move toward sophisticated thinking.^{6,7,13} Some of these studies also focused exclusively on undergraduate populations, in particular how these students interpret spectral data following either scaffolding or instructional intervention.^{9,12} Sample size remained relatively consistent across perspectives, with most studies including 15 to 29 participants (Table 1.1). One study aligning with the reasoning perspective constituted an exception to this trend, with 300 total participants.

In addition to study populations exhibiting a range of experience, investigations also adopted qualitative, quantitative, and mixed-method designs. The interpretation of spectral data is an inherently visual task, with NMR spectra serving as one of the most common visualizations used by synthetic chemists.² The ability to engage in spectroscopic structure elucidation thus also corresponds to the ability to comprehend related visualizations.² The nature of this task therefore necessitates methods that facilitate the investigation of thinking involved in the comprehension of representations. As a result, nearly all studies involve participants' completion of spectral interpretation tasks that incorporate relevant representations (i.e., spectra), mainly during one-on-one interviews but also through written responses to prompts. Data collection during these

interviews has taken on a number of forms, including the exclusive use of think-aloud protocols or eye tracking for qualitative or quantitative data collection, respectively. Think-aloud interviews involve participants providing a verbal report of their thoughts as they complete tasks.²¹ They are used to collect qualitative data on participants' thinking, including their conceptual understanding and problem-solving approach, among others. Eye tracking is used to collect predominantly quantitative data on participants' cognitive processes, which can then provide insight into aspects such as problem-solving strategies. Mixed-method studies collecting both qualitative and quantitative data used eye tracking concurrently with think-aloud protocols, as well as eye tracking followed by cued retrospective think-aloud (RTA) protocols. Cued RTA protocols involve participants watching a recording of their eye movements following the completion of an interpretation task and narrating in as much detail as possible what they were looking at and thinking about.²² The use of both qualitative and quantitative methods facilitates the triangulation of this data, providing rich and reliable insight into participants' thinking. Across these different study designs, eye-tracking data are often quantitatively analyzed using a variety of statistics, and interview data are qualitatively coded for participants' assumptions, heuristics, mental models, and problem-solving behaviors, among others.

Eye tracking is particularly well-suited for investigating the interpretation of spectra, as it is a leading research tool for investigating cognitive processes involved in the comprehension of visualizations in related fields such as cognitive psychology. Eye tracking involves measuring participants' eye movements as they complete visual-based tasks, where eye movements serve as a measure of participants' visual attention and the visualization patterns they use to interpret the sum of a stimulus.^{23,24} These cognitive processes often occur subconsciously and cannot be verbalized, though they provide significant insight into information participants find important, complex, or confusing. Further, the interpretation of spectra is a cognitively demanding task, especially for undergraduates initially learning this practice; when used independently or in tandem with cued RTA interviewing, eye tracking allows participants to complete tasks in silence as opposed to continuously thinking aloud, thus reducing their cognitive load. Both studies adopting the problem-solving perspective and reasoning perspective have collected data

via eye tracking. Problem-solving studies used this tool to investigate differences in visual attention and visualization patterns between experts and novices, and reasoning studies used eye tracking to understand how individuals with different levels of conceptual sophistication view spectral data differently.^{7,10,24}

Eye tracking affords certain benefits over standard think-aloud protocols, though it is not without limitations. Foremost, when used independently to collect exclusively quantitative data, significant inference is required on behalf of the researcher during subsequent analysis. For instance, participants may fixate on a spectral feature either because they find the information important or confusing. Without qualitative data to triangulate this quantitative data, the researcher is unable to reliably ascribe meaning to these eye movements. This limitation is mitigated to an extent with the use of cued RTA and concurrent think-aloud interviewing, though the benefit of allowing participants to work in silence is lost with the concurrent approach. Further, eye tracking generates a substantial amount of data, which makes sample size and the duration of data collection important considerations during study design. The sample size of studies utilizing eye tracking to investigate cognitive processes involved in the interpretation of spectra has ranged from just 13 to 27. This relatively small sample size makes statistical comparisons of eye movements between individuals with different levels of ability a challenge. To add to this difficulty, considerable differences in ability are often needed to result in measurable differences in eye movements.²⁵ The duration of data collection can also easily influence the size and thus manageability of a dataset, with even basic eye trackers measuring up to 60 datapoints per second. Imposing a time limit on task completion is one strategy for circumventing this issue, however the interpretation of spectra is a complex task that typically involves significant time. As a result, the majority of studies using eye tracking to investigate this practice did not impose such limits.^{10,11,13}

1.5 Empirical insights into teaching and learning spectroscopic structure elucidation

Studies adopting the problem-solving perspective have generated distinct yet equally informative findings when compared to studies adopting the reasoning perspective.

Results from both perspectives also complement one another to a degree and thus provide substantial insight into how instruction and curricular materials may be designed to support learning. By integrating these findings, instruction and curricular materials have the potential to support multiple aspects of cognition involved in spectroscopic structure elucidation. In addition, empirical insights into teaching this practice will facilitate the design of instructor education materials and further ensure instructional quality, and thus positive learning outcomes.

1.5.1 Insights into problem-solving aspects of spectroscopic structure elucidation

For studies adopting the first perspective, a number of findings relate to problem-solving behaviors involved in interpretation of IR and ^1H NMR spectra. These findings collectively suggest that a consistent, systematic approach serves as an important predictor for successful, expert-like problem-solving. In a foundational study of problem-solving approaches used by graduate students and faculty members to elucidate molecular structure using provided spectra, Cartrette and Bodner found that those who provided a correct structure were more likely to use the same steps during interpretation.⁶ Specifically, these participants were more likely to first determine the degree of unsaturation using a provided molecular formula, second identify functional groups using the provided IR spectrum, third determine connectivity using the provided ^1H NMR spectrum, and finally check the consistency of their proposed structure with the provided spectra. Findings from an eye-tracking study by Topczewski *et al.* further suggest that a consistent, systematic approach is requisite to expert-like problem-solving.⁷ This study investigated differences in problem-solving behaviors between undergraduates and chemistry graduate students, who it categorized as novices and experts, respectively. For the study, participants matched one of several molecules to a provided ^1H NMR spectrum while having their eye movements tracked. Results indicated that experts transitioned more between the correct molecular structure and ^1H NMR peaks compared to novices. In addition, a subset of novices transitioned more between provided molecular structures compared to experts and other novices, who instead transitioned between molecular structures and peaks. This result suggests that a subset of novices adopted an alternate strategy that involved searching amongst provided structures. The differential use of

strategies by novices, combined with experts' consistent transitioning between the correct structure and spectrum, suggests that experts collectively utilized a more efficient, systematic approach. An eye-tracking study by the author further supports these findings, as doctoral participants in this study engaged in informed, unidirectional processing of relevant information whereas undergraduates engage in the uniformed bidirectional processing of all information.¹⁰ Findings from a study by Domin and Bodner on graduate students' learning of advanced physical principles underlying 2D-NMR also point toward the importance of a systematic approach.⁸ In this study, individuals more successful at answering conceptual questions regularly constructed external representations that both resembled the representation provided in the instructional material and included all relevant information, whereas the less successful generated representations that incorporated prior knowledge and only partially resembled instructional materials.

In addition to a consistent and efficient approach, these studies demonstrated that individuals' ability to integrate information and the nature of their chemical knowledge are also important predictors of success. For instance, Cartrette and Bodner found a positive and significant correlation between the correct determination of molecular structure and the correct determination and usage of degrees of unsaturation, IR signal identification, coupling constants, multiplicity, chemical shift, and integration of ¹H NMR signals.⁶ They also found that successful participants demonstrated more procedural knowledge, particularly knowledge of how to use the N+1 rule and coupling constants. Less successful participants, on the other hand, demonstrated declarative knowledge of these aspects but were unable to flexibly and accurately use them during interpretation. These participants also tended to fixate on only one or two pieces of data, which the authors suggest may be due to their limited procedural knowledge. Further, the author's study demonstrated that graduate students possessed more sophisticated chemical understanding compared to undergraduate participants, and this more sophisticated understanding allowed them to more selectively focus their visual attention.¹⁰ Lastly, Domin and Bodner's finding that successful problem-solvers tend to construct representations that closely match instructional materials suggests that the organization of one's knowledge also predicts success.⁸

1.5.2 Insights into reasoning aspects of spectroscopic structure elucidation

For studies adopting the reasoning perspective, findings provide insight into how conceptual understanding relating to molecular features, spectral features, and the relationship between molecular structure and spectroscopic properties evolves as one moves toward sophisticated thinking. Collectively, these studies suggest that lower levels of sophistication are characterized by limited knowledge of the nature and behaviour of implicit molecular features (e.g., dynamic bonding interactions, dipole moments, symmetry, and proton exchange), the rigid view that spectral data should be absolute and invariable, and rule-based generalizations about explicit molecular features (e.g., the presence of certain atoms or functional groups) and their spectroscopic properties. Higher levels of sophistication, on the other hand, are characterized by sophisticated knowledge of the nature and behaviour of implicit molecular features, a flexible view of spectral data's variability, and a focus on implicit molecular features and their spectroscopic properties. Further, these studies also provide insight into modes of reasoning involved in judgement and decision-making inherent to spectroscopic analysis. Ultimately, findings suggest that lower levels of reasoning are characterized by decision making based on individual spectral features, whereas higher levels of reasoning are characterized by the consideration of multiple variables. Like problem-solving studies, these investigations focus primarily on the interpretation of IR and ^1H NMR spectra.

Findings from studies that focus exclusively on the interpretation of IR spectra provide significant insight into the evolution of conceptual sophistication. Overall, they suggest that such evolution corresponds to a shift from reasoning focused exclusively on static, explicit molecular features as the cause of spectral peaks to reasoning focused on dynamic, implicit molecular features. Explicit molecular features relate to directly observable information conveyed by structural formulae (e.g., the presence of a certain atom), whereas implicit molecular features are not represented symbolically and instead must be inferred (e.g., electronegativity differences between atoms).²⁶ Further, static features are motionless (e.g., the presence of a bond), whereas dynamic features involve motion (e.g., the stretching of a bond). This shift in reasoning is also accompanied by movement from the inaccurate view that spectral peaks result from energy within a given

molecule to the accurate view that spectroscopy depends upon external energy being absorbed by the molecule. For instance, one investigation by Cullipher and Seviau characterized the assumptions guiding undergraduate and graduate students' reasoning as they explained how provided molecular structures resulted in observed peaks on given IR spectra.¹³ Findings from this study suggest that the lowest level of conceptual sophistication corresponded to assumptions that the presence of certain atoms result in IR peaks, followed by assumptions that the presence of certain bonds and functional groups result in their appearance. These assumptions appeared mainly in the responses of undergraduate participants. The highest level of sophistication, reflected predominantly in the responses of graduate students, corresponded to assumptions that implicit molecular features such as potential vibrational motions, combined with the absorption of energy, result in the appearance of IR peaks. The authors describe this shift as movement from an 'atoms-as-components' and 'bonds-as-components' view of molecular structure to a 'bonding' perspective that reflects more sophisticated content knowledge of the dynamic nature of bonds.

Findings from another study by Wright and Oliver-Hoyo support and contribute to these findings.⁹ Through an analysis of the assumptions and mental models guiding undergraduates' thinking during the interpretation of IR spectra, this study also found that participants' assumptions reflected either an 'atoms-as-components,' 'bonds-as-components,' or 'bonding' perspective. In addition, participants used a number of mental models that further illustrate how conceptual understanding relating to structure-property relationships may progress. Less sophisticated mental models identified through this analysis reflected inaccurate beliefs that the provided molecules were a collection of static bonds, and that IR peaks resulted from energy contained within these bonds. More sophisticated mental models were characterized by the accurate views that bending and stretching of these bonds may result in changes in dipole moment depending on the symmetry of the molecule, and that the absorption of energy following such changes results in the appearance of IR peaks.

Findings from one study by the author (i.e., Connor *et al.*) provide additional insight into individuals' conceptual understanding relating to molecular features, spectral

features, and the relationship between molecular structure and spectroscopic properties, in particular understanding which comprises lower levels of sophistication.¹¹ This investigation found that five general categories of inaccurate assumptions in part guided undergraduates' reasoning during the interpretation of IR and ¹H NMR spectra. All categories identified in this investigation related to the relationship between molecular structure and spectroscopic properties to a degree, with four of these categories relating somewhat more to spectral features and one relating somewhat more to molecular features. A subsequent study by the author demonstrated that these five general categories guide both undergraduate and graduate students' reasoning during the interpretation of such spectra, where increasing understanding across categories corresponded to increasing knowledge of experimental and implicit chemical variables.¹⁰

The four categories relating to spectral features contained assumptions reflecting either a misunderstanding or lack of familiarity with (1) acceptable variability in spectral data, (2) acceptable contexts for applying the N+1 rule, (3) peak characteristics beyond absorption frequency and splitting and (4) fundamental principles underlying all spectroscopic data. Inaccurate assumptions about acceptable variability in spectral data included ideas that peaks should not exhibit variation in factors such as absorption frequency, signal intensity, and resolution. For instance, a number of participants in this investigation reasoned that IR peaks should be prominent and non-overlapping if a spectrum is to correspond to a molecule containing a particular functional group. These inaccurate assumptions reflected limited knowledge of implicit molecular features that influence peak intensity, combined with limited knowledge of the variable nature of spectroscopic data. Assumptions reflecting a misunderstanding of acceptable contexts for applying the N+1 rule included the belief that the this "rule" should apply when determining the multiplicity of signals for which it is not applicable. For instance, several participants in this investigation reasoned that protons attached to heteroatoms like nitrogen and oxygen, as well as vinyl protons, should have corresponding signals with splitting in accordance with the N+1 rule (Figure 1.2). These assumptions reflected limited knowledge of implicit molecular features that affect splitting, such as proton exchange, as well as limited understanding of the N+1 rule's restricted application to specific systems.

Chemists conducted a series of reactions to synthesize N-(2-hydroxyethyl)-propanamide. The chemists then analyzed the final product spectroscopically to determine if the synthesis was successful. Based on the spectroscopic data of the final product below (IR and ^1H NMR spectra), did the chemists successfully synthesize N-(2-hydroxyethyl)-propanamide?

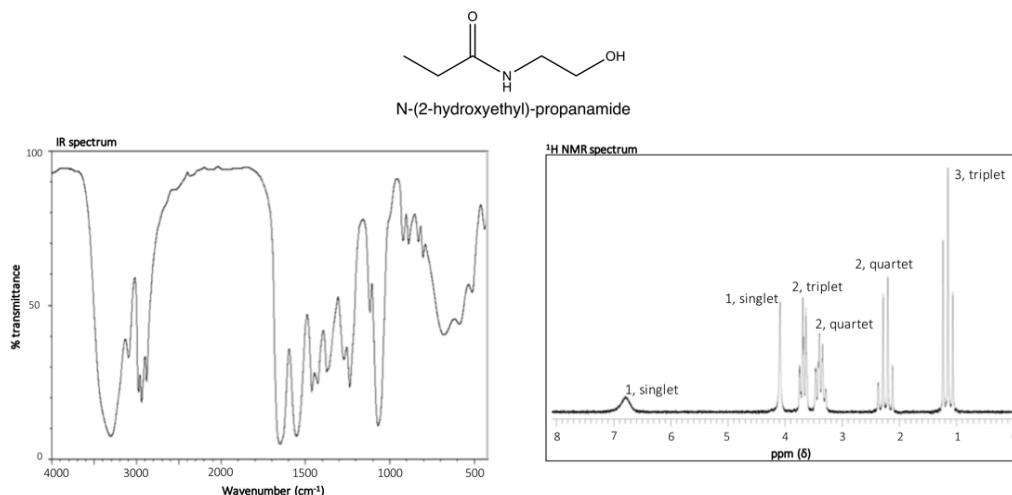


Figure 1.2. Spectral interpretation task in which undergraduates in the Connor et al. (2019) study assumed signals corresponding to hydrogen atoms on nitrogen and oxygen near 6.8 ppm and 4.2 ppm, respectively, should appear as triplets rather than singlets.

Participants' use of these assumptions, combined with their assumptions that spectral data should exhibit limited variability, suggests that lower levels of conceptual sophistication correspond to rule-based generalizations and limited knowledge of context. Use of these assumptions also suggests that simplifying the coupling phenomenon and then referring to this simplification as a "rule" rather than a "guideline" may serve to promote these lower levels of thinking among students.

Further, assumptions reflecting a lack of familiarity with additional peak characteristics typically involved participants misidentifying distinctive IR peaks as those corresponding to a functional group that absorbs electromagnetic radiation in a similar frequency region. For instance, a number of participants misidentified the intense, broad IR peak characteristic of a hydroxyl function group as a peak corresponding to a carbon-hydrogen bond (Figure 1.3). These assumptions reflect what Cullipher and Sevia refer to as the "bonds-as-components" perspective, which focuses on the presence of certain functional groups (i.e., an explicit molecular feature) as the cause of certain peak characteristics rather than implicit molecular features that would influence peak intensity, such as changes in dipole moment. Lastly, assumptions reflecting a misunderstanding of fundamental principles included ideas such as oxygen nuclei can generate signals in an ^1H NMR spectrum and de-shielding causes a shift right rather than left on an ^1H NMR

spectrum. These assumptions reflected limited knowledge of basic variables, either explicit or implicit, that influence spectral data.

The fifth category of assumptions related primarily to molecular features, in particular visuospatial aspects of structural formulae. Visuospatial thinking refers to the thought processes involved in the identification of spatial features such as the topicity of protons, in addition to the generation and recognition of structural formulae and other chemical symbols.²⁷ Structural formulae and NMR spectra are two of the most common visualizations used by chemists,² making visuospatial thinking an inherent aspect of spectroscopic structure elucidation. Yet, minimal research has focused on visuospatial thinking in this context. Visuospatial assumptions identified by the author related to participants limited ability to identify planes of symmetry and implicit hydrogen atoms within structural formulae. However, spectra also have spatial features which necessitate such thinking, such as discerning between two poorly resolved peaks. Additional research is therefore needed to better understand the nature of visuospatial thinking in this context and how related understanding may progress.

In addition to investigating assumptions, the author also characterized undergraduates' modes of reasoning, specifically their use of heuristic reasoning strategies, as they determined the success of syntheses using spectral data. Heuristic

Chemists conducted a series of reactions to synthesize 3-(allyloxy)propanal. The chemists then analyzed the final product spectroscopically to determine if the synthesis was successful. Based on the spectroscopic data of the final product below (IR and ¹H NMR spectra), did the chemists successfully synthesize 3-(allyloxy)propanal?

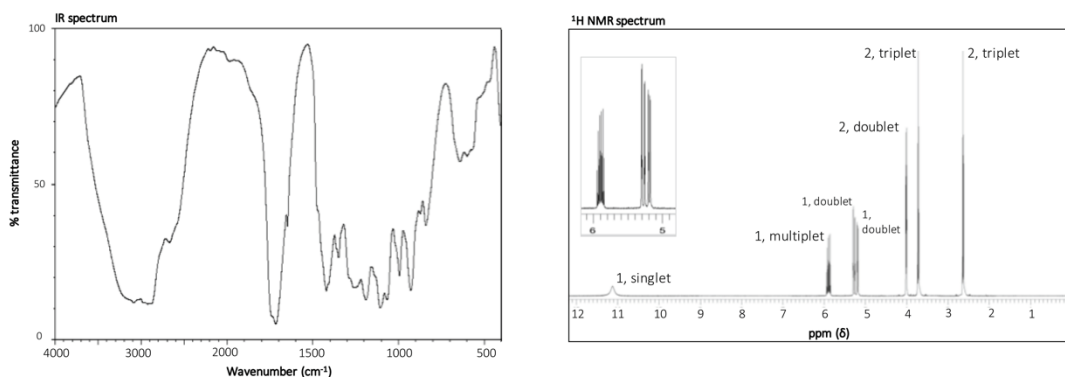
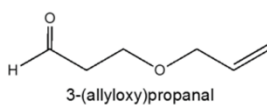


Figure 1.3. Spectral interpretation task in which undergraduates in the Connor et al. (2019) study assumed the broad IR peak characteristic of a hydroxyl group near 3000 cm⁻¹ corresponded to a carbon-hydrogen bond.

reasoning strategies are associated with Type 1 thinking in dual-process theory, a form of information processing that is fast, automatic, and independent of cognitive ability.²⁸ When used in combination with limited content knowledge, heuristics tend to result in biases in judgement and decision making.²⁹ The author found that participants used a number of heuristics as they determined the success of syntheses, and that the use of certain heuristics in combination with limited content knowledge resulted in constrained thinking and inaccurate decision making. Specifically, a number of participants who failed to determine the success of given syntheses held inaccurate assumptions that the N+1 rule should apply in contexts where it does not apply. These assumptions were a product of the generalization heuristic, which involves overextending learned rules without considering all relevant variables. Being rigid in their generalization, these participants then used individual spectral features that deviated from the N+1 rule as the entire basis of their inaccurate decision, an application of the one reason decision-making (ORDM) heuristic. Undergraduates' use of ORDM was also observed in the study by Wright and Oliver-Hoyo, where it appeared exclusively in the responses of participants with the less sophisticated mental model that molecules are a collection of static bonds.⁹ Participants with more sophisticated mental models instead engaged in multivariate reasoning, simultaneously considering features such as reduced mass and the abundance of certain bonds when explaining the appearance of IR spectra. Findings from these studies collectively suggest that inaccurate assumptions about the nature and behavior of molecular and spectral features often work in combination with ORDM during spectroscopic structure elucidation to constrain individuals' thinking.

Similar to the ORDM heuristic, participants in a study by Stowe and Cooper were found to be successful at evaluating IR, ¹³C, and ¹H NMR individually, though they struggled to propose a molecular structure consistent with all of this spectral data. As the chemical thinking framework suggests, progress in the complexity of thinking corresponds to a shift from the consideration of individual explicit variables when evaluating spectral data to the consideration of multiple implicit variables. Undergraduates' use of ORDM and their difficulty with integrating multiple pieces of spectral data in the above investigations thus aligns with lower-level modes of reasoning identified in this framework. The use of

multivariate reasoning by a subset of participants, however, suggests that higher levels of this progress variable can be attained through appropriate instruction.

1.5.3 General insights provided by problem-solving and reasoning perspectives

Studies investigating the learning of spectroscopic structure elucidation provide a number of general insights independent of epistemological perspective. Most notably, all studies that include participants from various educational and experience levels demonstrate differences in ability between individuals from these groups. For instance, findings from eye-tracking studies of both perspectives indicate that different groups view spectral data differently, suggesting that their cognitive processes involved in interpretation differ as well.^{7,13} Cartrette and Bodner also found that research experience and coursework were important predictors for successful interpretation.⁶ Clearly, the ability to interpret spectral data develops with education and experience in this practice. However, these studies further suggest that such ability develops at the graduate level and not within the undergraduate chemistry curriculum, underscoring need for the evidence-based design of instruction and curricular materials that effectively support learning.⁶ Findings from both perspectives also suggest that limited chemical knowledge results in an uninformed focus on both relevant and irrelevant information, followed by decision-making using just one to two pieces of spectral data.^{6,10,11} And then, even when all relevant information is evaluated, its integration poses as yet another challenging but essential step toward successful interpretation.^{6,12}

1.5.4 Insights into teaching spectroscopic structure elucidation

Classroom instruction plays a central role in students' learning of spectroscopic structure elucidation, suggesting that instruction quality serves as an important predictor of learning outcomes in this practice. Ensuring instruction quality is thus an essential aspect of supporting learning. The author's development of a formative assessment to measure students' lexical ability in ¹H NMR spectroscopy constitutes one important step toward improving instruction, as such a tool provides instructors with a means of providing feedback that promotes learning.¹⁴ Further, the author's empirical study of teaching

assistants' pedagogical content knowledge (PCK) in ^1H NMR spectroscopy provides some insight into how instructional quality can be improved to better support learning.¹⁶ PCK is a theoretical construct that entails instructors' knowledge for teaching particular content, including knowledge of the most useful and meaningful ways of representing subject matter in order to make it understandable to learners.³⁰ Instructors' PCK has been shown to positively correlate with both instruction quality and learning outcomes;^{31,32} improving instructors' PCK in spectroscopic structure elucidation thus serves as a means of improving instruction quality and learning outcomes in this practice.

The author found a positive and significant correlation between teaching assistants' PCK and their content knowledge in ^1H NMR, suggesting that the development of knowledge for teaching this technique requires related subject matter knowledge.¹⁶ This finding aligns with that of other investigations of PCK, which observed a similar correlation across a variety of topics.^{33,34} Moreover, this study also found that the development of PCK depends upon content knowledge required for specific ^1H NMR concepts (e.g., chemical equivalency) and problem types (e.g., determining if a spectrum corresponds to a structure); this finding differs from that of other PCK studies, which observed that knowledge for teaching is specific to topic rather than sub-topic or problem type.^{33,34} This difference may be attributed to the complex nature of ^1H NMR compared to introductory topics at the center of other investigations (e.g., chemical equilibria), as it is more advanced and requires both conceptual understanding and problem-solving skills.³⁵ This result ultimately suggests that instructor education materials for ^1H NMR should seek to cultivate knowledge for teaching specific ^1H NMR concepts and problem types in order to improve instruction quality. The design of these evidence-based materials will thus be an important step toward improving the teaching and learning of this practice.

Results from this study also indicated a positive and significant correlation between participants' teaching experience in ^1H NMR and their PCK. Notably however, PCK did not correlate with overall teaching experience, which included chemistry courses in which ^1H NMR is not covered. These findings suggest that knowledge for teaching ^1H NMR is cultivated specifically through experience teaching this practice and not through

instruction of tangentially related chemistry concepts. Teaching assistants and other instructors with prior teaching experience in ^1H NMR may therefore have greater knowledge for teaching this practice and be more capable of supporting learning. Further, this correlation underscores the need for evidence-based instructor education materials; instructors with significant teaching experience in ^1H NMR are not always available, meaning that novice instructors should have access to resources that will help them effectively support students' learning. This study was specific to ^1H NMR, so the degree to which these findings transfer to other spectroscopic techniques is therefore uncertain. Additional research is therefore needed to determine if these findings should inform instructor education across all spectroscopic techniques.

1.6 Instructional innovations

In addition to receiving attention in chemistry education research, a number of instructional innovations have been developed to support students' learning of spectroscopic structure elucidation. These innovations range from in-class, scaffolded spectral interpretation activities to virtual reality simulations involving data collection with an IR spectrometer.^{36,37} These resources can be grouped into two general categories: (1) scaffolding strategies and (2) laboratory activities and experiments. They are designed around a range of learning objectives, including cultivation of the ability to analyze and interpret spectral data, proficiency with practical considerations such as use of instrumentation and sample preparation, and knowledge of physical principles underlying spectroscopy. A number of innovations are highlighted herein to illustrate the types of resources available, as well as to provide general insight into essential components necessary to effectively support learning. These innovations primarily support students' triangulation of data across spectra and connection between molecular structure and spectroscopic properties. Notably, the majority do not directly support the development of sophisticated chemical knowledge relating to implicit molecular features and spectroscopic properties, multivariate reasoning involving these implicit features, and use of a consistent, systematic approach.

1.6.1 Scaffolding strategies

Scaffolding strategies for spectroscopic structure elucidation come in a variety of forms, including in-class group activities, online modules, and written prompts. As noted, an empirical study of textbooks' potential for scaffolding ^1H NMR spectral interpretation revealed that they neither provide sufficient practice with interpreting different spectral features (e.g., peak integration) nor order this practice in a systematic way.¹⁵ While this investigation focused only on ^1H NMR, it is possible that this lack of scaffolding extends to coverage of other spectroscopic methods; scaffolding innovations are thus likely necessary to adequately support learning.

Notably, one scaffolding innovation was incorporated into Stowe and Cooper's investigation of undergraduates' argumentation using spectral data, thus providing some empirical support for the efficacy, or lack thereof, of certain approaches.¹² Findings from this study point toward general guidelines for the design and selection of new and existing approaches, respectively. Ultimately, they suggest that significant scaffolding is needed to support students' triangulation of data from multiple sources in addition to their analysis of individual spectra. For this investigation, the authors developed and administered a variety of written prompts designed to scaffold participants' elucidation of molecular structure using IR, ^{13}C , and ^1H NMR spectra. The development of these prompts serves as a useful exemplar for the development of other scaffolding strategies for spectral interpretation, as well as for the design of formative and summative assessments. Specifically, they were designed using the 3-Dimensional Learning Assessment Protocol (3D-LAP), a set of criteria that are shown to effectively elicit students' engagement in science and engineering practices defined in the National Research Council's Framework for K-12 Science Education.³⁸ The prompts incorporated four criteria for eliciting evidence of engagement in argumentation from evidence; in general, they instructed participants to (1) make a claim (i.e. provide a molecular structure) consistent with (2) an observation (i.e. provided spectra), and then to (3) indicate evidence that supports the claim (i.e. identify spectral features that suggest a given molecular structure) and (4) provide reasoning about why evidence supports the claim (i.e. explain why these features suggest this structure). Prompts ranged in their

amount of scaffolding; a low structure prompt that did not incorporate the four criteria was used to compare the effect of scaffolding on undergraduates' ability to engage in argumentation. A moderately structured prompt provided the general instructions described above, and a highly structured prompt provide these instructions and also scaffolded participants' selection of spectral evidence and reasoning. The number of spectra provided with the moderately structured prompt was also varied to evaluate whether the amount of provided data affected individuals' ability to triangulate information. Results from this study indicated that even with significant scaffolding, participants were unable to integrate this information and provide a molecular structure consistent with all identified evidence. Further, participants were more able to provide such a molecular structure when given three spectra (IR, ^{13}C , and ^1H NMR) as opposed to two (IR and ^{13}C NMR). These findings suggest that scaffolding strategies should seek to guide students through the process of data triangulation in addition to the analysis and interpretation of individual spectra. They also suggest that the challenge of integrating spectral data lies beyond individuals' limited processing capacity and that simply varying the amount of information provided will not effectively support learning.

An innovation by Winschel *et al.* offers one promising approach for scaffolding both the integration of multiple sources of data and analysis of individual spectra.³⁹ This strategy was used in a second-semester organic chemistry laboratory course and involved cooperative learning, an evidence-based instructional practice that promotes the development of problem-solving skills. As part of this innovation, students completed a weekly jigsaw activity in an online environment. For each activity, students were divided into groups of four to six, and each student analyzed an individual piece of data on a spectrum. They then posted their analysis on a discussion board where they worked collaboratively to elucidate the corresponding molecular structure. The complexity of each activity increased every week, where students began with analyzing an ^1H NMR spectrum and then moved to combination problems involving an ^1H NMR spectrum and supplementary IR or ^{13}C NMR spectra. Analysis of student activity in the online environment, together with attitudinal surveys, indicated that students found the innovation to be a productive tool for learning spectroscopy. A similar innovation

developed by Flynn offers another strategy for promoting students' ability to analyze individual spectra, though it could be easily adapted to cultivate their ability to integrate data from multiple sources.³⁷ This strategy was used in a large, lecture-based classroom and designed to help students interpret ^1H NMR spectra. During the activity, students were presented with a spectrum and molecular formula corresponding to an unknown compound. They then determined molecular fragments with the help of a table that required them to deduce chemical shift, integration, and multiplicity of each signal. Students then created sticky notes of each molecular fragment and arranged them in different ways to determine the most plausible structure. The use of a document camera and classroom response system facilitated class-wide collaboration.

Further, an instructional innovation by Angawi combined cooperative learning with the use of a variety of written prompts to scaffold ^1H NMR spectral interpretation.⁴⁰ Unlike the previously described innovations, this semester-long approach was designed for an upper-level undergraduate spectroscopy course. Throughout the semester, students worked in small groups to complete a variety of tasks involving the interpretation ^1H NMR spectra. Written prompts guided each task and required students to assign spectra to *cis* and *trans* isomers, predict the observed spectra of molecules by tabulating expected data (e.g., chemical shift) in a provided table, and elucidate molecular structure using a provided table, molecular formula, and spectrum. Students also received targeted instruction on problem-solving approaches and their application during spectral interpretation. Course grades of students who completed the innovation were statistically higher than those who did not, providing some evidence of the approach's efficacy. However, it is uncertain whether the structure of written prompts, cooperative environment, or problem-solving instruction resulted in improved learning. Stowe and Cooper's finding that prompt structure did not influence participants' ability to propose accurate molecular structures suggests that the cooperative environment or problem-solving instruction used in this approach were essential for supporting learning.¹²

A number of free online programs are also available to scaffold students' learning of organic structure determination via spectroscopic analysis, including interactive tutorials and free spectral databases. These online resources are designed for a range of

educational levels and spectroscopic techniques. Debska and Guzowska-Swider provide a comprehensive overview of these available resources, noting programs such as ^1H NMR Spectroscopy, which allows students to predict spectra and compare these with reference spectra, and the SDBS Integrated Spectral Data Base System for Organic Compounds, an integrated spectral library that includes ^1H NMR, ^{13}C NMR, FT-IR, and laser-Raman spectra.^{41–43} Other programs, such as iSpec, cultivate students' understanding of practical considerations involved in spectral analysis and interpretation, including how to process spectra and weigh the costs and benefits of conducting multiple analyses.⁴⁴

1.6.2 Laboratory experiments and activities

In addition to scaffolding approaches, a number of laboratory experiments and activities have been developed to cultivate students' ability relating to spectroscopic analysis. Notably, Wright and Oliver-Hoyo's investigation of undergraduates' mental models during IR spectral interpretation provides insight into the design of effective laboratory activities, as individuals in this study completed a hands-on learning activity prior to participation.^{9,45} This activity involved use of a physical model that incorporates the harmonic oscillator model and Hooke's law to study dynamic vibrations within diatomic molecules. The model combined whiffle balls, magnets, and springs to illustrate the effects of bond order, polarity, and reduced mass on IR peak position and the effects of bond abundance and dipole on IR peak intensity. The activity required manipulation of the physical model in a variety of ways to illustrate different effects, such as changing spring strength to illustrate the effect of bond order on vibration frequency. Following each manipulation of the physical model, students were then instructed to use this physical observation to explain the effect of each variable on peak position or intensity for a given spectrum. The model was tested with over four hundred first-semester organic chemistry students, the majority of which were able to abstract principles from the model and use these principles to determine an approximate value for peak position and intensity. A number of undergraduates who completed the activity and then participated in the subsequent investigation also exhibited the more sophisticated mental model that molecules exhibit dynamic bonding interactions. This more sophisticated mental model was observed

predominantly amongst graduate students in another study that was not preceded by such an activity.¹³ The prevalence of the more sophisticated mental model among undergraduates after participating in this activity suggests that it successfully cultivated students' understanding of physical principles underlying spectroscopy and their ability to reason about structure-property relationships. The potential of this activity for promoting learning may be attributed to a variety of sources, including effective visualization of dynamic sub-microscopic interactions and scaffolding to help students use this physical understanding to relate structure to spectroscopic properties.

As suggested by findings from Wright and Oliver-Hoyo's investigation, laboratory experiments that effectively cultivate students' ability relating to spectroscopic analysis are likely those that directly scaffold the connection between molecular structure and spectroscopic properties. Many laboratory experiments relating to spectroscopic analysis accomplish this objective, and collectively they span a range of techniques. For instance, one second-year organic chemistry laboratory has been designed to introduce students to ¹³C NMR and scaffold their connection between the molecular structure of alcohols and chemical shift.⁴⁶ Further, many laboratories scaffold the connection between molecular structure and properties in the context of practical applications. Examples include a second-year laboratory in which students use ¹H NMR to determine the structure of natural amino acids and either a first- or second-year laboratory in which students determine average chain length and degree of unsaturation of common edible oils using ¹³C NMR.^{47,48} Notably, these experiments also supports students' learning of sample preparation and instrumentation usage, other skills essential to spectroscopic structural elucidation.

1.7 Implications

Previous studies on the teaching and learning of spectroscopic analysis provide a number of implications for the direction of future research. They also serve to inform the design of instruction, as briefly described in the previous section on instructional innovations. Addressing these implications in both future research and the classroom will be essential

for further advancing our understanding of how ability in this practice develops and ultimately supporting learning.

1.7.1 Implications for research

In addition to providing a foundation for understanding the development of ability in spectroscopic structure elucidation, previous studies also point toward necessary and productive avenues for future research. These avenues correspond to additional facets of ability yet to be investigated, either due to inherent limitations of previous study designs or lack of intentional investigative focus. One of the most promising directions involves the investigation of chemical knowledge, reasoning, and problem-solving strategies used for spectral interpretation in more authentic scientific contexts. As noted, a number of previous investigations required participants to complete spectral interpretation tasks that do not directly mirror how spectra are most often used within the scientific community. For instance, synthetic chemists often have knowledge of the compound they intend to synthesize and then evaluate spectra for evidence of their target product.³ This evaluation involves predication and identification of expected peaks based on the target product's molecular structure, as well as consideration of byproduct and unexpected product formation, the purity of starting material, and product yield, among other factors.³ A number of studies on the learning of this practice, however, involve complete structural elucidation using molecular formulae and spectra. Chemistry education researchers are increasingly adopting a resources-based view of cognition, which acknowledges that context directly influences the knowledge and reasoning one uses to approach a task or problem.⁴⁹ If research is to more completely characterize the development of ability in spectroscopic analysis and meaningfully inform instruction, then it must investigate knowledge and reasoning in these authentic contexts. An evaluation of findings situated in more authentic contexts will then allow researchers to assess the transferability of earlier findings. Further, situating spectral interpretation tasks in a more authentic context would also allow for a meaningful investigation of individuals' visuospatial assumptions, including how they evolve in sophistication. As noted, these assumptions are relatively uncharacterized though particularly relevant for spectroscopic structural elucidation.

Two studies by Connor *et al.* make an initial step toward investigating reasoning in more authentic scientific contexts, as they characterized participants' chemical assumptions while determining the success of syntheses using spectra.^{10,11} One of these studies also characterized the heuristic reasoning strategies that constrained undergraduates' reasoning during this practice. However, additional research is needed to understand both students' and practicing chemists' use of productive heuristic reasoning strategies in such a context. Experienced chemists regularly use heuristics to facilitate their decision making, albeit with knowledge of appropriate contexts in which they can be applied.⁵⁰ Characterization of the heuristics that guide practicing chemists' interpretation of spectra could then be used to inform instruction that promotes this expert-like thinking. A similar investigation among students would allow for the identification of productive reasoning strategies which could be leveraged to further promote ability.

Additional facets of ability yet to be explored include individuals' conceptual understanding of introductory physical principles underlying NMR, including nuclear spin states, nuclear magnetic moments, the absorption of energy, spin-spin coupling, and the timescale of NMR, among others. Studies that investigated individuals' conceptual understanding as they interpreted spectra provide a number of insights into reasoning involving physical principles underlying IR but not those underlying NMR; more structured conceptual questions such as those used by Domin and Bodner may thus be necessary to elicit such understanding.⁸ Knowledge of introductory physical principles is essential for evaluating spectral features such as the multiplicity of signals, a feature for which undergraduates often rigidly apply the N+1 rule without considering other variables.¹¹ Investigating the teaching and learning of introductory physical principles underlying NMR will thus be essential for effectively promoting sophisticated, multivariate thinking among students.

Further, the ability to use words to identify and analyze features of spectra, what Kozma and Russell deem as an essential component of chemists' representational competence,² is also understudied. The importance of this ability is due to the function that spectra serve among synthetic chemists; not only are they representations of submicroscopic phenomena, but they are also tools that support social discourse.² For

instance, synthetic chemists often use NMR spectra to demonstrate to others that a target compound was successfully synthesized. If chemists were unfamiliar with the specialized language specific to spectroscopic techniques, they would be unable to engage in this discourse. Further, Kozma and Russell found that undergraduate chemistry students do not engage in such discourse when in the laboratory, with their discussion instead focusing on practical considerations such as the setup of equipment.² One possible cause for this lack of engagement may be due to limited familiarity with this complex terminology. Investigating how familiarity with this terminology forms and how such knowledge may be cultivated in the classroom will thus be essential for designing instruction that supports the development of relevant ability. The formative assessment developed by the author to measure lexical ability in ¹H NMR spectroscopy will serve as an essential research tool in such investigations.¹⁴

1.7.2 Implications for instruction

Implications for the design of instruction were briefly noted in the previous discussion of instructional innovations; in part, they include the importance of scaffolding students' ability to triangulate data across spectra and make connections between molecular structure and spectroscopic properties. Previous research provides a number of additional implications beyond these considerations that current instructional innovations do not yet incorporate. Collectively, findings suggest that instruction should be designed to (1) cultivate sophisticated chemical knowledge relating to the nature and behaviour of implicit molecular features and spectroscopic properties, as well as their relationship to one another (2) promote multivariate reasoning ability that involves the evaluation of multiple implicit chemical features and their spectroscopic properties, and (3) foster consistency in this evaluation. Together, these components will support students' learning and promote their ability to engage in this essential disciplinary practice.

Supporting students' ability to relate implicit molecular features to corresponding spectroscopic data will involve the design of classroom instruction and curricular materials that foster sophisticated conceptual understanding of the nature and behaviour of such features and data, as well as their relationship to one another. Studies demonstrated a

number of dimensions that underly such conceptual understanding.^{9-11,13} These dimensions include knowledge of: the dynamic nature of bonding and its effect on IR spectral data; implicit visuospatial aspects of molecules such as topology and symmetry, as well as their effect on spectral data; flexibility of the N+1 rule and implicit molecular features that result in deviations; the variable nature of spectral data and implicit molecular features that influence absorption frequency; and characteristic spectral features such as the IR peak intensity of hydroxyl groups and implicit molecular features that influence absorption intensity.¹¹ Studies have yet to investigate individuals' conceptual understanding of introductory physical principles underlying NMR (e.g., spin-spin coupling), though such understanding is also necessary for relating implicit molecular features to NMR spectral features. Fostering understanding and familiarity across these dimensions will provide a conceptual foundation requisite to reasoning focused on relationships between implicit molecular features and spectroscopic properties. In addition, it will help shift students from reasoning that focuses only on the evaluation of explicit molecular features and unfamiliar spectral data toward more expert-like thinking.

In addition to fostering conceptual understanding across all relevant dimensions, instruction must also be designed to support students' ability to engage in multivariate, analytical reasoning. A number of studies demonstrated that students engage in unproductive heuristic reasoning strategies when evaluating spectroscopic data, with the most common strategy involving decision making based on a single piece of evidence (i.e. ORDM).^{11,12,45} Effective instruction will thus incorporate strategies that shift students from the inappropriate use of this heuristic toward thinking that involves the evaluation and weighing of multiple pieces of evidence. Researchers have identified several strategies for shifting students toward multivariate reasoning, and they could readily be incorporated into instruction on spectroscopic structure elucidation. For instance, research in cognitive psychology suggests that having students consider the opposite or assess the correctness of any decision helps to adjust for decision biases resulting from the use of heuristics.^{51,52} Chemistry education research also suggests that having students predict how others may incorrectly respond to a question promotes analytical thinking.⁵³ These simple strategies could readily be incorporated into classroom

instruction and curricular materials. Further, research on cognitive biases in medicine offers a more targeted approach, where findings suggest that practitioners engage in more analytical thinking when provided with descriptions of common heuristics and examples of how they result in biases in clinical contexts. As Connor *et al.* has noted, this approach could be adapted for instruction on spectroscopic structure elucidation, where students are provided with descriptions of common heuristics and examples how they result in biases when evaluating given spectra.¹¹

Studies on problem-solving aspects of spectroscopic structure elucidation collectively demonstrate that an informed, consistent, and systematic approach is requisite for successful, expert-like spectral interpretation. Both classroom instruction and curricular materials should thus support and encourage students' use of such an approach. Scaffolding innovations hold particular promise for cultivating this aspect of ability, as they have the potential to provide ordered, highly structured support. Previously published innovations, however, do not directly scaffold all three of these components. The design of such an approach will thus be an important future step toward promoting relevant ability.

Lastly, previous studies provide important implications for the design of assessments and instructor preparation, two essential yet somewhat unexplored components of effective instruction on spectroscopic structure elucidation. With regard to assessments, findings suggest that limited chemical knowledge of an individual spectral feature has the potential to derail students' interpretation efforts, particularly if they are also engaging in unproductive heuristic reasoning strategies.¹¹ Assessments should thus be design to assess students' knowledge in individual dimensions of understanding in order to provide instructors with a more accurate measure of students' ability and a comprehensive view of content areas that require additional instructional focus. As previously noted, the utility and efficacy of assessments would also benefit from the incorporation of criteria shown to elicit evidence of argumentation.¹⁸ Such inclusion will further help ensure that assessments provide instructors with an accurate measure of ability. With regard to instructor preparation, findings suggest that instructional experience specifically in spectroscopic analysis results greater knowledge for teaching this

practice.¹⁶ These findings also suggest that knowledge for teaching this practice is sub-topic and problem-type specific, meaning that instructor education materials should aim to cultivate knowledge for teaching relevant sub-topics and problem types in order to improve instruction quality and student learning outcomes.

1.8 Conclusions

Spectroscopic structure elucidation will remain an essential practice for synthetic chemists, and research on the development of relevant ability will play an important role in informing the design of instruction. Thus far, findings have demonstrated that sophisticated chemical knowledge of implicit molecular features and their spectroscopic properties, multivariate reasoning, and consistency and efficiency in approach are hallmarks of this ability. As researchers work to investigate additional facets of ability, instruction should be designed to support learning of these three components. A number of promising instructional innovations for supporting learning of individual components were identified herein, in addition to implications for the design of future innovations that support the three together. When the evidence-based design of instruction is employed in combination with the similar design of instructor education materials and assessment, teaching and learning should take on a form that efficiently supports students' ability to engage in this essential practice.

1.9 References

- (1) Schummer, J. The Impact of Instrumentation on Chemical Species Identity: From Chemical Substances to Molecular Species. In *From Classical to Modern Chemistry: The Instrumental Revolution*; Morris, P., Ed.; The Royal Society of Chemistry: Cambridge, 2002; pp 188–211.
- (2) Kozma, R.; Russell, J. Students Becoming Chemists: Developing Representational Competence. In *Visualization in Science Education*; Gilbert, J. K., Ed.; Springer: Dordrecht, 2005; pp 121–146. https://doi.org/10.1007/1-4020-3613-2_4.
- (3) 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. *Chem. Educ. Res. Pract.* **2012**, *13* (3), 179–185. <https://doi.org/10.1039/C2RP90001G>.

- (4) Steigel, A. Mechanistic Studies of Rearrangements and Exchange Reactions by Dynamic NMR Spectroscopy. In *NMR Basic Principles and Progress*; Diehl, P., Fluck, E., Kosfeld, R., Eds.; Springer-Verlag: Berlin, 1982; pp 1–44. https://doi.org/10.1524/zpch.1972.80.1_2.110.
- (5) Binsch, G.; Kessler, H. The Kinetic and Mechanistic Evaluation of NMR Spectra. *Angew. Chemie Int. Ed. English* **1980**, *19* (6), 411–428. <https://doi.org/10.1002/anie.198004113>.
- (6) Cartrette, D. P.; Bodner, G. M. Non-Mathematical Problem Solving in Organic Chemistry. *J. Res. Sci. Teach.* **2010**, *47* (6), 643–660. <https://doi.org/10.1002/tea.20306>.
- (7) Topczewski, J. J.; Topczewski, A. M.; Tang, H.; Kendhammer, L. K.; Pienta, N. J. NMR Spectra through the Eyes of a Student: Eye Tracking Applied to NMR Items. *J. Chem. Educ.* **2017**, *94* (1), 29–37. <https://doi.org/10.1021/acs.jchemed.6b00528>.
- (8) Domin, D.; Bodner, G. Using Students' Representations Constructed during Problem Solving to Infer Conceptual Understanding. *J. Chem. Educ.* **2012**, *89* (7), 837–843. <https://doi.org/10.1021/ed1006037>.
- (9) Wright, L. C.; Oliver-Hoyo, M. T. Student Assumptions and Mental Models Encountered in IR Spectroscopy Instruction. *Chem. Educ. Res. Pract.* **2020**, *21* (1), 426–437. <https://doi.org/10.1039/c9rp00113a>.
- (10) Connor, M. C.; Glass, B. H.; Finkenstaedt-Quinn, S. A.; Shultz, G. V. Developing Expertise in ¹H NMR Spectral Interpretation. *J. Org. Chem.* **2020**, *acs.joc.0c01398*. <https://doi.org/10.1021/acs.joc.0c01398>.
- (11) 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>.
- (12) Stowe, R. L.; Cooper, M. M. Arguing from Spectroscopic Evidence. *J. Chem. Educ.* **2019**, *96* (10), 2072–2085. <https://doi.org/10.1021/acs.jchemed.9b00550>.
- (13) Cullipher, S.; Sevian, H. Atoms versus Bonds: How Students Look at Spectra. *J. Chem. Educ.* **2015**, *92* (12), 1996–2005. <https://doi.org/10.1021/acs.jchemed.5b00529>.
- (14) Connor, M. C.; Glass, B. H.; Shultz, G. V. Development of the NMR Lexical Representational Competence (NMR-LRC) Instrument as a Formative Assessment of Lexical Ability in ¹H NMR Spectroscopy. *Prep.*
- (15) Anderson, S. Y. C.; Ong, W. S. Y.; Momsen, J. L. Support for Instructional

- Scaffolding with ^1H NMR Spectral Features in Organic Chemistry Textbook Problems Shannon. *Chem. Educ. Res. Pract.* **2020**, Advance Article. <https://doi.org/10.1039/c9rp00252a>.
- (16) Connor, M. C.; Shultz, G. V. Teaching Assistants' Topic-Specific Pedagogical Content Knowledge in ^1H NMR Spectroscopy. *Chem. Educ. Res. Pract.* **2018**, *19* (3), 653–669. <https://doi.org/10.1039/c7rp00204a>.
- (17) Kozma, R. The Material Features of Multiple Representations and Their Cognitive and Social Affordances for Science Understanding. *Learn. Instr.* **2003**, *13* (2), 205–226. [https://doi.org/10.1016/s0959-4752\(02\)00021-x](https://doi.org/10.1016/s0959-4752(02)00021-x).
- (18) 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>.
- (19) Sevian, H.; Talanquer, V. Rethinking Chemistry: A Learning Progression on Chemical Thinking. *Chem. Educ. Res. Pract.* **2014**, *15*, 10–23. <https://doi.org/10.1039/c3rp00111c>.
- (20) Wilson, M.; Berkeley, U. C.; June, R.; May, A. Measuring Progressions: Assessment Structures Underlying a Learning Progression. **2009**, *46* (6), 716–730. <https://doi.org/10.1002/tea.20318>.
- (21) Ericsson, K. A.; Simon, H. A.; Ericsson; Simon. Verbal Reports as Data. *Psychol. Rev.* **1980**, *87* (3), 215–251. <https://doi.org/10.1037/0033-295X.87.3.215>.
- (22) van Gog, T.; Paas, F.; van Merriënboer, J. J. G. G.; Witte, P. Uncovering the Problem-Solving Process: Cued Retrospective Reporting versus Concurrent and Retrospective Reporting. *J. Exp. Psychol. Appl.* **2005**, *11* (4), 237–244. <https://doi.org/10.1037/1076-898X.11.4.237>.
- (23) Just, M. A.; Carpenter, P. A. A Theory of Reading: From Eye Fixations to Comprehension. *Psychol. Rev.* **1980**, *87* (4), 329–354. <https://doi.org/10.1037/0033-295X.87.4.329>.
- (24) Cullipher, S.; Hansen, S. J. R.; Vandenplas, J. R. Eye Tracking as a Research Tool: An Introduction. In *Eye Tracking for the Chemistry Education Researcher*; American Chemical Society: Washington, DC, 2018; pp 1–9.
- (25) Gegenfurtner, A.; Lehtinen, E.; Säljö, R. Expertise Differences in the Comprehension of Visualizations: A Meta-Analysis of Eye-Tracking Research in Professional Domains. *Educ Psychol Rev* **2011**, *23*, 523–552. <https://doi.org/10.1007/s10648-011-9174-7>.
- (26) 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>.
- (27) Harle, M.; Towns, M. A Review of Spatial Ability Literature, Its Connection to Chemistry, and Implications for Instruction. *J. Chem. Educ.* **2011**, *88* (3), 351–360. <https://doi.org/10.1021/ed900003n>.
- (28) Evans, J. S. B. T.; Stanovich, K. E. Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspect. Psychol. Sci.* **2013**, *8* (3), 223–241. <https://doi.org/10.1177/1745691612460685>.
- (29) Tversky, A.; Kahneman, D. Judgment under Uncertainty: Heuristics and Biases. *Science* (80-). **1974**, *185* (4157), 1124–1131. <https://doi.org/10.4324/9781912282562>.
- (30) Shulman, L. Those Who Understand: Knowledge Growth in Teaching. *Educ. Res.* **1986**, *15* (2), 4–14. <https://doi.org/10.3102/0013189X015002004>.
- (31) Hill, H. C.; Rowan, B.; Ball, D. L. Effects of Teachers' Mathematical Knowledge for Teaching on Student Achievement. *Am. Educ. Res. J.* **2005**, *42* (2), 371–406. <https://doi.org/10.3102/00028312042002371>.
- (32) Park, S.; Jang, J. Y.; Chen, Y. C.; Jung, J. Is Pedagogical Content Knowledge (PCK) Necessary for Reformed Science Teaching?: Evidence from an Empirical Study. *Res. Sci. Educ.* **2010**, *41*, 245–260. <https://doi.org/10.1007/s11165-009-9163-8>.
- (33) Hale, L. V. A.; Lutter, J. C.; Shultz, G. V. The Development of a Tool for Measuring Graduate Students' Topic Specific Pedagogical Content Knowledge of Thin Layer Chromatography. *Chem. Educ. Res. Pract.* **2016**, *17*, 700–710. <https://doi.org/10.1039/C5RP00190K>.
- (34) Jüttner, M.; Boone, W.; Park, S.; Neuhaus, B. J. Development and Use of a Test Instrument to Measure Biology Teachers' Content Knowledge (CK) and Pedagogical Content Knowledge (PCK). *Educ. Assessment, Eval. Account.* **2013**, *25*, 45–67. <https://doi.org/10.1007/s11092-013-9157-y>.
- (35) Mavhunga, E.; Rollnick, M. Improving PCK of Chemical Equilibrium in Pre-Service Teachers Improving PCK of Chemical Equilibrium in Pre-Service Teachers. *African J. Res. Math. Sci. Technol. Educ.* **2013**, *17* (1–2), 113–125. <https://doi.org/10.1080/10288457.2013.828406>.
- (36) Dunnagan, C. L.; Dannenberg, D. A.; Cuales, M. P.; Earnest, A. D.; Gurnsey, R. M.; Gallardo-Williams, M. T. Production and Evaluation of a Realistic Immersive Virtual Reality Organic Chemistry Laboratory Experience: Infrared Spectroscopy. *J. Chem. Educ.* **2020**, *97* (1), 258–262.

<https://doi.org/10.1021/acs.jchemed.9b00705>.

- (37) Flynn, A. B. NMR Interpretation: Getting from Spectrum to Structure. *J. Chem. Educ.* **2012**, *89* (9), 1210–1212. <https://doi.org/10.1021/ed3000974>.
- (38) Lavery, J. T.; Underwood, S. M.; Matz, R. L.; Posey, L. A.; Carmel, J. H.; Caballero, M. D.; Fata-Hartley, C. L.; Ebert-May, D.; Jardeleza, S. E.; Cooper, M. M. Characterizing College Science Assessments: The Three-Dimensional Learning Assessment Protocol. *PLoS One* **2016**, *11* (9), 1–21. <https://doi.org/10.1371/journal.pone.0162333>.
- (39) Winschel, G. A.; Everett, R. K.; Coppola, B. P.; Shultz, G. V.; Lonn, S. Using Jigsaw-Style Spectroscopy Problem-Solving to Elucidate Molecular Structure through Online Cooperative Learning. *J. Chem. Educ.* **2015**, *92*, 1188–1193. <https://doi.org/10.1021/acs.jchemed.5b00114>.
- (40) Angawi, R. F. Using a Problem Solving-Cooperative Learning Approach to Improve Students' Skills for Interpreting ¹H NMR Spectra of Unknown Compounds in an Organic Spectroscopy Course. *J. Chem. Educ.* **2014**, *91*, 823–829. <https://doi.org/10.1021/ed4004436>.
- (41) Debska, B.; Guzowska-Swider, B. Molecular Structure from ¹H NMR Spectra: Education Aided by Internet Programs. *J. Chem. Educ.* **2007**, *84* (3), 556–560. <https://doi.org/10.1021/ed084p556>.
- (42) National Institute of Advanced Industrial Science and Technology. SDBSWeb <http://sdb.sdb.aist.go.jp> (accessed May 4, 2020).
- (43) Wong, Y.-L. ¹H NMR Interpretation Tutorial <http://users.wfu.edu/ylwong/chem/nmr/h1/> (accessed May 4, 2020).
- (44) Vosegaard, T. ISpec: A Web-Based Activity for Spectroscopy Teaching. *J. Chem. Educ.* **2018**, *95* (1), 97–103. <https://doi.org/10.1021/acs.jchemed.7b00482>.
- (45) Wright, L. C.; Oliver-Hoyo, M. T. Supporting the Teaching of Infrared Spectroscopy Concepts Using a Physical Model. *J. Chem. Educ.* **2019**, *96* (5), 1015–1021. <https://doi.org/10.1021/acs.jchemed.8b00805>.
- (46) Chamberlain, P. H. Identification of an Alcohol with ¹³C NMR Spectroscopy. *J. Chem. Educ.* **2013**, *90* (10), 1365–1367. <https://doi.org/10.1021/ed300833s>.
- (47) Fry, C. G.; Hofstetter, H.; Bowman, M. D. A First Laboratory Utilizing NMR for Undergraduate Education: Characterization of Edible Fats and Oils by Quantitative ¹³C NMR. *J. Chem. Educ.* **2017**, *94* (9), 1319–1323. <https://doi.org/10.1021/acs.jchemed.7b00057>.

- (48) Zivkovic, A.; Bandolik, J. J.; Skerhut, A. J.; Coesfeld, C.; Zivkovic, N.; Raos, M.; Stark, H. Introducing Students to NMR Methods Using Low-Field ^1H NMR Spectroscopy to Determine the Structure and the Identity of Natural Amino Acids. *J. Chem. Educ.* **2017**, *94* (1), 115–120. <https://doi.org/10.1021/acs.jchemed.6b00168>.
- (49) Hammer, D.; Elby, A.; Scherr, R. E.; Redish, E. F.; Hammer, D.; Elby, A.; Scherr, R. E.; Redish, E. F. Resources, Framing , and Transfer. In *Transfer of Learning from a Modern Multidisciplinary Perspective*; Mestre, J. R., Ed.; Information Age Publishing: Greenwich, 2005; pp 89–119.
- (50) Talanquer, V. Chemistry Education: Ten Heuristics to Tame. *J. Chem. Educ.* **2014**, *91* (8), 1091–1097. <https://doi.org/10.1021/ed4008765>.
- (51) Mussweiler, T.; Strack, F.; Pfeiffer, T. Overcoming the Inevitable Anchoring Effect: Considering the Opposite Compensates for Selective Accessibility. *Personal. Soc. Psychol. Bull.* **2000**, *26* (9), 1142–1150.
- (52) Thompson, V. A.; Turner, J. A. P.; Pennycook, G. Intuition, Reason, and Metacognition. *Cogn. Psychol.* **2011**, *63* (3), 107–140. <https://doi.org/10.1016/j.cogpsych.2011.06.001>.
- (53) Talanquer, V. Concept Inventories: Predicting the Wrong Answer May Boost Performance. *J. Chem. Educ.* **2017**, *94*, 1805–1810. <https://doi.org/10.1021/acs.jchemed.7b00427>.

Chapter 2

Constraints on Organic Chemistry Students' Reasoning During IR and ^1H NMR Spectral Interpretation

2.1 Initial remarks

This chapter corresponds to the first of two studies investigating the development of expertise in ^1H NMR spectral interpretation. These studies characterized multiple facets of expertise among undergraduate and doctoral chemistry students to provide a comprehensive view of how such proficiency may develop. Among these facets were conceptual understanding in terms of underlying chemical assumptions, heuristic reasoning strategies, and cognitive processes such as visual attention. Notably, both studies also investigated the interpretation of complementary IR spectra, though results from this series of investigations relate predominantly to ^1H NMR spectroscopy. The inclusion of complementary IR spectra primarily afforded insight into how individuals learn to triangulate ^1H NMR spectra with other sources of spectral data, an essential skill among organic chemists and a hallmark of expertise ^1H NMR spectral interpretation.^{1,2} Further, this inclusion of IR spectra provided insight into the degree to which findings may transfer to other spectral interpretation tasks in chemistry (e.g., mass spectra).

The first study sought to identify inaccurate chemical assumptions, heuristics, and sets of these assumptions and heuristics used by undergraduates when interpreting ^1H NMR and complementary IR spectra. Underlying assumptions and heuristic reasoning strategies facilitate decision making, an inherent aspect of spectral interpretation. Further, sets of inaccurate assumptions and heuristic reasoning strategies often work together to act as barriers to learning and, in turn, developing expert-like thinking. The first study thus served to provide insight into potential cognitive barriers associated with first learning the practice of spectral interpretation. Findings from this study suggest that five categories of

inaccurate chemical assumptions in part guided undergraduates' reasoning. These assumptions were often rule-based overgeneralizations of learned principles (e.g., expected number of signals based on the number of chemical environments, chemical shift, the N+1 rule, etc.) which focused on context rather than underlying chemical principles. Further, undergraduates' thinking was constrained when they rigidly applied these rule-based overgeneralizations to ultimately make decisions using just one variable. In these instances, participants restricted their analysis to spectral features that deviated from overgeneralized learned principles while failing to evaluate implicit chemical features resulting in variability. These findings provide insight into the design of introductory-level instruction that helps novices both avoid and overcome these cognitive barriers to learning. The second study, which corresponds to Chapter 3, builds upon these findings by investigating how conceptual understanding develops beyond these inaccurate chemical assumptions. Specifically, the second study characterizes how all underlying chemical assumptions among undergraduate and doctoral chemistry students evolve with increasing expertise.

This chapter first appeared as a research article in *Chemistry Education Research and Practice*, and the original publication and copyright information are provided below. The original publication was modified to adhere to Rackham dissertation formatting requirements, though no additional changes were made. Dr. Solaire Finkenstaedt-Quinn assisted with the design of codebooks and qualitative coding of interview data, as well as provided feedback on the clarity and organization of the manuscript prior to journal submission. All remaining work, including study design, data collection, thematic analysis, quantitative data analysis, and writing of the manuscript were completed independently by the author.

Original publication and copyright information

Reproduced from “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>” with permission from the Royal Society of Chemistry.

2.2 Abstract

Promoting students' ability to engage in discipline-specific practices is a central goal of chemistry education. Yet if instruction is to meaningfully foster such ability, we must first understand students' reasoning during these practices. By characterizing constraints on chemistry students' reasoning, we can design instruction that targets this constrained reasoning and ultimately promotes more sophisticated ways of thinking. For this study, we investigated reasoning used by 18 organic chemistry students at a large university in the United States as they evaluated the success of chemical syntheses through IR and ^1H NMR spectral interpretation, a common task of practicing chemists. Students completed a series of interpretation tasks while having their eye movements tracked and then participated in semi-structured, cued retrospective think-aloud (RTA) interviews about their reasoning during spectral interpretation. RTA interviews were analyzed qualitatively to characterize invalid chemical assumptions and heuristic reasoning strategies used by participants, both of which science education literature identifies as fundamental constraints to learning. The most problematic assumptions and heuristics, i.e., those used more frequently by unsuccessful participants, were then identified through statistical analysis. Findings suggest that the most problematic constraints on students' reasoning during spectral interpretation constitute a combination of particular invalid chemical assumptions and heuristic reasoning strategies.

2.3 Introduction

Characterization of molecular structure is a fundamental practice of chemistry that is typically accomplished through spectroscopic analysis, where infrared (IR) and nuclear magnetic resonance (NMR) spectroscopy serve as key structural determination methods. The importance of IR and NMR spectroscopy to the undergraduate chemistry curriculum is therefore not surprising, with many organic chemistry textbooks devoting at least one chapter to these characterization methods.³ Spectral interpretation is an inherent aspect of this practice, yet despite its importance to practicing chemists and the instructional focus on spectroscopic techniques, few research studies investigate student knowledge and learning of this aspect broadly or for IR and NMR specifically. This relative lack of

research is problematic given the complex nature of spectral interpretation; for students to correctly interpret spectral data, they must not only be able to recognize functional groups, determine electronegativity effects, and identify molecular symmetry but also understand how molecules interact with electromagnetic radiation. In addition, students must be able to apply their knowledge of spectroscopy and molecular properties to graphical representations as well as translate between molecular and graphical representations as they reason.

2.3.1 Teaching and learning IR and NMR spectroscopy

The vast majority of literature on teaching and learning IR and NMR spectroscopy focuses on scaffolding strategies and laboratory activities, with minimal investigation of learning outcomes.^{4–11} However, a small number of studies have investigated how undergraduates, graduate students, and faculty members interpret IR and NMR spectra.^{12–14} Findings from these studies provide some insight into how individuals learn to engage in this aspect of spectroscopic analysis.

In an investigation of graduate students' and faculty members' spectral interpretation approaches for combined IR and ^1H NMR problems, Cartrette and Bodner found that successful participants often had more experience solving complex spectra and applied a consistent approach across problems.¹² Additionally, successful participants were more flexible in their understanding of the "N+1 rule" for ^1H NMR spectral interpretation and could effectively explain deviations from the rule. Cullipher and Sevia used eye tracking and think-aloud interviewing to investigate undergraduates' and graduate students' reasoning as they related molecular structures to corresponding IR spectra.¹³ This study found that participants in given educational levels (1) relied on different assumptions about structure-property relationships as they reasoned and (2) viewed spectral data with different gaze patterns, both of which suggest different approaches to evaluating spectra. Topczewski *et al.* also used eye tracking to investigate interpretation approaches used by undergraduates and graduate students, in particular the approaches used to match organic molecules to appropriate ^1H NMR spectra.¹⁴ This study also found differences in gaze patterns between less advanced and more advanced

participants, implying different approaches to interpreting spectra between the two groups. These studies provide a useful foundation for understanding how individuals learn to interpret spectral data. In addition, the range of spectral interpretation ability demonstrated by participants in each study suggests that such ability develops along a progression; if stages of this progression can be mapped, then instruction can be designed to cultivate students' ability to interpret spectral data.

In addition to the aforementioned studies focusing on learning IR and ^1H NMR spectroscopy, research by Connor and Shultz investigated teaching assistants' knowledge for teaching ^1H NMR spectroscopy.¹⁵ They found that as the fraction of teaching experience in ^1H NMR spectroscopy increased relative to other courses, teaching assistants' knowledge for teaching this topic also increased, irrespective of general teaching experience or organic chemistry research experience. This finding is promising given that it provides insight into how other instructors may cultivate knowledge for teaching spectroscopy and in turn improve instruction quality and student learning outcomes. However, additional research focusing on students' reasoning during spectral interpretation is needed in order to inform instructor education and thus expedite the development of this knowledge.

2.4 Theoretical framework

2.4.1 Categorization of mental representations

Research in cognitive psychology suggests that as individuals interact with an entity (e.g., an object, event, state, person, idea, etc.) they construct a mental representation of this entity that corresponds to a given category; it is this categorization that facilitates the organization of experiences.^{16,17} Researchers argue that category membership is not simply determined by the surface similarities of entities, but rather by underlying knowledge structures or theories in which the representations are fixated.¹⁸ Future reasoning is then guided by the assumptions one has about the properties and behaviour of entities belonging to a category.¹⁹ However, because categories are defined and

governed by underlying knowledge structures, these assumptions may also serve to constrain reasoning.²⁰

Findings from discipline-based education research (DBER) illustrate how underlying knowledge structures influence the categorization of mental representations. For example, Chi *et al.* found differences in the categorization of physics problems between experts and novices, as well as differences in the knowledge associated with these categories.²¹ Novices in this study tended to group problems using explicit features provided in the problem, whereas experts generated groupings based on relevant physics principles. Further, Galloway *et al.* found differences in the categorization of organic chemistry reactions between organic chemistry students and professors; students in this study tended to group reactions using surface features, whereas professors generated categories solely for process-oriented reasons.²² Stains and Talanquer also found differences in the categorization of chemical reactions at various stages of expertise, with undergraduate participants in this study tending to group reactions by surface features (e.g., “aqueous reactant”, “produce water,” etc.) and graduate students forming groups based on traditional, discipline-based reaction types (e.g., redox, acid-base, etc.).²³ Findings from these studies suggest that in order to design chemistry instruction that promotes expert-like thinking, the underlying knowledge structures which govern categorization should be characterized.

2.4.2 Assumptions and heuristics as cognitive constraints

A number of studies have investigated chemistry students' assumptions in order to characterize underlying knowledge structures in the discipline as well as gain insight into the cognitive elements that constrain reasoning.^{13,24,25} This literature defines assumptions as “presuppositions about the properties and behavior of the entities and phenomena in the domain”;¹⁹ these assumptions can range from being intuitive in nature, much like phenomenological primitives,²⁶ to learned principles. In addition to those aligning with scientifically accepted views, assumptions may also be inaccurate. Maeyer and Talanquer refer to assumptions that reflect inaccurate ideas about chemical entities as

spurious chemical assumptions; according to these authors, this class of assumptions often arises from misunderstanding or generalizing learned chemical principles.¹⁹

Assumptions about the nature of entities often work together with heuristic reasoning strategies during judgement and decision-making, especially when relevant background knowledge is lacking;²⁴ sets of these assumptions and heuristic reasoning strategies constitute the fundamental constraints to learning in a domain.²⁵ Heuristic reasoning strategies, or heuristics, are simplification and effort-reduction methods used by individuals to decrease the amount of information to process during decision-making.²⁷ According to the dual-process theory of cognition, reasoning is guided by two types of thinking: Type 1 and Type 2 processing.²⁸ The use of heuristics is associated with Type 1 processing, which tends to be fast, automatic, and independent of cognitive ability. This system is autonomous and does not require working memory.²⁹ Type 2 processing, on the other hand, tends to be slow, systematic, and dependent on cognitive ability. This system requires working memory.²⁹

Day-to-day decisions are often facilitated by heuristic reasoning strategies associated with Type 1 processing; it is this type of thinking that allows us to complete common tasks without excessive cognitive load. Further, expertise in a field is not characterized by a lack of heuristic reasoning but rather the effective use of heuristics in appropriate contexts.¹⁹ However, their use may also result in cognitive bias and errors,³⁰ as demonstrated by a number of studies that investigated the use of heuristics in.^{31–33} Chemistry students' use of heuristics, in addition to their assumptions, therefore merits investigation. By characterizing these cognitive constraints, instruction can be designed to target this reasoning and in turn promote more sophisticated ways of thinking. In addition, the characterization of cognitive constraints at various stages of learning in an area can assist in the construction of a learning progression that guides instruction and supports students' development of knowledge.^{25,34} This work provides insight into the heuristics and invalid assumptions which may co-occur with valid assumptions at the lower anchor of a learning progression on spectral interpretation.

2.5 Research questions

In an effort to characterize constraints on organic chemistry students' reasoning during IR and ^1H NMR spectral interpretation, this study addressed the following research questions:

- (1) What invalid chemical assumptions or heuristic reasoning strategies (if any) do undergraduate organic chemistry students use when determining the success of a synthesis using IR and ^1H NMR spectra?
- (2) What invalid chemical assumptions or heuristic reasoning strategies (if any) most severely constrain organic chemistry students' reasoning when determining the success of a synthesis using IR and ^1H NMR spectra?

For this investigation, invalid chemical assumptions included what Maeyer and Talanquer define as spurious chemical assumptions, or "invalid ideas about the properties of chemical entities or reactions, often resulting from misinterpretations and overgeneralizations of chemical principles."¹⁹ In addition, invalid chemical assumptions included assumptions reflecting any intuitive knowledge that contradicts scientifically accepted principles.

It is important to note that students may also hold scientifically accurate assumptions which guide their thinking along productive avenues toward expertise. If these assumptions are characterized, instruction can build upon them to further cultivate sophisticated thinking. In addition, these assumptions may also restrict thinking given that individuals may rely on them to provide local explanatory coherence rather than achieve conceptual coherence.²⁵ For these reasons, chemistry education studies have typically investigated students' valid assumptions in addition to their invalid assumptions.^{19,25} This study is limited to the investigation of invalid assumptions in order to provide a detailed account of any findings while also maintaining their accessibility. Providing a richly detailed, accessible description of findings will ensure their utility and allow for assessments of transferability. Further, by investigating invalid assumptions and heuristics we can first identify any potential significant barriers to analytical thinking, as

well as gain some insight into how students may use heuristics effectively. This study thus serves as a productive *initial* step in the process of understanding students' reasoning surrounding this complex practice. Characterization of the range of conceptual sophistication demonstrated by organic chemistry students during spectral interpretation, which includes both scientifically accurate and inaccurate assumptions, will be the focus of future work.

2.6 Methods

2.6.1 Sample and setting

Eighteen undergraduates from a large, public Midwestern university participated in the study. Seventeen undergraduates were enrolled in an organic chemistry II laboratory course at the time of data collection, and one undergraduate had completed the course in a prior semester. Participants were recruited from four sections of the course. Three sections followed a traditional design and were a combination of chemistry majors and non-majors, with each section taught by a different instructor. The fourth section followed a course-based undergraduate research experience (CURE) design and was a combination of chemistry majors and non-majors, with this section taught by one of the three instructors mentioned above.³⁵ Eleven students were recruited from the traditional sections, and six students were recruited from the CURE section. Participants were recruited via email and in-class announcements by the first author. The undergraduate who completed the course in a prior semester was recruited via snowball sampling.³⁶ Of the 18 undergraduates that volunteered to participate, all were interviewed; the study population was therefore a convenience sample. Responses from the undergraduate who completed the course in a prior semester did not noticeably differ from that of the larger sample population, so they were included in subsequent data analysis. Of the 18 participants, there were nine males and nine females. Participants represented a variety of ethnicities, which is a general representation of the larger student population at this institution. All individuals voluntarily consented to participate in the study and IRB approval was obtained.

IR and ^1H NMR spectroscopy were taught in detail in the organic chemistry II laboratory course at the institution in which the study took place. Instructors of each section covered content relevant to the course in a weekly one-hour laboratory lecture. IR spectroscopy was covered in this lecture during Weeks 4 and 6 of the semester, and ^1H NMR spectroscopy was covered during Weeks 8, 11, and 12. As part of instruction on IR spectroscopy, students were taught to (1) identify the main components of an IR spectrum (e.g., peak characteristics, units, and regions); (2) identify major functional groups and bonds in the functional group region; (3) interpret authentic spectra collected in lab and identify if it corresponds to a product, starting material, or reaction solvent; (4) match a set of compounds to the appropriate IR spectra; (5) use an IR spectrum along with other characterization methods to identify unknown compounds and (6) use an IR spectrum, with and without other information, to predict molecular structure. As part of instruction on ^1H NMR spectroscopy, students were taught to (1) interpret features of a spectrum (e.g., number of peaks, peak position, integration, first-order splitting, some second-order splitting, and splitting of OH and NH hydrogens) to determine molecular fragments or the complete structure of a compound; (2) interpret coupling constants and use them to differentiate between structural isomers; (3) match a set of compounds to appropriate NMR spectra; (4) use an NMR spectrum, with and without other information, to predict molecular structure; (5) use an authentic NMR spectrum, along with an IR spectrum and thin layer chromatography, to identify an unknown compound and (6) compare an authentic NMR spectrum obtained in lab to spectral data from the literature. Students were also provided with a coursepack containing optional practice problems involving IR and ^1H NMR spectral interpretation.

2.6.2 Data collection

For this investigation, we used cued retrospective think-aloud (RTA) interviewing to collect qualitative data on students' reasoning during spectral interpretation. Cued RTA interviewing is a qualitative technique paired with eye tracking to characterize individuals' thinking,³⁷ where eye tracking involves measuring individuals' eye movements as they complete a visual-based task.³⁸ Cued RTA interview protocols involve participants watching a recording of their eye movements following the completion of a visual-based

task and verbalizing in as much detail as possible what they were viewing and thinking as they completed the task. van Gog *et al.* demonstrated that cued RTA interviewing is an effective tool for eliciting problem-solving process information when compared to concurrent think-aloud reporting and standard retrospective reporting.³⁹ In addition, cued RTA interviewing is a particularly well-suited tool for investigating individuals' thinking as they complete a complex task such as spectral interpretation because it allows participants to work on their own and in silence, as opposed to concurrent think-aloud interviewing which requires participants to verbalize their thoughts in-the-moment and thus increase their cognitive load.

Each participant took part in one 30-60 minute session in which they completed three spectral interpretation tasks while having their eye movements tracked. Following the completion of each interpretation task, participants completed a semi-structured, cued RTA interview in which they watched a recording of their eye movements and described in detail what they were focusing on and thinking about during each task.⁴⁰ Cued RTA interviews were conducted using the Tobii Studio 3.4.8 RTA feature, which allows for simultaneous audio recording and playback of Tobii Studio eye movement recordings.⁴¹ Data collected in this investigation included audio-visual recordings of RTA interview responses and information relating to participants' research experience interpreting IR and ¹H NMR spectra. All data was collected during a four-week period following instruction on ¹H NMR spectroscopy. Data collection continued during this period until 18 individuals had participated; participants expressed no new reasoning at this point, indicating data saturation was achieved.³⁶

Prior to the start of each session, participants were given an overview of the task format. To ensure that the description of the task was interpreted as intended, participants were asked to describe the task and its objective in their own words. Participants were then given an explanation of the cued RTA interview protocol and informed that each interpretation task would be followed by a cued RTA interview. Participants were not time-restricted as they completed interpretation tasks, and all tasks were presented in a randomized order. Prior to the start of each RTA interview, participants were informed that they could pause the recording of eye movements at any time during the interview if

they needed more time to speak. The interviewer was also able to pause the recording of eye movements in order to further probe students' reasoning. Audio-visual recordings of RTA responses included a video screen capture of eye movement recordings viewed by participants during the RTA interview overlaid with an audio narration of their verbalized thoughts.

2.6.3 Description of interpretation tasks

The three interpretation tasks included in this study were of an identical format (Figure 2.1). Each task included a short prompt explaining that chemists first attempted to synthesize a given compound and then analysed their final product spectroscopically to determine if the synthesis was successful. The prompt then instructed participants to determine if the synthesis of the desired product was successful using the provided spectroscopic data (IR and ^1H NMR spectra). This problem type was selected given that determining the outcome of a synthesis using spectroscopic data aligns with the common day-to-day problems of practicing organic chemists;⁴² by incorporating authenticity into the tasks, any findings may more meaningfully inform classroom instruction.

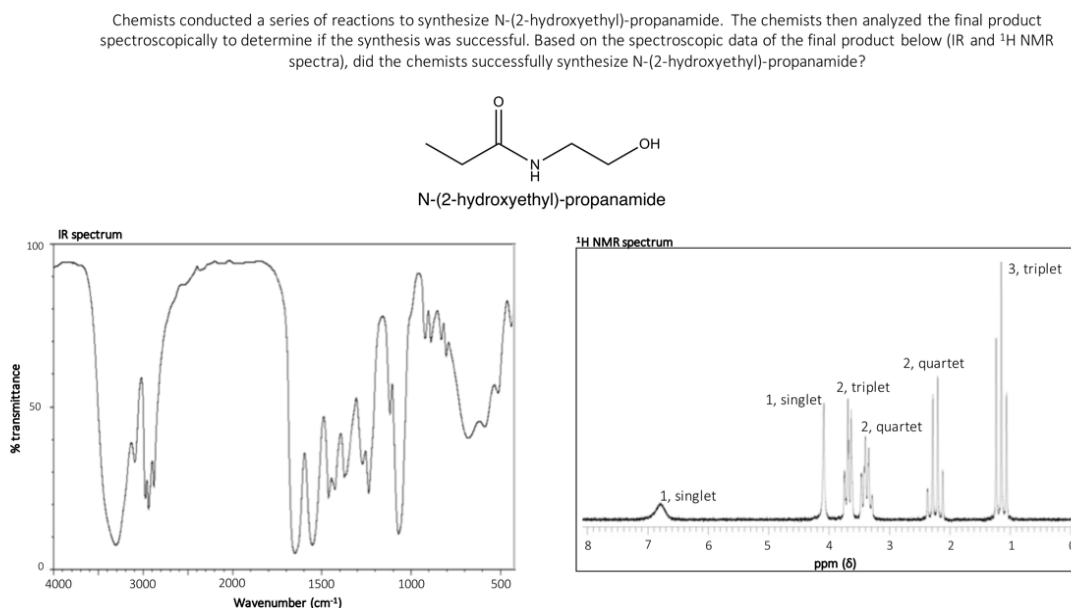


Figure 2.1. Spectral interpretation task (Synthesis 1) asking participants to determine if N-(2-hydroxyethyl)-propanamide was successfully synthesized using the provided IR spectrum and ^1H NMR spectrum.

Tasks were labelled as Synthesis 1 (Figure 2.1), Synthesis 2 (Supporting Information, Figure 2.4), and Synthesis 3 (Supporting Information, Figure 2.5). Molecules corresponding to the desired product of each synthesis are provided in Figure 2.2. All spectra were obtained from the Spectral Database for Organic Compounds and were free of signals due to solvent or impurities;⁴³ these spectra were selected in order to reduce participants' cognitive load and allow for the completion of already complex tasks. All spectra are reproduced herein with permission from SDBSWeb. Integration values and multiplicities were included on all ¹H NMR resonances given that the authors wished to investigate students' reasoning and not their ability to distinguish between individual peaks. The labels served to further reduce participants' cognitive load. In addition, reference tables containing characteristic IR absorption values and ¹H NMR chemical shift values⁴⁴ were included with each task given that content knowledge can act as confounding variable during task completion.⁴⁵ These tables also mirrored resources available to students when interpreting spectra in the context of the course. After completing the task, participants could respond with "yes, the product was synthesized", "no, the product was not synthesized", or "not enough information to tell."

As part of a larger study, a faculty member with more than ten years of teaching experience in IR and NMR spectroscopy was interviewed to provide insight into ¹H NMR spectral features that often create difficulty for undergraduates. Molecules with ¹H NMR spectra that included these potentially difficult spectral features were incorporated into tasks for this investigation in order to increase the likelihood of eliciting invalid chemical assumptions and problematic heuristic reasoning strategies among participants. For

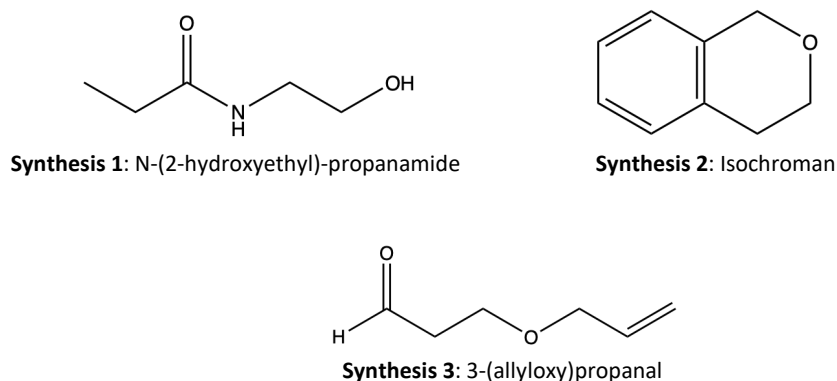


Figure 2.2. Compounds identified as desired products in each interpretation task.

Synthesis 1, participants evaluated the synthesis of N-(2-hydroxyethyl)-propanamide (Figure 2.2). The provided IR and ^1H NMR spectra corresponded to this molecule (Figure 2.1) and the correct response to Synthesis 1 was “yes, the product was synthesized.” This molecule was selected because it contains an amide functional group that results in splitting patterns that deviate from the “N+1 rule,” a feature the consulted faculty member identified as difficult for students. Participants in this study received classroom instruction on deviations from the “N+1 rule,” however the authors hypothesized that such features may still pose difficulty for participants. For Synthesis 2, participants evaluated the synthesis of isochroman (Figure 2.2). The provided IR and ^1H NMR spectra corresponded to isochroman (Supporting Information, Figure 2.4) and the correct response to Synthesis 2 was “yes, the product was synthesized.” Isochroman was selected because its ^1H NMR spectrum contains overlapping signals resulting from aromatic hydrogens, another feature the consulted faculty member identified as difficult for students. Participants had also received classroom instruction on this phenomenon. Lastly, students evaluated the synthesis of 3-(allyloxy)propanal for Synthesis 3 (Figure 2.2). The provided IR and ^1H NMR spectra corresponded to 3-allyloxypropionic acid (Supporting Information, Figure 2.5) and the correct response to Synthesis 3 was “no, the product was not synthesized.” This molecule was selected given that it contains a variety of functional groups, which the authors hypothesized would increase difficulty.

Prior to the study, each task was piloted with four undergraduates having recently completed the organic chemistry II laboratory course in order to ensure that the prompt was interpreted as intended and that no task was too easy or too difficult for the study population. No task received either all correct or incorrect responses, suggesting that they were of an appropriate level of difficulty and that participants' inability to make annotations did not inhibit their ability to interpret spectra. Four interpretation tasks were initially developed and piloted, however pilot study members reported fatigue after the third task, so only three tasks were included in the study.

2.6.4 Qualitative analysis of RTA interviews

A mixed-methods approach with a conversion design was used to investigate invalid chemical assumptions and heuristic reasoning strategies that constrained undergraduates' reasoning during spectral interpretation.³⁶ RTA interviews were first analysed qualitatively to identify invalid chemical assumptions and heuristics used by participants. Frequencies of responses containing assumptions and heuristics were then analysed quantitatively to identify any assumptions or heuristics that most severely constrained participants' reasoning.

RTA interviews were transcribed verbatim, and audio-visual recordings of the interviews were used when necessary to clarify any ambiguous references to spectral data in the transcripts. The first author inductively coded all transcripts for invalid chemical assumptions. During this process, the author generated *in vivo* codes and descriptive codes that corresponded to specific invalid ideas about chemical and spectral features expressed by participants.^{46,47} Codes and definitions were then refined in order to combine closely related invalid ideas into single codes. The second author then deductively coded all transcripts using the revised codes and definitions as well as inductively coded transcripts to identify any invalid ideas not identified by the first author. To establish reliability of the coding, the first and second author then discussed and revised all codes until 100% agreement was reached. NVivo 11 software was used throughout the coding process.⁴⁷ After consensus was established, the first author then identified themes among the invalid chemical assumption codes using constant comparative analysis.⁴⁸ Written analytic memos and regular discussions with the second author facilitated the identification of themes.⁴⁷

Following this analysis, themes and their contributing invalid chemical assumptions were shared with five external experts to establish the validity and transferability of the findings to other interpretation tasks and instructional contexts. External experts were instructors of first and second-semester organic chemistry laboratory and lecture courses from four institutions of various types (one public doctoral-granting university in the Midwest, one public doctoral-granting university in Canada, one

private doctoral-granting university in the Midwest, and one private liberal arts college in the Midwest) who provided feedback on the perceived extent to which the identified themes constrain their own students' reasoning. All external experts cover IR and ^1H NMR spectroscopy in their respective courses. Experts had teaching experience in IR and ^1H NMR spectroscopy ranging from two years to nearly 20 years.

To code for heuristic reasoning strategies, the first author developed an initial list of heuristics and corresponding definitions using existing literature on heuristic reasoning in chemistry.^{19,30,49} The first author then deductively coded all transcripts using this list. Any heuristics that did not appear in responses were then removed from the initial list, and definitions of remaining heuristics were revised in order to clearly operationalize each heuristic. The first and second author then deductively coded all transcripts using the refined list. The authors then discussed and revised all codes until 100% agreement was reached.

2.6.5 Quantitative analysis of assumptions and heuristics

After the first and second authors reached consensus for all codes, frequencies of responses containing given invalid chemical assumptions and heuristic reasoning strategies were tabulated for each interpretation task. For this tabulation, invalid chemical assumptions were grouped into previously identified themes. In order to investigate if the use of certain assumptions or heuristics was task-specific, a two-sided Fisher's exact test and a Pearson χ^2 test of independence were used to determine if certain interpretation tasks disproportionately elicited specific invalid chemical assumptions or heuristic reasoning strategies. A Pearson χ^2 test of independence was used for the analysis of heuristic frequency distributions between tasks, whereas a Fisher's exact test was used for the analysis of assumption frequency distributions between tasks. A Fisher's exact test was used for the later analysis because the total number of assumptions did not meet the minimum requirements for the Pearson χ^2 test of independence, and a two-sided Fisher's exact test is recommended in lieu of a Pearson χ^2 test of independence when the total number of observations is less than 20.⁵⁰ Statistical significance was set at 0.05

for all significance testing. All statistical analyses were completed using the R Stats Package in RStudio Version 1.1.453.⁵¹

In order to identify assumptions or heuristics that most severely constrained organic chemistry students' reasoning, one-sided Fisher's exact tests were used to determine if certain invalid chemical assumptions or heuristic reasoning strategies appeared in incorrect responses significantly more than they appeared in correct responses. A one-sided test is used in lieu of a two-sided test when frequencies are expected to be greater for a given group.⁵⁰ Incorrect responses involved incorrectly determining the success of given syntheses (e.g., selecting "no, the product was not synthesized" when the IR and ¹H NMR spectra corresponded to the target molecule), whereas correct responses involved correctly determining the success of given syntheses (e.g., selecting "yes, the product was synthesized" when the IR and ¹H NMR spectra corresponded to the target molecule). Responses in which the "not enough information to tell" option was selected were omitted from this analysis. Odds ratios were evaluated *post hoc* as a measure of effect size for assumptions and heuristics that appeared relatively more in incorrect responses. The Haldane-Anscombe correction was used for the determination of odds ratios due to some frequencies equalling zero.⁵² Small, moderate, and large effects corresponded to odds ratios equalling 1.68, 3.47, and 6.71, respectively.⁵³ While some correct responses exhibited constraints on reasoning (e.g., invalid assumptions that either did not influence participants' decisions or resulted in participants responding correctly "for the wrong reasons"), incorrect responses represent an extreme case of constrained reasoning; identification of invalid assumptions and heuristics common to incorrect responses may thus provide insight into the cognitive elements that most severely constrain organic chemistry students' reasoning during spectral interpretation.

2.7 Results and discussion

Of the 18 participants, only five correctly determined the success of all three syntheses. Nearly half of participants ($n = 8$) correctly determined the success of two of the three syntheses, and four participants correctly determined the success of only one synthesis.

Lastly, one participant did not correctly determine the success of any synthesis. This distribution further suggests that interpretation tasks were of an appropriate level of difficulty for this population and that undergraduates' reasoning during spectral interpretation merits investigation. Synthesis 1 and 2 appeared to be of equal difficulty, with six incorrect responses to each of these tasks. Synthesis 3 appeared to be slightly less difficult, with only four incorrect responses to this task. Further, each synthesis had one response indicating "not enough information to tell."

2.7.1 Qualitative findings: invalid chemical assumptions

Through the inductive coding of RTA interview responses, we identified a total of 20 unique invalid chemical assumptions. Of these assumptions, 12 related to ^1H NMR spectroscopy, 5 related to IR spectroscopy, and 3 related to molecular structure (Table 2.1). From these 20 coded assumptions, we identified five themes that more comprehensively explain the invalid chemical assumptions that constrained students' reasoning during spectral interpretation (Table 2.1). Invalid chemical assumptions were related to specific spectral features included in this study. However, both the use of interpretation tasks that incorporated a variety of spectral features and the identification of common themes contribute to the transferability of our findings to other IR and ^1H NMR spectral interpretation tasks for this study population. Validation of these themes by external experts further contributes to the transferability of these findings to other interpretation tasks, as well as to other instructional contexts. These themes are described in detail below.

Theme I: Assumptions that the "N+1 rule" should hold. Each ^1H NMR spectrum included in this study incorporated one spectral feature for which the "N+1 rule" fails to hold. The "N+1 rule" is a guideline commonly included in ^1H NMR instructional materials for determining signal multiplicity,⁴⁴ however a number of exceptions to this "rule" exist. A majority of interviewees ($n = 13$) incorrectly indicated that the failure of the "N+1 rule" to hold was problematic and that this failure suggested that given syntheses were unsuccessful (Table 2.1). For Synthesis 1 (Figure 2.1), four students regarded the

Table 2.1. Invalid chemical assumptions that constrained students' reasoning and themes among these assumptions. n-values in the third column correspond to the number of participants (n=18) who used specific assumptions at least once. n-values under each theme correspond to the number of participants with responses contributing to each theme.

Theme	Contributing invalid chemical assumptions	n	Synthesis
Assumptions that the "N+1 rule" should hold (n = 13)	• NH and/or OH should not appear as singlets	5	1
	• CH ₂ groups between NH and OH should appear as quartets	4	1
	• The aromatic ring has too few corresponding NMR peaks	6	2
	• Double bonds should obey the "N+1 rule"	5	3
Assumptions that spectral data should be absolute (n= 9)	• IR peaks should be prominent if the functional group is present	9	1, 2, 3
	• Chemical shift values should match the reference material	1	2
	• The number of chemically equivalent hydrogen sets should match the number of peaks	1	2
	• A messy IR spectrum suggests an unsuccessful synthesis	1	3
Visuospatial invalid assumptions (n = 8)	• Isochroman is symmetric	7	2
	• Incorrect number of hydrogen atoms attached to methylene and methine carbon atoms	1	2
Practical invalid assumptions (n = 8)	• There is an IR peak corresponding to a halogen functional group	2	1, 3
	• The NH singlet corresponds to an artefact the NMR spectrometer "picked up"	1	1
	• The IR peak near 3000 cm ⁻¹ corresponds to the OH functional group	1	2
	• The broad IR peak near 3000 cm ⁻¹ corresponds to the CH functional group (n=4) or water (n=1)	5	3
Fundamental invalid assumptions (n = 6)	• Parts of a molecule vary in concentration	1	1
	• Incorrect splitting knowledge: connected hydrogen atoms determine multiplicity	1	1, 2
	• Incorrect splitting knowledge: multiplicity determined by absolute number of adjacent hydrogen atoms ("N") rather than "N+1"	1	2, 3
	• De-shielding causes a shift right	1	2
	• Oxygen nuclei generate ¹ H NMR signals	1	2
	• Doublets are part of doublet of doublets	3	3

appearance of singlets corresponding to OH and NH hydrogens as problematic, stating that these hydrogens should appear as triplets given their number of nearest neighbours. One of these students, Frances, explains how this deviation from the "N+1 rule" influenced her evaluation:

“And then, I basically, I concluded that both of the single hydrogens that were on the alcohol and on the NH, they did have neighboring hydrogens next to them and because of that, because of that, those peaks couldn't be singlets. My reasoning for the question.”

One additional student regarded the singlet corresponding to the NH hydrogen as problematic yet recognized that signals corresponding to OH hydrogens do not always undergo splitting. Further, four students correctly paired corresponding singlets to the NH and OH hydrogens, however they stated that the two hydrogen groups between these functional groups should appear as either two triplets or two quartets (and not as one quartet and one triplet, as they appear in the spectrum). All six participants who incorrectly determined that Synthesis 1 was unsuccessful relied on one of these invalid assumptions during their reasoning.

Invalid chemical assumptions that contribute to Theme I also appeared in RTA responses for Synthesis 2 and Synthesis 3 (Supporting Information, Figure 2.4 and 2.5). For Synthesis 2, nearly-equivalent aromatic protons in isochroman give rise to one multiplet and one doublet rather than two doublets and two triplets as the “N+1 rule” indicates. One-third of students (n = 6) incorrectly deduced that the actual splitting pattern may imply an unsuccessful synthesis. Again, Frances explains how this deviation from the “N+1” influenced her evaluation:

Frances: So, I think again I was double checking using the “N+1 rule” just trying to figure the different environments. And then counting in my head, just trying to see the different environments.

Interviewer: Did the “N+1 rule” seem to be checking out for you?

Frances: I believe that on the right part of the molecule it was working on it. At least on the left part of the molecule, when I was analyzing it, there didn't seem like there was ... because on the spectrum it's listed in the three in one, which I didn't really... to me didn't make any sense just because it didn't seem like there was any way, at least in mind, to quantify that.

Of the six participants that incorrectly determined that Synthesis 2 was unsuccessful, five relied on this invalid assumption during their reasoning, further suggesting that students' reasoning is constrained by the notion that the “N+1 rule” should generally hold.

Students' reasoning was further constrained by this notion in RTA responses to Synthesis 3. For this interpretation task, several students (n = 5) reasoned using the

incorrect assumption that the “N+1 rule” should apply to vinylic hydrogen atoms. One student, John, explains that the appearance of two doublets corresponding to terminal vinylic hydrogen atoms in 3-(allyloxy)propanal in part led him to question the success of this synthesis:

“...So I figured there are... too many integrations of one. And I figure that might not be right. Yeah there are too many ones and too many splits over here. It shouldn't be [so] many splits. I figured that is the wrong thing.”

As illustrated above, assumptions that the “N+1 rule” should hold appeared in all three interpretation tasks, providing some indication that these assumptions constrain students’ reasoning in a number of contexts. This result aligns with the finding by Cartrette and Bodner that unsuccessful participants were less flexible in their understanding of the “N+1 rule” when compared to successful participants, providing additional evidence of transferability to other interpretation tasks and instructional contexts.¹²

Theme 2: Assumptions that spectral data should be absolute. Students’ reasoning was further constrained by invalid assumptions that certain spectral data should be prominent or definite if corresponding molecular features are present in the synthesized product. Half of students (n = 9) incorporated such assumptions into their reasoning. These invalid chemical assumptions contribute to Theme 2 (Table 2.1), the notion that spectral data should be absolute if molecular features are present. The most prevalent of these assumptions was that IR peaks should be readily distinguishable if corresponding functional groups are present in the synthesized product, with half of students (n = 9) incorrectly identifying IR peaks of low intensity or overlapping IR peaks as evidence of unsuccessful syntheses. Audrey’s response illustrates how this assumption influenced her reasoning for Synthesis 3:

“That stretch around [1700 cm⁻¹] is kind of the combination of ... It's on the high end of the C double bond C stretch and the low end of the carbonyl stretch, I didn't like the fact that it was like one. And I was like, ‘You'd probably see something different.’ And so I think that in the end was what led me to say, ‘No.’”

A number of less prevalent assumptions also contributed to Theme 2. Similar to the previously described invalid assumption, one student, Stephen, incorrectly reasoned that

a complex IR spectrum in Synthesis 3 (Supporting Information, Figure 2.5) provided some evidence of an unsuccessful synthesis:

Stephen: *Yeah, it was really the IR that in the end made me decide no. I think I was just really confused by the NMR, so I ended up, yeah.*

Interviewer: *So maybe another IR due to contamination maybe? Or...*

Stephen: *Yeah, maybe. It just seemed - yeah, kind of messy to me.*

Interviewer: *Okay, so more "messy" than you normally see?*

Stephen: *Yeah.*

Another student incorrectly reasoned that Synthesis 2 was unsuccessful using the invalid assumption that ^1H NMR chemical shift values should exactly match values provided in the reference table. Lastly, one student identified a mismatch in the number of chemically equivalent hydrogen groups and ^1H NMR resonances in Synthesis 2 as evidence of an unsuccessful synthesis, further suggesting this notion constrained students' reasoning.

Theme 3: Visuospatial invalid assumptions. Students' invalid chemical assumptions relating to their visuospatial thinking also appeared to constrain reasoning (Table 2.1). A surprising number of students ($n = 7$) reasoned using the invalid assumption that isochroman possesses molecular symmetry in their response to Synthesis 2. One student, Madelyn, explains how this assumption influenced her reasoning:

"And, then I moved straight to NMR. See what I did. Here's what I counted, right off the bat, the peaks. The phenyl I counted wrong a bunch of times because of the symmetric. There should be two on the phenyl. Three on the other ring. Three. That lined up good with that."

Of the six students who incorrectly determined that Synthesis 2 was unsuccessful, four relied on this invalid assumption in their reasoning. No invalid assumptions relating to symmetry appeared in responses to Syntheses 1 and 3, likely because of the distinct asymmetry of the molecules in each corresponding task. Chemistry students' difficulty with visuospatial thinking is widely reported in chemistry education literature,^{54,55} however much of this literature relating to organic chemistry focuses on students' difficulty with forming three-dimensional mental images while visualizing two-dimensional molecular structures or performing mental rotation tasks. Students' inability to recognize the

asymmetry of isochroman (a task that does not require mental rotation) suggests that students' difficulty with visuospatial thinking in the context of organic chemistry may extend to less complex visualization tasks. Further, this difficulty may serve to constrain students' reasoning during spectral interpretation.

Theme 4: Practical invalid assumptions. Nearly half of students ($n = 8$) reasoned using invalid chemical assumptions that likely arose from a lack of practical experience interpreting spectral data (Table 2.1). These invalid assumptions most commonly took the form of students incorrectly identifying characteristic IR peaks. For Synthesis 3, four participants incorrectly identified the broad IR peak near 3000 cm^{-1} , a peak characteristic of the OH group of a carboxylic acid, as corresponding to the CH functional group. IR peaks corresponding to the CH functional group are notably less broad and intense than those corresponding to this OH.⁵⁶ Of the four students who incorrectly stated that Synthesis 3 was successful, two reasoned using this assumption. Similar to this assumption, Nancy incorrectly reasoned that Synthesis 2 was unsuccessful after misidentifying an IR peak corresponding to the CH functional group as belonging to the OH group of a carboxylic acid. She explains how this assumption influenced her reasoning:

"I know a broader peak around $3000\text{ [cm}^{-1}]$ usually corresponds to an OH, and I didn't see that [in the molecule]. I know that that's not necessary, but that peak kind of looks like what I've seen before with an OH. I didn't see that.... It was really that three [corresponding to integration of the multiplet NMR peak], I think, that I was basing my decision off of. The three and then this peak here [the IR peak near 3000 cm^{-1}]."

A small number of students ($n = 2$) also incorrectly reasoned that Syntheses 1 and 3 were unsuccessful due to the presence of apparent IR peaks corresponding to a halogen functional group in the fingerprint region of the spectra. Further, one student rationalized the presence of the unexpected broad singlet corresponding to the NH hydrogen in Synthesis 1 as an artifact that the spectrometer detected. This student, Robert, explains his reasoning below:

"Yeah, so I went back to the NMR, because I was really stuck on the NH being a singlet. And it was concerning to me that it was only a singlet and it was so small. And it wasn't a

real peak, I guess, it was more like it was just kind of a small thing that was picked up by the machine.”

Robert’s notion that the spectrometer can detect phenomena other than the absorption of electromagnetic energy by hydrogen nuclei further contributes to the practical invalid assumptions that constrained students’ reasoning.

Theme 5: Fundamental invalid assumptions. The last class of assumptions that constrained students’ reasoning were fundamental misunderstandings about basic NMR principles ($n = 6$, Table 2.1). The most common of these fundamental invalid assumptions ($n = 3$) was that the two doublets corresponding to each vinylic hydrogen atom in Synthesis 3 comprised a set of doublet of doublets. In addition, one student reasoned using the assumption that the number of hydrogen atoms on a carbon atom (rather than the adjacent carbon atoms) gives rise to a signal’s splitting pattern. Other assumptions held by individual students were (1) specific parts of a molecule may vary in concentration and result in unexpected peaks, (2) de-shielding causes a shift right rather than left on the NMR spectrum, (3) oxygen nuclei give rise to ^1H NMR signals, and (4) multiplicity is determined by the absolute number of adjacent hydrogen atoms (“N”) and not the number of adjacent hydrogen atoms plus one (“N+1”).

External expert validation. All experts stated that the themes accurately reflected problematic reasoning they have encountered among their own students. However, two of the five experts stated that the identified themes did not capture invalid chemical assumptions that are problematic during the interpretation of more authentic ^1H NMR spectra in undergraduate laboratory courses (e.g., spectra containing solvent peaks, unlabelled multiplicities, or raw integration values). These assumptions were not captured due to the format of the interpretation tasks and are discussed in the Limitations section below. Further, one expert stated that their students often invalidly assume that aldehyde hydrogens (in reference to Synthesis 3) and OH and NH hydrogens (in reference to Synthesis 1) should always appear as singlets. A number of participants in this study assumed that aldehyde hydrogens, as well as OH and NH hydrogens, appear as singlets; however, it is unclear from the interview data if participants assumed these hydrogens *always* appear as singlets. In addition, students received instruction explaining that

aldehyde hydrogens often appear as singlets due to a combination of low instrument resolution and the small coupling between aldehyde hydrogens and hydrogens on adjacent carbons.⁵⁶ Such reasoning was therefore not coded as an invalid assumption given that it would require significant inference by the authors. Nevertheless, these codes would fall under Theme 2 (i.e., assumptions that spectral data should be absolute) and thus do not discredit this study's findings. The inability to identify any problematic reasoning that required inference is further discussed in the Limitations section below.

Lastly, two of the five experts stated that they had not observed a small number of specific invalid chemical assumptions among their own students. These assumptions included participants' notion that parts of a molecule can vary in concentration, that IR peaks should be prominent if the functional group is present, and that halogen peaks were present in certain IR spectra. However, these experts explained that they may not have observed these assumptions among their own students given that they have not asked questions that would elicit such reasoning.

2.7.2 Qualitative findings: heuristic reasoning strategies

Through the deductive coding of heuristic reasoning strategies, we identified eight heuristics used in at least 20% of responses (Table 2.2). All participants used at least one heuristic in each response, though the way participants used them varied with individual and context. Talanquer divides common heuristics used by chemistry students for judgement and decision-making into three general groups: (1) fundamental associative processes, (2) inductive judgements, and (3) affective judgements.⁴⁹ Rather than present a list of heuristic strategies used by our participants, we aim to demonstrate how individuals used heuristics from each of these groups as they evaluated the success of syntheses via spectral interpretation.

Fundamental associative processes. Participants most commonly employed processing fluency and associative activation when evaluating each synthesis, both of which fall under the category of fundamental associative processes (Table 2.2).

Table 2.2. Heuristics identified in at least 20% of RTA interview transcripts, corresponding definitions, the number of participants who used corresponding heuristics at least once (n = 18), and total number of responses containing each heuristic (N = 54).

	Heuristic	Description	Participants	Responses (Total)
<i>Fundamental associative processes</i>	Processing fluency	Readily making sense of any salient molecular or spectral features	18	52
	Associative activation	Associating one observed spectral or molecular feature with a corresponding feature	17	42
<i>Inductive judgements</i>	Generalization	Overgeneralizing learned rules or patterns without considering all variables that may be involved	16	29
	Representativeness	Using some (but not all) spectral features to decide if entire spectra correspond to a molecule	15	18
	Reduction	Eliminating spectral features as information to process when alternative molecules share similar spectral features	14	20
	Rigidity	Using knowledge that has worked in the past and failing to consider other approaches	11	16
	One-reasoning decision making (ORDM)	Considering multiple spectral features while reasoning, but ultimately basing a decision on one spectral feature	10	11
<i>Affective judgements</i>	Affect	Experiencing positive or negative emotions evoked by spectral data	14	21

All participants used processing fluency in at least two out of three responses, and 17 out of 18 participants used associative activation in at least two of their responses. The prevalence of these heuristics is not surprising given that they often work together to support other heuristic reasoning.⁴⁹ Processing fluency refers to the ease with which an individual processes either explicit or implicit cues. In the context of spectral interpretation, use of this heuristic took the form of participants readily making sense of salient spectral and molecular features using either existing knowledge of such features or provided reference material. As in Robert's response to Synthesis 1 below, the heuristic often appeared at the beginning of participants' responses and focused more on explicit rather than implicit features:

“Okay, so I started by looking at the molecule and counting all of the hydrogens, and comparing it to the NMR, to look at the integration and the splitting again. I feel like that's the easiest way to start.... And so I saw that there's the right amount of integration values and it looks like they all correlate to the peaks as they should.”

In this example, Robert uses existing knowledge of chemical equivalency and its effect on the appearance of ^1H NMR signals to conduct what he deems as an easy, initial evaluation of the NMR spectrum. His focus on explicit features is unsurprising given that experts rather than novices tend to process implicit features more readily.⁴⁹

Where processing fluency refers to the ease with which information is processed, associative activation refers to the processing mechanism by which associations are automatically evoked through interaction with some stimulus.⁵⁷ For this study, associative activation took the form of participants either (1) associating a spectral or molecular feature with a corresponding characteristic feature using existing knowledge of such combinations or (2) explicitly connecting spectral features to those observed previously in instructional materials, laboratory, or other contexts. Associative activation therefore extends beyond readily processing spectral features using any source of information (i.e., existing knowledge regarding basic principles or reference material) to encompass the use of activated existing knowledge of spectral and molecular features. Ralph's response to Synthesis 1 illustrates this distinction. In his response below, he first observed the broad IR peak near 3300 cm^{-1} and correctly associated it with the OH functional group using existing knowledge of this combination. He then observed an IR peak near 1700 cm^{-1} and correctly associated it with the carbonyl functional group, a combination he observed previously in the laboratory:

“And so I saw there was an OH peak, I remember that. That was like one of the only peaks I remember, by memory.... But yeah, so this one and then also a carbonyl peak in this area. $1700\text{ [cm}^{-1}\text{]}$ area. I do remember that from lab. So that was kind of like, okay, those both match.”

Associative activation typically occurs alongside processing fluency given that the mechanism typically involves processing information with ease, and it is difficult to present evidence of this heuristic in isolation of the other.⁴⁹ Ralph's response illustrates both associative activation and processing fluency, as do all excerpts identified as associate

activation in this study; however, not all excerpts coded as processing fluency necessarily involve associative activation. Further, research on heuristic reasoning suggests that strongly activated information tends to disproportionately influence decision-making.⁵⁸ Associative activation can serve as an effective heuristic when used in correct contexts like in Ralphs's case above; however, it can be problematic if used inappropriately. For instance, Nancy incorrectly associated the IR peak near 3000 cm^{-1} with the OH functional group for Synthesis 2. As noted in the description of participants' practical invalid assumptions, this narrower peak actually corresponds to the CH functional group. Nancy then relied on this association and the target molecule's lack of an OH group to incorrectly determine that the synthesis was unsuccessful.

Inductive judgements. Five heuristics identified in participants' responses contribute to inductive reasoning; generalization, rigidity, representativeness, reduction, and one-reason decision making (ORDM). Participants applied these heuristics less frequently than fundamental associative processes, however their use was still prevalent. Participants most commonly used generalization, with this heuristic appearing in at least one out of three responses for 16 out of 18 participants. Generalization involves extending previously observed patterns or rules to potentially unfamiliar situations or contexts in order to make a judgement, and among novices in a field it tends to entail the over-extension of learned rules or principles.⁴⁹ In the context of this study, this heuristic often manifested as students' over-extension of the "N+1 rule" to given ^1H NMR signals without considering other variables such as amide bond coupling behaviour, the near-equivalent chemical environment of aromatic hydrogens, or the chemical inequivalence of vinylic hydrogen atoms. Other common generalizations involved claims that IR peaks should be prominent if functional groups are present, chemical shift values should match reference material exactly, and the number of chemically equivalent hydrogen atom sets should match the number of ^1H NMR signals. It should be noted that a number of invalid chemical assumptions (Table 2.1) resulted from generalization, however not all invalid assumptions were a product of this heuristic.

Rigidity is related to generalization and involves relying on problem-solving approaches that have worked in the past while failing to consider other strategies in new

contexts. Over half of participants (11 out of 18) used this heuristic in at least one response. When participants applied the rigidity heuristic, they most often relied on invalid chemical assumptions resulting from generalizations to make a final decision about the success of syntheses. Rigidity and generalization heuristics therefore often co-occurred, yet participants could still apply the generalization heuristic without being rigid in their generalization. Notably, participants that relied on both the generalization and rigidity heuristics tended to incorrectly determine the success of syntheses rather than simply question the success.

Participants also used the representativeness heuristic frequently, with over two-thirds of participants (15 out of 18) applying it in at least one response. This heuristic involves using easily processed information to determine whether an object belongs to a given class;⁵⁹ if the object is judged to belong, a decision is then made using properties of the class. In the context of this study, the representativeness heuristic involved participants evaluating a limited number of spectral features to determine if they corresponded to any of the many features they would expect from the target molecule. If participants found that selected features corresponded to some features expected from the target molecule, they judged their selected features to be an adequate representation of the expected spectra. They then made a decision about the synthesis using this judgement. When applying this heuristic, participants failed to evaluate one or more explicit spectral cues and ultimately stopped evaluating spectral data once they felt there was enough evidence to make a decision. Explicit spectral cues constituted prominent spectral features that provided some indication of each synthesis's success and that were evaluated by the majority of participants (e.g., the large IR peak corresponding to the OH functional group in Synthesis 3). This heuristic is useful given that it allows individuals to make a judgement when they lack necessary background information, as in Kim's correct response to Synthesis 1 below:

“And it looked good, like the numbers mostly worked out. I was going back over everything. Again, I'm not super confident on my ¹H NMR, but from what I knew it looked pretty good. So [I] decided that it had been synthesized.”

In this response, Kim judged the integration values and number of peaks in the NMR spectrum to be an adequate representation of the NMR spectrum that would be expected. She then used this judgement to correctly determine the synthesis's success. However, use of the heuristic became problematic when participants disregarded critical spectral features when determining representativeness, as in Chris's incorrect response to Synthesis 3:

“And then I just started looking again at the multiplet of integration value of one, because I didn't really know where that came from at first, and I still wasn't totally sure. I thought it could have something to do with the oxygen or proton transfer, but I wasn't really sure where it would come from, but I thought based on the other evidence I found that it was ... I don't know. I could place six out of seven peaks and the IR matched up close enough that I thought it was a good representation of the molecule.”

In Chris's response, he judged six out of seven ^1H NMR peaks and selected IR peaks to be an adequate representation of each expected spectra while disregarding spectral features corresponding to a carboxylic acid functional group. He then used this judgement to incorrectly determine the success of the synthesis.

Like the representativeness heuristic, the reduction heuristic also involves reducing the amount of information to be processed. More specifically, the reduction heuristic involves eliminating cues that are shared among alternative options as information to process.³⁰ When individuals used the reduction heuristic in this study, they explicitly chose to disregard certain spectral features they considered as either characteristic of more than one molecular feature or uncharacteristic of any particular molecular feature. For example, participants often chose to disregard absorption peaks in the fingerprint region of IR spectra as information to process given this region's complexity. Reduction of this information did not appear to inhibit participants' reasoning. However, use of the reduction heuristic became problematic when participants failed to recognize spectral features that were characteristic of a molecular feature and then eliminated them as information to process. For example, Shelia failed to recognize the IR peak characteristic of the OH functional group in Synthesis 3 (one indication of the unsuccessful synthesis of the target molecule) and then disregarded its presence to incorrectly determine the synthesis was successful:

"I'm looking at that 3,000 [cm⁻¹] peak, and I'm having a hard time piecing together what it might be. I think it might be an alkane, but it's not like a big functional group that we talked about a lot, like anything that's really special."

Participants' use of the reduction heuristic in both unproblematic and problematic ways aligns with the notion that although experts and novices both use heuristics, novices often lack knowledge of the appropriate contexts in which heuristics can be successfully applied.⁶⁰

One-reason decision making (ORDM) was the least frequently used heuristic in this group, with 10 participants using it approximately once; however, its use had a noticeable influence on participants' decision-making. This heuristic involves looking for one 'clever' cue, and then using only this cue to make a decision; its use may further entail the search for more than one cue, however a decision is made using a single feature.⁶¹ Participants in this study used ORDM as they assessed multiple spectral features during their evaluation but ultimately based their decision on just one feature. Use of the ORDM heuristic often involved participants relying on an invalid chemical assumption resulting from a generalization to make a decision, as Audrey's incorrect response to Synthesis 1 illustrates:

Audrey: And then I just decided they should all be quartets and if they weren't that wasn't what we had.

Interviewer: Okay. So it seems like you decided that before you looked at other pieces of information. What was your rationale for saying, "Okay, these don't match up. These should all be quartets. Let me look at the IR"?

Audrey: Yeah, so...I feel better about the IR kind of, and so if I could disprove it with the IR that would just like add to my confidence, I guess, with it.

Interviewer: Gotcha.

Audrey: But, then like going back I was like, "well they all do all have like those three [adjacent hydrogen atoms] so we should see quartets for all of them. So I decided that was good enough.

During this evaluation, Audrey searched for additional features in the IR spectrum that could provide some indication of the synthesis's success, however she based her decision only on an unexpected ¹H NMR splitting pattern. This example illustrates that ORDM becomes problematic when relevant background knowledge is lacking, or in this case with the co-occurrence of an invalid assumption.

Affective judgements. Only one heuristic identified in over 20% of participants' responses fell under the category of affective judgements (Table 2.2). This heuristic, termed affect, involves relying on one's positive or negative impressions to make decisions. This heuristic facilitates decision-making given that relying on one's impressions is often easier than systematically evaluating the weight of several cues, however such reasoning may result in illogical judgements.⁴⁹ A majority of participants (14 out of 18) applied this heuristic in at least one response, and its use took the form of individuals expressing positive or negative impressions of whether spectra corresponded with target molecules. In most cases, participants expressed positive or negative impressions about the data but the impact of such impressions on their decisions could only be inferred. For instance, in Shelia's incorrect response to Synthesis 2, she expressed confusion regarding the large ¹H NMR peak corresponding to hydrogen atoms on the benzene ring:

"...at the 7 ppm, that's kind of tripping me up. I'm looking more at the zoomed in version again. I just, that benzene ring is really tripping me up and it may be symmetrical, it may not be but still, the multiplet still doesn't make sense to me."

Whether this negative impression contributed to her incorrect decision is uncertain. However, the authors still coded such reasoning as the affect heuristic in order to over-estimate rather than under-estimate the influence of such impressions on decision-making. Nonetheless, some responses did demonstrate the direct influence of individuals' impressions on their decision-making. For example, Robert expressed that he felt positively about the provided spectral data for Synthesis 1 and used this feeling to inform his decision:

"I chose yes because I guess that I feel like it was there in a small amount.... So I could redo the NMR with a higher concentration to see if it was what I thought it was or not. And so I just kind of had a gut feeling that it was there."

As illustrated in the description of qualitative findings, participants used a number of invalid chemical assumptions and heuristic reasoning strategies during spectral interpretation. Invalid chemical assumptions and heuristics constrained participants' reasoning to various degrees, with some assumptions and heuristics appearing to result

in the incorrect determination of each synthesis's success and others having no obvious impact on decision-making. Further, use of some heuristics was problematic in certain contexts and productive in others, like the use of the representativeness heuristic by Kim and Chris as described above.

2.7.3 Quantitative findings: identification of invalid chemical assumptions and heuristics that most severely constrained reasoning

Following the qualitative analysis, frequencies of responses containing invalid chemical assumptions and heuristics were analysed quantitatively in order to (1) determine if particular interpretation tasks disproportionately elicited certain invalid chemical assumptions or heuristics and (2) identify assumptions or heuristics that most severely constrained participants' reasoning.

The extent to which interpretation tasks disproportionately elicited certain assumptions or heuristics was investigated in order to establish that participants' reasoning was not dictated by problem-specific features. Demonstrating that assumptions and heuristics were used with similar distributions across a variety of tasks serves as a means to establish transferability of any findings to other IR and ^1H NMR spectral interpretation tasks for this study population, in particular to tasks involving molecules with a similar variety of functional groups. Frequencies of responses containing invalid chemical assumptions and heuristics for each interpretation task are provided in the Supporting (Tables 2.4 and 2.5, respectively). A two-sided Fisher's exact test was used to determine if certain interpretation tasks disproportionately elicited certain invalid chemical assumptions. The distributions of assumption frequencies varied significantly with interpretation task ($p = 0.009$, two-sided Fisher's exact test); however, when visuospatial invalid assumptions were omitted from the analysis, distributions of assumption frequencies did not vary significantly ($p = 0.323$, two-sided Fisher's exact test). This lack of significance suggests that the tasks included in this study may only disproportionately elicit visuospatial invalid assumptions, as evinced by the exclusive appearance of these assumptions in responses to Synthesis 2 (Supporting Information, Table 2.4). Although a Fisher's exact test is used in lieu of a Pearson's χ^2 test of

independence when sample size is sufficiently small, a Pearson's χ^2 test of independence and *post hoc* residual analysis of invalid chemical assumption frequencies also revealed that the interpretation tasks only disproportionately elicited visuospatial invalid assumptions. Further, a Pearson's χ^2 test of independence was used to determine if certain interpretation tasks disproportionately elicited certain heuristics. The distribution of heuristic frequencies did not vary significantly with interpretation task ($\chi^2 = 9.03$, $p = 0.83$), indicating that interpretation tasks did not disproportionately elicit certain heuristics (Supporting Information, Table 2.5). The fact that interpretation tasks only disproportionately elicited visuospatial invalid assumptions and no other assumptions or heuristics provides additional evidence that findings from this study are transferable to other similar spectral interpretation tasks for this study population.

One-sided Fisher's exact tests were used to identify any assumptions or heuristics that tended to appear in incorrect responses and not in correct responses for each task. By identifying such assumptions or heuristics, the most problematic constraints on organic chemistry students' reasoning could be identified. P-values corresponding to all one-sided Fisher's exact tests are provided in Table 2.3. Significant p-values correspond to given assumptions or heuristics for which (1) the proportion of incorrect responses that used the assumption or heuristic is statistically greater than the proportion of incorrect responses that did not use the assumption or heuristic and (2) the proportion of correct responses that used an assumption or heuristic is statistically less than the proportion of correct responses that did not use the assumption or heuristic (Table 2.3). In other words, significant p-values correspond to assumptions or heuristics for which responding incorrectly was associated with whether the assumption or heuristic was used and responding correctly was associated with whether the assumption or heuristic was not used; significant p-values thus allow for identification of the assumptions and heuristics that tended to appear in incorrect responses and not in correct responses. Odds ratios were evaluated *post hoc* as a measure of effect size for statistically significant Fisher's exact tests (Table 2.3).⁵⁰ The odds ratio corresponds to the odds of using an assumption or heuristic and responding incorrectly versus the odds of using an assumption or heuristic and responding correctly. All odds ratios associated with significant p-values far

exceeded the criteria for a large effect size (odds ratios > 6.71),^{53,62} indicating a substantially large effect of using particular assumptions or heuristics on the ultimate accuracy of one's response. These large effect sizes were expected given that (1) a large effect would be necessary to result in statistically significant p-values for this relatively small sample size and (2) some assumptions and heuristics appeared exclusively in a majority of incorrect responses and not in correct responses.

For Syntheses 1 and 2, incorrect responses contained assumptions that the “N+1 rule” should hold, as well as use of the generalization, rigidity, and one-reason decision making (ORDM) heuristics at a statistically significant level (Table 2.3). In addition,

Table 2.3. P-values and odds ratios corresponding to one-sided Fisher's exact tests for assumptions and heuristics associated with incorrect responses. *corresponds to significance at the $p < 0.05$ level, **corresponds to significance at the $p < 0.01$ level, and ***corresponds to significance at the $p < 0.001$ level. †corresponds to odds ratios associated with significant p-values. All odds ratios associated with significant p-values far exceeded the criteria for a large effect size (odds ratios > 6.71). n/a^a corresponds to contingency tables with frequencies of zero in all cells. n/a^b corresponds to contingency tables in which all frequencies in one row or column were zero.

Assumptions and heuristics	Fisher's exact test p-values			Odds ratios		
	Synthesis 1	Synthesis 2	Synthesis 3	Synthesis 1	Synthesis 2	Synthesis 3
Assumptions						
Assumptions that the “N+1 rule” should hold	0.002**	< 0.001***	0.792	49.4†	84.3†	0.9
Assumptions that spectral data should be absolute	n/a ^a	0.001**	0.999	n/a ^a	41.4†	0.2
Practical invalid assumptions	0.999	0.353	0.099	0.5	6.3	7.0
Visuospatial invalid assumptions	n/a ^a	0.145	n/a ^a	n/a ^a	4.4	n/a ^a
Fundamental invalid assumptions	0.999	0.999	0.299	0.3	0.2	3.3
Heuristics						
Processing fluency	n/a ^b	n/a ^b	0.765	n/a ^b	n/a ^b	1.1
Associative activation	0.890	0.928	0.421	0.5	0.4	3.0
Generalization	0.017*	< 0.001***	0.985	21.7†	91.0†	0.2
Representativeness	0.999	0.999	0.015*	0.0	0.3	27.0†
Affect	0.841	0.841	0.999	0.7	2.4	0.1
Reduction	0.999	0.925	0.080	0.1	0.5	10.4
Rigidity	< 0.001***	< 0.001***	0.999	299.0†	84.3†	0.2
One-reasoning decision making	< 0.001***	0.029*	0.999	299.0†	23.0†	0.5

incorrect responses to Synthesis 2 contained assumptions that spectral data should be absolute (Table 2.3). The presence of *both* assumptions and heuristics at a statistically significant level aligns with research demonstrating that problematic reasoning among students is a product of multiple factors including heuristics and not simply misconceptions.³² This finding also aligns with research stating that *sets* of assumptions and heuristics constitute fundamental constraints to learning.²⁵ Incorrect responses to Synthesis 3 only contained the representativeness heuristic at a statistically significant level, possibly because of the smaller number of incorrect responses to this task. For this task, practical invalid assumptions and the reduction heuristic had p-values close to the 0.05 criteria for statistical significance (0.099 and 0.080, respectively), but it is uncertain if a larger number of incorrect responses would have resulted in significant p-values.

Further, assumptions and heuristics identified through statistical analysis did not appear in isolation from one another but rather co-occurred in incorrect responses, as Figure 2.3 illustrates. Combinations of assumptions and heuristics in responses to Syntheses 1 and 2 took on a similar form, where participants first expressed an invalid chemical assumption resulting from a generalization, demonstrated rigidity with respect to this generalization, and then ultimately made a decision using only the spectral feature which violated their invalid assumption. A number of excerpts provided in the Qualitative findings section illustrate these combinations, as does Frances' previously described response to Synthesis 1:

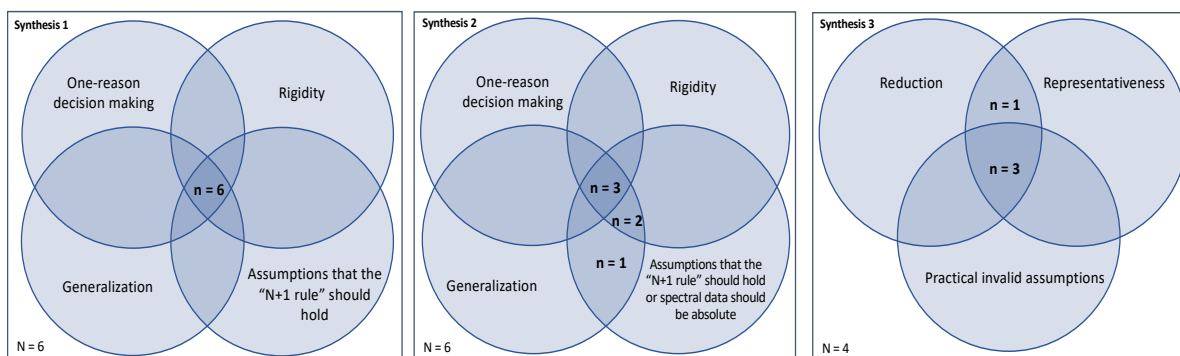


Figure 2.3. Co-occurrence of assumptions and heuristics in incorrect responses. The total number of incorrect responses (N) to a given task is indicated at the bottom of each corresponding Venn diagram. The number of incorrect responses containing particular assumptions and heuristics (n) is indicated in corresponding intersections of each diagram. Incorrect responses to Synthesis 2 contained both assumptions that the "N+1 rule" should hold and assumptions that spectral data should be absolute at statistically significant level; of these responses, two contained only assumptions that the "N+1 rule" should hold, one contained only assumptions that spectral data should be absolute, and three contained both types of assumptions.

“And then, I basically, I concluded that both of the single hydrogens that were on the alcohol and on the NH, they did have neighboring hydrogens next to them and because of that, because of that, those peaks couldn't be singlets. My reasoning for the question.”

As individuals make decisions, they typically identify cues and then assess their weight, or importance (Shah and Oppenheimer, 2008). Participants who held these invalid, rule-based assumptions possibly found spectral features which violated such assumptions to be accessible, highly weighted cues. The significant weight of such cues then potentially resulted in the use of ORDM and rigidity, both of which are effort-reduction heuristics, to facilitate decision-making.

Combinations of assumptions and heuristics in incorrect responses to Synthesis 3 took the form of participants making a practical invalid assumption regarding the large IR peak near 3000 cm^{-1} , subsequently reducing this IR peak as information to process, and then failing to evaluate other significant spectral data before making a decision. Madelyn's evaluation of Synthesis 3 illustrates this combination. Madelyn first correctly identified the IR peak near 3000 cm^{-1} as corresponding to the OH functional group:

“I started looking at the broadest [IR] peak first because usually I correspond it to an OH, so right off the bat I kind of felt that it wasn't getting synthesized.”

However, she then judged the NMR spectrum to be an adequate representation of the molecule while failing to evaluate chemical shift values that provided additional evidence of an OH functional group. She ultimately ended her evaluation by incorrectly rationalizing that the IR peak near 3000 cm^{-1} could instead correspond to the CH functional group:

“At this point I went back to the IR [spectrum] to maybe see if that peak [near 3000 cm^{-1}] could represent something else, cause usually that's an OH group. So I went back to the table and saw that also CH sometimes is a medium intensity. I was thinking maybe that would be a CH, especially if it's like an aldehyde....”

This invalid assumption facilitated Madelyn's reduction of the IR peak as information to process and, when combined with the representativeness heuristic used to evaluate the NMR spectrum, resulted in her incorrect response.

From the quantitative analysis, assumptions and heuristics that most severely constrained reasoning appear to vary somewhat with spectra and thus depend upon context. For two of the three interpretation tasks (Syntheses 1 and 2), incorrect response contained a combination of assumptions that the “N+1 rule” should hold and the generalization, rigidity, and ORD M heuristics at a statistically significant level. For one of these two tasks (Synthesis 2), assumptions that spectral data should be absolute also appeared in incorrect responses at a statistically significant level. For one of the three tasks (Synthesis 3), incorrect responses contained practical invalid assumptions and the reduction and representativeness heuristics, though only the representativeness heuristic appeared at a statistically significant level. As noted above, spectral interpretation tasks did not disproportionately elicit these assumptions or heuristics among the study population, suggesting that use of these particular assumptions and heuristics is a reflection of severely constrained reasoning and not reasoning used by all students. However, it is uncertain if there are *additional* combinations of such constraints given that the use of these combinations appears to depend somewhat on context. Additional investigations are needed to further characterize any additional severely constrained reasoning.

2.8 Conclusions

This study investigated constraints on organic chemistry students’ reasoning during IR and ¹H NMR spectral interpretation, in particular the invalid chemical assumptions and heuristic reasoning strategies used by students when evaluating the success of chemical syntheses using spectral data. A mixed-methods approach with a conversion design was used to first qualitatively characterize the invalid chemical assumptions and heuristic reasoning strategies used by study participants during spectral interpretation. Themes among invalid chemical assumptions were identified in order to more comprehensively characterize participants’ reasoning. Frequencies of responses containing given assumptions and heuristics were then analysed quantitatively to identify assumptions and heuristics that most severely constrained reasoning. Findings from both analyses provide insight into reasoning that may in part represent the lower anchor of a learning progression on spectral interpretation.

Findings from the initial qualitative analysis provide insight into which assumptions and heuristics organic chemistry students use during spectral interpretation. For this analysis, 20 invalid chemical assumptions were identified in participants' responses. Five themes emerged from these invalid chemical assumptions that more comprehensively illustrate constraints on participants' reasoning during these tasks: (1) assumptions that the "N+1 rule" should hold, (2) assumptions that spectral data should be absolute, (3) visuospatial invalid assumptions, (4) practical invalid assumptions, and (5) fundamental invalid assumptions. These themes were validated by external experts who provided insight into the validity and transferability of findings to other interpretation tasks and instructional contexts. Eight heuristic reasoning strategies were also identified during this initial qualitative analysis, all of which fall into one of three categories of heuristic reasoning described by Talanquer: (1) fundamental associative processes, (2) inductive judgements, and (3) affective judgements.⁴⁹ Heuristic reasoning strategies constrained participants' reasoning to various degrees, with some heuristics appearing to result in the incorrect determination of each synthesis's success (e.g., one-reason decision making and rigidity) and others having no obvious impact on decision-making (e.g., processing fluency). Further, use of some heuristics appeared problematic in certain contexts and supportive of correct decision-making in others (e.g., representativeness and affect).

The quantitative analysis of invalid chemical assumption and heuristic frequencies provided insight into cognitive elements that most severely constrained participants' reasoning. These assumptions and heuristics are those that tended to appear in incorrect responses and not in correct responses. While some constraints also appeared in correct responses (e.g., invalid chemical assumptions that ultimately did not influence decision-making), incorrect responses represent an extreme case of constrained reasoning. From this quantitative analysis, incorrect responses more often contained assumptions that the "N+1 rule" should hold, assumptions that spectral data should be absolute, and the generalization, rigidity, ORD, and representativeness heuristics when compared to correct responses. The prevalence of rule-based assumptions and effort-reduction heuristics in incorrect responses may have resulted from less engagement with optional practice problems in the provided coursepack, in particular since these problems often

included unexpected spectral features; however, we do not have data to support this claim. Further, these assumptions and heuristics tended to occur in combination with one another, as Figure 2.3 illustrates; in other words, the use of both assumptions *and* heuristics appears to result in incorrect responses rather than the use of assumptions or heuristics in isolation. This co-occurrence aligns with previous research on students' reasoning in chemistry which states that problematic reasoning is not just a collection of misconceptions but a combination of multiple factors including heuristics,³² and that sets of particular assumptions and heuristics constitute the fundamental constraints to learning.²⁵

2.9 Limitations

The design of interpretation tasks had inherent limitations. Integration values and multiplicities were provided for all ¹H NMR resonances in order to investigate students' reasoning and not their ability to distinguish between individual peaks. In addition, all provided spectra were free of signals resulting from solvent or impurities in order to reduce participants' cognitive load. By using these clean, labelled spectra, any potential invalid chemical assumptions relating to the interpretation of more authentic spectra (i.e., those containing unlabelled multiplicities, integration values, peaks due to solvent or impurities, etc.) were not elicited. Two of five external experts stated that such assumptions were held by their own students but were not captured by this study. Findings from this study may therefore only partially transfer to the interpretation of more authentic spectra.

Moreover, participants only completed three interpretation tasks due to fatigue reported by pilot study members. Although interpretation tasks included a variety of spectral features in order to elicit a range of reasoning from participants, it is possible that some invalid chemical assumptions or heuristic reasoning strategies were not captured given this small number of tasks. Further, Pearson's χ^2 test of independence and Fisher's exact test were used to establish that tasks did not disproportionately elicit certain assumptions or heuristics. However, combinations of assumptions and heuristics appearing in incorrect responses did tend to vary with context. Further investigations are therefore needed to characterize any additional combinations.

In addition to not eliciting all possible reasoning, some potentially problematic reasoning was not coded as an invalid chemical assumption or heuristic as it required significant inference by the authors. For instance, while some participants stated that OH and NH hydrogens should appear as singlets, it was unclear if these participants invalidly assumed that these hydrogens *always* appear as singlets regardless of solvent effects. Similarly, some participants may have used certain heuristics subconsciously,³⁰ however such use was not captured. One exception was made for the affect heuristic, which was coded regardless of whether participants' emotions directly influenced their ultimate decision. The prevalence of certain assumptions or heuristics may therefore have been underestimated for this study. It is also possible that participants failed to verbalize or remember their reasoning during RTA interviews, however this limitation was mitigated by the interviewer's use of probing questions and students' ability to pause the eye movement recording during the interview and reflect on their thinking.

Lastly, the study population was a convenience sample and therefore may not reflect reasoning used by all organic chemistry students, in particular those in other instructional contexts. The review of invalid chemical assumption themes by external experts in part contributes to the transferability of the findings to other instructional contexts. However, we have included a detailed description of the instructional setting and data in the form of participant quotes to allow for judgements regarding transfer.

2.10 Implications for teaching and research

2.10.1 Implications for teaching

Findings from this study provide additional evidence that problematic reasoning among students is not solely a product of any misconceptions they hold, but rather a combination of their underlying assumptions and heuristics. If instruction is to foster students' ability to interpret spectra, it must therefore explicitly address these assumptions as well as actively promote students' shift from Type 1 to Type 2 thinking. There are a number of promising strategies for shifting decision makers from Type 1 to Type 2 thinking;^{63,64} for instance, research in cognitive psychology suggests that prompting individuals to "consider the

opposite” of any decision they are about to make can promote Type 2 thinking and correct for decision biases resulting from the use of heuristics.⁶⁵ This research also suggests that having individuals assess the rightness of their decision promotes Type 2 thinking, where individuals with low feelings of rightness demonstrate increased rethinking times and increased probability of answer change.⁶⁶ Further, research in chemistry education demonstrates that having chemistry students predict how incorrect students may respond to given questions positively influences performance on these questions;⁶⁷ this finding further suggests that even simple interventions may help students spend more time evaluating cues and reflecting on their decisions. Instructors could incorporate any of these strategies into course materials on spectral interpretation (e.g., clicker questions, practice problems, exam questions, etc.) with minimal effort.

In addition to these easily-adopted strategies, research on cognitive biases in medicine offers targeted approaches for promoting Type 2 thinking.⁶⁸ One of these approaches involves providing practitioners with detailed descriptions of common, problematic heuristics along with several clinical examples that illustrate how their use negatively impacts decision-making.⁶⁸ Transferring this approach to instruction on spectral interpretation would involve instructors providing students with detailed descriptions of common, problematic heuristics along with example spectra which illustrate how the use of each heuristic results in erroneous decision-making. Findings from this study provide insight into the most problematic heuristics used by organic chemistry students and may thus inform such instruction.

2.10.2 Implications for research

To the best of the authors’ knowledge, this is the first study to use RTA interviewing in combination with eye tracking to collect qualitative data on students’ reasoning in chemistry. The abundance and complexity of assumptions and heuristics captured using this data collection method indicate that it serves as a valid and promising tool to investigate students’ reasoning for complex chemistry tasks, in particular those for which standard think-aloud techniques may overburden participants’ cognitive load.

In addition, findings from this study lay groundwork for the development of a learning progression on spectral interpretation. This study focused only on reasoning at the lower anchor, in particular the invalid assumptions and heuristics used at this level. Additional studies are therefore needed to characterize the valid assumptions that guide students' reasoning at the lower anchor as well as how knowledge develops beyond this level. Further, findings demonstrate that problematic reasoning is not only a product of misconceptions but also of heuristic reasoning strategies. Research to characterize students' reasoning in chemistry should therefore extend beyond generating inventories of misconceptions.

2.11 Supporting Information

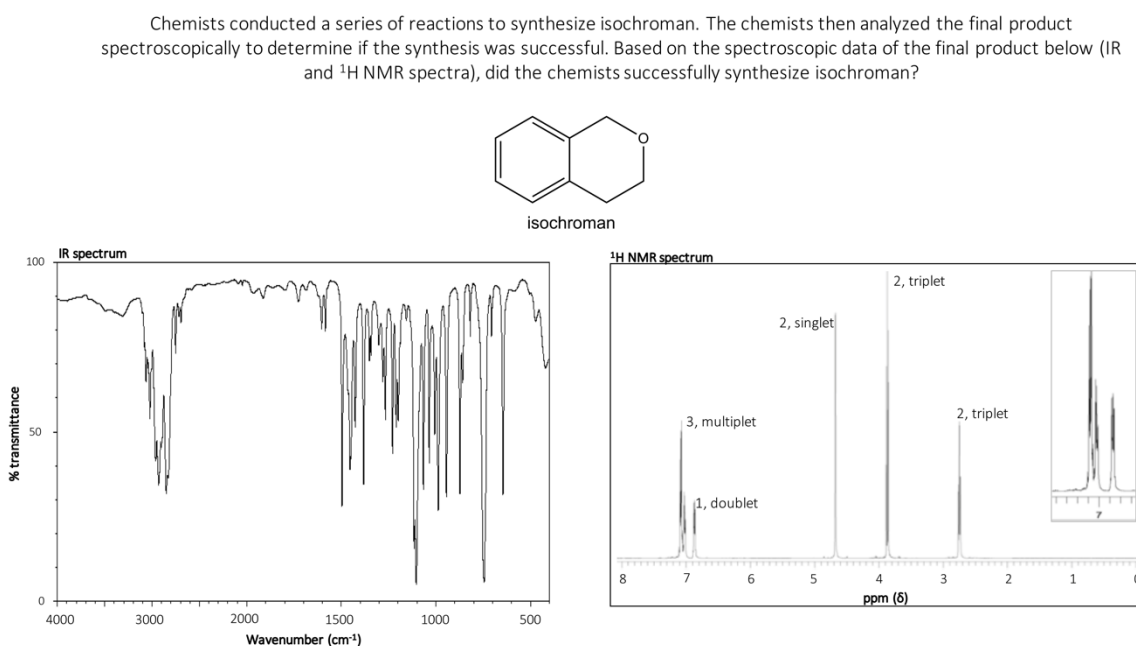


Figure 2.4. Spectral interpretation task (Synthesis 2) asking participants to determine if isochroman was successfully synthesized using the provided IR spectrum and ^1H NMR spectrum.

Chemists conducted a series of reactions to synthesize 3-(allyloxy)propanal. The chemists then analyzed the final product spectroscopically to determine if the synthesis was successful. Based on the spectroscopic data of the final product below (IR and ^1H NMR spectra), did the chemists successfully synthesize 3-(allyloxy)propanal?

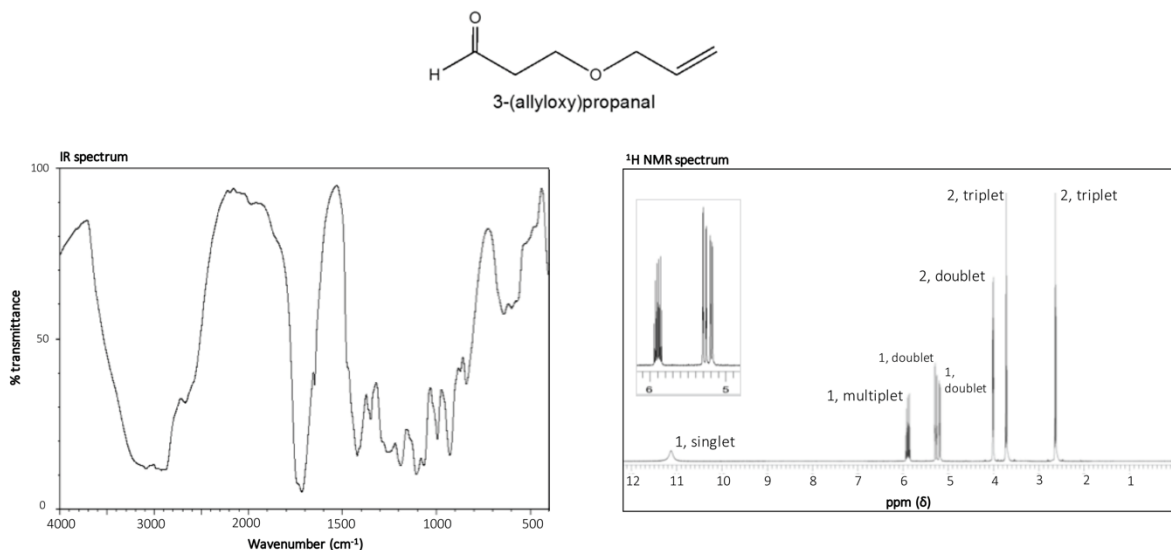


Figure 2.5. Spectral interpretation task (Synthesis 3) asking participants to determine if 3-(allyloxy)propanal was successfully synthesized using the provided IR spectrum and ^1H NMR spectrum.

Table 2.4. Invalid chemical assumption themes, the number of participants who used corresponding assumptions at least once ($n=18$), and frequencies of responses containing corresponding assumptions. The distributions of frequencies vary significantly with interpretation task ($p = 0.009$, two-tailed Fisher's exact test). However, when visuospatial invalid assumptions are omitted, distributions of frequencies do not vary significantly with task ($p = 0.323$, two-tailed Fisher's exact test); this lack of significance suggests that tasks may only disproportionately elicit visuospatial invalid assumptions.

Invalid chemical assumptions	Participants	Synthesis	Synthesis	Synthesis	Total Responses
		1	2	3	
Assumptions that the "N+1 rule" should hold	13	8	6	5	19
Assumptions that spectral data should be absolute	9	1	5	4	10
Practical invalid assumptions	8	2	1	6	9
Visuospatial invalid assumptions	8	0	8	0	8
Fundamental invalid assumptions	6	2	3	3	8

Table 2.5. Heuristics identified in at least 20% of RTA interview transcripts, the number of participants who used corresponding heuristics at least once (n=18), and frequencies of responses containing corresponding heuristics. The distribution of heuristic frequencies did not vary significantly with interpretation task ($\chi^2 = 9.03$, $p = 0.83$).

Heuristic	Participants	Synthesis 1	Synthesis 2	Synthesis 3	Total Responses
Processing fluency	18	18	17	17	52
Associative activation	17	16	11	15	42
Generalization	16	11	8	10	29
Representativeness	15	9	2	7	18
Affect	14	8	6	7	21
Reduction	14	4	6	10	20
Rigidity	11	6	6	4	16
One-reasoning decision making	10	6	3	2	11

2.12 References

- (1) Kozma, R. The Material Features of Multiple Representations and Their Cognitive and Social Affordances for Science Understanding. *Learn. Instr.* **2003**, *13* (2), 205–226. [https://doi.org/10.1016/s0959-4752\(02\)00021-x](https://doi.org/10.1016/s0959-4752(02)00021-x).
- (2) Kozma, R.; Russell, J. Students Becoming Chemists: Developing Representational Competence. In *Visualization in Science Education*; Gilbert, J. K., Ed.; Springer: Dordrecht, 2005; pp 121–146. https://doi.org/10.1007/1-4020-3613-2_4.
- (3) Alexander, C. W.; Asleson, G. L.; Doig, M. T.; Heldrich, F. J. Spectroscopic Instruction in Introductory Organic Chemistry: Results of a National Survey. *J. Chem. Educ.* **1999**, *76* (9), 1294–1296. <https://doi.org/10.1021/ed076p1294>.
- (4) Parmentier, L. E.; Lisensky, G. C.; Spencer, B. A Guided Inquiry Approach to NMR Spectroscopy. *J. Chem. Educ.* **1998**, *75* (4), 470–471. <https://doi.org/10.1021/ed075p470>.
- (5) Debska, B.; Guzowska-Swider, B. Molecular Structure from ^1H NMR Spectra: Education Aided by Internet Programs. *J. Chem. Educ.* **2007**, *84* (3), 556–560. <https://doi.org/10.1021/ed084p556>.
- (6) Veeraraghavan, S. NMR Spectroscopy and Its Value: A Primer. *J. Chem. Educ.*

- 2008**, 85 (4), 537–540. <https://doi.org/10.1021/ed085p537>.
- (7) Livengood, K.; Lewallen, D. W.; Leatherman, J.; Maxwell, J. L. The Use and Evaluation of Scaffolding, Student Centered-Learning, Behaviorism, and Constructivism to Teach Nuclear Magnetic Resonance and IR Spectroscopy in a Two-Semester Organic Chemistry Course. *J. Chem. Educ.* **2012**, 89 (8), 1001–1006. <https://doi.org/10.1021/ed200638g>.
 - (8) Azman, A. M.; Esteb, J. J. A Coin-Flipping Analogy and Web App for Teaching Spin-Spin Splitting in ¹H NMR Spectroscopy. *J. Chem. Educ.* **2016**, 93 (8), 1478–1482. <https://doi.org/10.1021/acs.jchemed.6b00133>.
 - (9) Erhart, S. E.; McCarrick, R. M.; Lorigan, G. A.; Yeziarski, E. J. Citrus Quality Control: An NMR/MRI Problem-Based Experiment. *J. Chem. Educ.* **2016**, 93 (2), 335–339. <https://doi.org/10.1021/acs.jchemed.5b00251>.
 - (10) Graham, K. J.; Mcintee, E. J.; Schaller, C. P. Web-Based 2D NMR Spectroscopy Practice Problems. *J. Chem. Educ.* **2016**, 93 (8), 1483–1485. <https://doi.org/10.1021/acs.jchemed.6b00007>.
 - (11) Vosegaard, T. ISpec: A Web-Based Activity for Spectroscopy Teaching. *J. Chem. Educ.* **2018**, 95 (1), 97–103. <https://doi.org/10.1021/acs.jchemed.7b00482>.
 - (12) Cartrette, D. P.; Bodner, G. M. Non-Mathematical Problem Solving in Organic Chemistry. *J. Res. Sci. Teach.* **2010**, 47 (6), 643–660. <https://doi.org/10.1002/tea.20306>.
 - (13) Cullipher, S.; Seviran, H. Atoms versus Bonds: How Students Look at Spectra. *J. Chem. Educ.* **2015**, 92 (12), 1996–2005. <https://doi.org/10.1021/acs.jchemed.5b00529>.
 - (14) Topczewski, J. J.; Topczewski, A. M.; Tang, H.; Kendhammer, L. K.; Pienta, N. J. NMR Spectra through the Eyes of a Student: Eye Tracking Applied to NMR Items. *J. Chem. Educ.* **2017**, 94 (1), 29–37. <https://doi.org/10.1021/acs.jchemed.6b00528>.
 - (15) Connor, M. C.; Shultz, G. V. Teaching Assistants' Topic-Specific Pedagogical Content Knowledge in ¹H NMR Spectroscopy. *Chem. Educ. Res. Pract.* **2018**, 19 (3), 653–669. <https://doi.org/10.1039/c7rp00204a>.
 - (16) Margolis, E. A Reassessment of the Shift from the Classical Theory of Concepts to Prototype Theory. *Cognition* **1994**, 51 (1), 73–89. [https://doi.org/10.1016/0010-0277\(94\)90009-4](https://doi.org/10.1016/0010-0277(94)90009-4).
 - (17) Gelman, S. A. Learning from Others: Children's Construction of Concepts. *Annu. Rev. Psychol.* **2009**, 60, 115–140. <https://doi.org/10.1146/annurev.psych.59.103006.093659>.

- (18) Murphy, G. L.; Medin, D. L. The Role of Theories in Conceptual Change. *Psychol. Rev.* **1985**, *92* (3), 289–316. <https://doi.org/10.1016/j.jff.2018.03.021>.
- (19) 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>.
- (20) Vosniadou, S. Capturing and Modeling the Process of Conceptual Change. *Learn. Instr.* **1994**, *4* (1), 45–69. [https://doi.org/10.1016/0959-4752\(94\)90018-3](https://doi.org/10.1016/0959-4752(94)90018-3).
- (21) Chi, Michelene, T. H.; Feltovich, Paul, J.; Glaser, R. Categorization and Representation of Physics Problems by Experts and Novices. *Cogn. Sci.* **1979**, *5*, 121–152. https://doi.org/10.1207/s15516709cog0502_2.
- (22) 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.* **2018**. <https://doi.org/10.1039/C8RP00120K>.
- (23) Stains, M.; Talanquer, V. Classification of Chemical Reactions: Stages of Expertise. *J. Res. Sci. Teach.* **2008**, *45* (7), 771–793. <https://doi.org/10.1002/tea.20221>.
- (24) Talanquer, V. Commonsense Chemistry: A Model for Understanding Students' Alternative Conceptions. *J. Chem. Educ.* **2006**, *83* (5), 811. <https://doi.org/10.1021/ed083p811>.
- (25) Talanquer, V. On Cognitive Constraints and Learning Progressions: The Case of "Structure of Matter." *Int. J. Sci. Educ.* **2009**, *31* (15), 2123–2136. <https://doi.org/10.1080/09500690802578025>.
- (26) diSessa, A. A. Toward an Epistemology of Physics. *Cogn. Instr.* **1993**, *10* (2–3), 105–225. <https://doi.org/10.1080/07370008.1985.9649008>.
- (27) Shah, A. K.; Oppenheimer, D. M. Heuristics Made Easy: An Effort-Reduction Framework. *Psychol. Bull.* **2008**, *134* (2), 207–222. <https://doi.org/10.1037/0033-2909.134.2.207>.
- (28) Osman, M.; Stavy, R. Development of Intuitive Rules: Evaluating the Application of the Dual-System Framework to Understanding Children's Intuitive Reasoning. *Psychon. Bull. Rev.* **2006**, *13* (6), 935–953. <https://doi.org/10.3758/BF03213907>.
- (29) Evans, J. S. B. T.; Stanovich, K. E. Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspect. Psychol. Sci.* **2013**, *8* (3), 223–241. <https://doi.org/10.1177/1745691612460685>.
- (30) McClary, L.; Talanquer, V. Heuristic Reasoning in Chemistry: Making Decisions about Acid Strength. *Int. J. Sci. Educ.* **2011**, *33* (10), 1433–1454.

<https://doi.org/10.1080/09500693.2010.528463>.

- (31) Taber, K. S. College Students' conceptions of Chemical Stability: The Widespread Adoption of a Heuristic Rule out of Context and beyond Its Range of Application. *Int. J. Sci. Educ.* **2009**, *31* (10), 1333–1358. <https://doi.org/10.1080/09500690801975594>.
- (32) Cooper, M. M.; Corley, L. M.; Underwood, S. M. 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>.
- (33) Graulich, N. Intuitive Judgments Govern Students' Answering Patterns in Multiple-Choice Exercises in Organic Chemistry. **2015**. <https://doi.org/10.1021/ed500641n>.
- (34) Smith, C. L.; Wiser, M.; Anderson, C. W.; Krajcik, J. FOCUS ARTICLE: Implications of Research on Children's Learning for Standards and Assessment: A Proposed Learning Progression for Matter and the Atomic-Molecular Theory. *Meas. Interdiscip. Res. Perspect.* **2006**, *4* (1–2), 1–98. <https://doi.org/10.1080/15366367.2006.9678570>.
- (35) Auchincloss, L. C.; Laursen, S. L.; Branchaw, J. L.; Eagan, K.; Graham, M.; Hanauer, D. I.; Lawrie, G.; Mclinn, C. M.; Pelaez, N.; Rowland, S.; et al. Assessment of Course-Based Undergraduate Research Experiences: A Meeting Report. *CBE—Life Sci. Educ.* **2014**, *13*, 29–40. <https://doi.org/10.1187/cbe.14-01-0004>.
- (36) Cohen, L.; Manion, L.; Morrison, K. *Research Methods in Education*, 7th ed.; Routledge: London, 2011.
- (37) Hyrskykari, A.; Ovaska, S.; Majaranta, P.; Rähkä, K.-J.; Lehtinen, M. Gaze Path Stimulation in Retrospective Think-Aloud. *J. Eye Mov. Res.* **2008**, *2* (4), 1–18.
- (38) Just, M. A.; Carpenter, P. A. A Theory of Reading: From Eye Fixations to Comprehension. *Psychol. Rev.* **1980**, *87* (4), 329–354. <https://doi.org/10.1037/0033-295X.87.4.329>.
- (39) van Gog, T.; Paas, F.; van Merriënboer, J. J. G. G.; Witte, P. Uncovering the Problem-Solving Process: Cued Retrospective Reporting versus Concurrent and Retrospective Reporting. *J. Exp. Psychol. Appl.* **2005**, *11* (4), 237–244. <https://doi.org/10.1037/1076-898X.11.4.237>.
- (40) Guan, Z.; Lee, S.; Cuddihy, E.; Ramey, J. The Validity of the Stimulated Retrospective Think-Aloud Method as Measured by Eye Tracking. *Proc. SIGCHI Conf. Hum. Factors Comput. Syst. - CHI '06* **2006**, 1253. <https://doi.org/10.1145/1124772.1124961>.
- (41) Tobii Technology. Eye Tracking for Research <https://www.tobiipro.com/> (accessed

Feb 26, 2020).

- (42) 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. *Chem. Educ. Res. Pract.* **2012**, *13* (3), 179–185. <https://doi.org/10.1039/C2RP90001G>.
- (43) National Institute of Advanced Industrial Science and Technology. SDBSWeb <http://sdb.sdb.aist.go.jp> (accessed May 4, 2020).
- (44) Bruice, P. Y. *Organic Chemistry*, 6th ed.; Prentice Hall: Upper Saddle River, 2011.
- (45) Bowen, C. W.; Bodner, G. M. Problem Solving Processes Used by Students in Organic Synthesis. *Int. J. Sci. Educ.* **1991**, *13*, 143–158.
- (46) Miles, M. B.; Huberman, A. M. *Qualitative Data Analysis*, 2nd ed.; Sage Publications: Thousand Oaks, 1994.
- (47) Saldana, J. *The Coding Manual for Qualitative Researchers*; Sage Publications: Los Angeles, 2016.
- (48) Creswell, J. W.; Poth, C. N. *Qualitative Inquiry and Research Design*, 4th ed.; Sage Publications: Los Angeles, 2018.
- (49) Talanquer, V. Chemistry Education: Ten Heuristics to Tame. *J. Chem. Educ.* **2014**, *91* (8), 1091–1097. <https://doi.org/10.1021/ed4008765>.
- (50) Sheskin, D. J. *Handbook of Parametric and Nonparametric Statistical Procedures*, 5th ed.; CRC Press: Boca Raton, 2011.
- (51) R Core Team. R: A Language and Environment for Statistical Computing. Vienna 2019.
- (52) Lawson, R. Small Sample Confidence Intervals for the Odds Ratio. *Commun. Stat. Part B Simul. Comput.* **2004**, *33* (4), 1095–1113. <https://doi.org/10.1081/SAC-200040691>.
- (53) Chen, H.; Cohen, P.; Chen, S. How Big Is a Big Odds Ratio? Interpreting the Magnitudes of Odds Ratios in Epidemiological Studies. *Commun. Stat. Simul. Comput.* **2010**, *39* (4), 860–864. <https://doi.org/10.1080/03610911003650383>.
- (54) Wu, H. K.; Shah, P. Exploring Visuospatial Thinking in Chemistry Learning. *Sci. Educ.* **2004**, *88* (3), 465–492. <https://doi.org/10.1002/sce.10126>.
- (55) Harle, M.; Towns, M. A Review of Spatial Ability Literature, Its Connection to Chemistry, and Implications for Instruction. *J. Chem. Educ.* **2011**, *88* (3), 351–360.

<https://doi.org/10.1021/ed900003n>.

- (56) Pavia, D. L.; Lampman, G. M.; Kriz, G. S.; Vyvyan, J. R. *Introduction to Spectroscopy*, 5th ed.; Cengage Learning: Stamford, 2015.
- (57) Morewedge, C. K.; Kahneman, D. Associative Processes in Intuitive Judgment. *Trends Cogn. Sci.* **2010**, *14* (10), 435–440. <https://doi.org/10.1016/j.tics.2010.07.004>.
- (58) Heckler, A. F. The Ubiquitous Patterns of Incorrect Answers to Science Questions: The Role of Automatic, Bottom-Up Processes. In *Psychology of Learning and Motivation*; Mestre, J. P., Ross, B. H., Eds.; Academic Press: London, 2011; Vol. 55, pp 227–267. <https://doi.org/10.1016/B978-0-12-387691-1.00008-9>.
- (59) Tversky, A.; Kahneman, D. Judgment under Uncertainty: Heuristics and Biases. *Science* (80-.). **1974**, *185* (4157), 1124–1131. <https://doi.org/10.4324/9781912282562>.
- (60) Kahneman, D.; Klein, G. Conditions for Intuitive Expertise: A Failure to Disagree. *Am. Psychol.* **2009**, *64* (6), 515–526. <https://doi.org/10.1037/a0016755>.
- (61) Gigerenzer, G.; Gaissmaier, W. Heuristic Decision Making. *Annu. Rev. Psychol.* **2011**, No. 62, 451–482. <https://doi.org/10.1146/annurev-psych-120709-145346>.
- (62) Ferguson, C. J. An Effect Size Primer: A Guide for Clinicians and Researchers. *Prof. Psychol. Res. Pract.* **2009**, *40* (5), 532–538. <https://doi.org/10.1037/a0015808>.
- (63) Lerner, J.; Tetlock, P. Accounting for the Effects of Accountability. *Psychol. Bull.* **1999**, *125*, 255–275.
- (64) Milkman, K. L.; Chugh, D.; Bazerman, M. H. How Can Decision Making Be Improved? *Perspect. Psychol. Sci.* **2009**, *4* (4), 379–383.
- (65) Mussweiler, T.; Strack, F.; Pfeiffer, T. Overcoming the Inevitable Anchoring Effect: Considering the Opposite Compensates for Selective Accessibility. *Personal. Soc. Psychol. Bull.* **2000**, *26* (9), 1142–1150.
- (66) Thompson, V. A.; Turner, J. A. P.; Pennycook, G. Intuition, Reason, and Metacognition. *Cogn. Psychol.* **2011**, *63* (3), 107–140. <https://doi.org/10.1016/j.cogpsych.2011.06.001>.
- (67) Talanquer, V. Concept Inventories: Predicting the Wrong Answer May Boost Performance. *J. Chem. Educ.* **2017**, *94*, 1805–1810. <https://doi.org/10.1021/acs.jchemed.7b00427>.

- (68) Croskerry, P. The Importance of Cognitive Errors in Diagnosis and Strategies to Minimize Them. *Acad. Med.* **2003**, 78 (8), 775–780.

Chapter 3

Developing Expertise in ^1H NMR Spectral Interpretation

3.1 Initial remarks

This chapter corresponds to the second of two studies investigating the development of expertise in ^1H NMR spectral interpretation. While the first study sought to identify cognitive barriers associated with initially learning this practice, the second study sought to characterize how conceptual understanding and information processing evolve with increasing expertise once these barriers are surpassed. The second study thus provided a detailed map of developing expertise in ^1H NMR spectral interpretation, while the first study provided insight into potential roadblocks near the beginning of this map. Findings from both studies are important for designing undergraduate and graduate-level instruction that fosters relevant expertise.

Specifically, the second study investigated undergraduate and doctoral chemistry students' cognitive processes and underlying chemical assumptions during the interpretation of ^1H NMR spectra and complementary IR spectra. This second study builds upon the first study in a number of ways. Data collected from undergraduate participants during the first study included eye movements and retrospective think-aloud (RTA) interviews, though data analysis in this first study was restricted to the inaccurate chemical assumptions expressed in RTA interviews. For this second investigation, data analysis focused on these participants' eye movements and all chemical assumptions expressed in RTA interviews. Doctoral chemistry students were also included as participants in this second investigation. Doctoral participants completed a data collection procedure identical to that of undergraduate participants, in which they completed the same spectral interpretation tasks while having their eye movements tracked and then verbalized their thinking during RTA interviews. The data corpus for the second study

thus included eye movements and RTA interviews from undergraduate and doctoral chemistry students.

Results from the second study suggest that five areas of conceptual understanding are necessary for interpreting ^1H NMR spectra, and that progress in each area corresponds to increasing knowledge of experimental and implicit chemical variables. Notably, the five identified areas parallel the five categories of inaccurate chemical assumptions that in part guided undergraduates' thinking in the first study; potential barriers to learning thus likely exist across all five of these areas. Results further suggest that limited understanding in these areas results in the uninformed processing of all features within a spectrum, whereas more sophisticated understanding allows for the efficient and selective processing of relevant spectral data. These findings provide insight into how the organic chemistry community may cultivate expertise in ^1H NMR spectral interpretation among its newest members.

This chapter first appeared as a research article in the *Journal of Organic Chemistry*, and original publication and copyright information are provided below. All changes to the original publication were cosmetic and involved adjusting formatting to adhere to Rackham dissertation requirements. Benjamin Glass and Dr. Solaire Finkenstaedt-Quinn assisted with the design of codebooks and qualitative coding of interview data, as well as provided feedback on the manuscript prior to journal submission. Benjamin Glass also assisted with visualizing quantitative results. Further, Dr. Solaire Finkenstaedt-Quinn assisted with data collection involving doctoral participants. All remaining work, including study design, thematic analyses, quantitative data analysis, and writing of the manuscript were completed independently by the author.

Original publication and copyright information

Reproduced with permission from “Connor, M. C.; Glass, B. H.; Finkenstaedt-Quinn, S. A.; Shultz, G. V. Developing Expertise in ^1H NMR Spectral Interpretation. *J. Org. Chem.* **2020**. <https://doi.org/10.1021/acs.joc.0c01398>.” Copyright 2020. American Chemical Society.

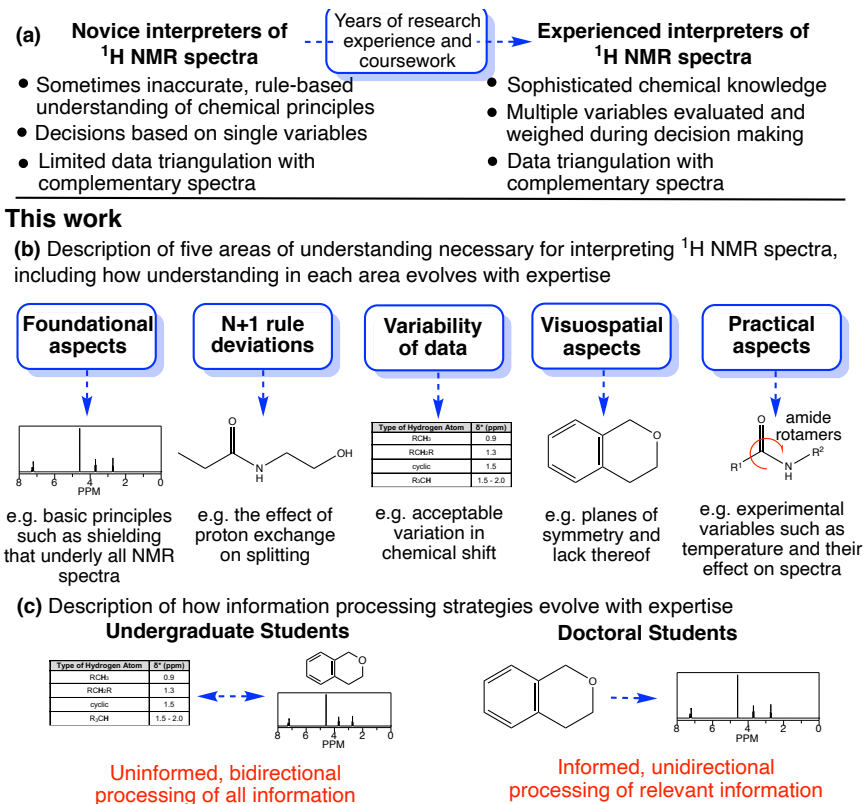
3.2 Abstract

Advancements in organic chemistry depend upon chemists' ability to interpret NMR spectra, though research demonstrates that cultivating such proficiency requires years of graduate-level study. The organic chemistry community thus needs insight into how this expertise develops to expedite learning among its newest members. This study investigated undergraduate and doctoral chemistry students' understanding and information processing during the interpretation of ^1H NMR spectra and complementary IR spectra. Eighteen undergraduate and seven doctoral chemistry students evaluated the outcome of a series of syntheses using spectra corresponding to the products. Eye movements were measured to identify differences in cognitive processes between undergraduate and doctoral participants, and interviews were conducted to elucidate the chemical assumptions that guided participants' reasoning. Results suggest five areas of understanding are necessary for interpreting spectra, and progress in understanding corresponds to increasing knowledge of experimental and implicit chemical variables. Undergraduate participants exhibited uninformed bidirectional processing of all information, whereas doctoral participants exhibited informed unidirectional processing of relevant information. These findings imply the community can support novices' development of expertise by cultivating relevant understanding and encouraging use of informed interpretation strategies, including preliminary evaluation of relevant variables, prediction of expected spectral features, and search for complementary data across spectra.

3.3 Introduction

Nuclear magnetic resonance (NMR) spectroscopy is an essential analytical tool in organic chemistry, where the technique's primary value lies in its potential for determining chemical identity through characterizing molecular structure.¹ NMR serves as a principle means of investigating aspects such as molecular conformation and enantiopurity, and the continuous development of new methods expands upon this potential to encompass novel applications such as quantifying metal-ligand complexes' relative Lewis acidity and real-time enzymatic reaction monitoring.¹⁻⁴ The technique thus exhibits remarkable

Scheme 3.1. Empirical insights into developing expertise in ^1H NMR spectral interpretation



disciplinary value, though unfortunately, research demonstrates that individuals often struggle to successfully evaluate NMR spectra in the absence of significant research and classroom experience.⁵ The *Journal of Organic Chemistry* has published multiple educational resources for interpreting NMR spectra to support its readership's development of this essential ability.⁶⁻⁸ In addition to these resources, the organic chemistry community requires an understanding of how expertise in this practice develops. Such insight will help advisors, mentors, and organic chemistry instructors provide guidance and instruction that expedites this development among their students and mentees.

3.3.1 Mapping the development of expertise in ^1H NMR spectral interpretation

Several studies within the past decade investigated the development of expertise in NMR spectral interpretation, with most focusing on how undergraduates, doctoral students, or faculty members evaluate ^1H NMR spectra (Scheme 3.1, a).^{5,9-11} Findings from these studies suggest that individuals eventually develop such expertise, though they typically

only acquire proficiency after years of graduate-level study. These findings also suggest that this development corresponds to an increasingly sophisticated (1) conceptual understanding of relevant chemical principles and (2) reasoning ability that involves evaluating multiple implicit variables. Our study provides a detailed map of this development among undergraduate and doctoral chemistry students (Scheme 3.1, b-c). Learning to interpret ^1H NMR spectra takes place both in the research laboratory and the classroom, making these findings relevant to the entire organic chemistry community. Advisors and mentors can use this map to guide incoming graduate students and mentees toward the expert-like interpretation of spectra necessary for advancements in the field. Graduate-level instructors can then use this map to design instructional materials that provide additional supplementary preparation for first and second-year students. Lastly, undergraduate-level instructors can also use this map to design instructional materials, helping to ensure that incoming graduate students and individuals entering industry directly from their undergraduate institutions can engage in this essential practice.

3.3.2 Mapping the evolution of conceptual understanding

This study aimed to provide a detailed map of how expertise in ^1H NMR spectral interpretation develops among undergraduate and doctoral chemistry students, with a focus on how requisite conceptual understanding may evolve with increasing expertise. Research suggests that underlying concepts are often verbalisms for undergraduate and first-year doctoral students in organic chemistry.¹² Further, more advanced doctoral students are capable of rationalizing spectroscopic data, though they often do so without understanding the physical principles underlying phenomena.¹² Conversely, practicing organic chemists often exhibit deep conceptual understanding that allows them to think about molecules and chemical processes at the molecular level, which then allows them to make chemical predictions necessary for advancements in the field.¹² Investigating how relevant conceptual understanding develops and subsequently cultivating this understanding among the field's newest members is thus essential for supporting their development of expertise in interpreting ^1H NMR spectra.

To investigate the nature and evolution of conceptual understanding, this study draws upon the cognitive theory of categorization. Research in cognitive psychology suggests that as individuals interact with any entity, they construct a mental representation of the entity that corresponds to a given category.^{13,14} This categorization is informed by one's underlying knowledge about the entity at hand.^{13,14} Assumptions one has about the nature and behavior of entities belonging to the category then guide subsequent reasoning.¹⁵ For example, a novice interpreter of ^1H NMR spectra may observe a hydroxyl hydrogen atom (i.e., an entity) within a structural formula. Due to their limited knowledge, they may then consider it as simply a hydrogen atom (i.e., their categorization), for which they assume splitting is determined using the N+1 rule (i.e., their assumption about all entities belonging to this category). Future reasoning would then involve their failure to identify the corresponding unsplit peak on a spectrum.

Conversely, an experienced interpreter of ^1H NMR spectra may observe the hydroxyl hydrogen atom (i.e., an entity) and use their developed knowledge to categorize it as molecular feature which undergoes proton exchange (i.e., their categorization). Their assumption that this implicit molecular feature affects splitting would then facilitate their identification of the corresponding peak (i.e., their assumption about all entities belong to this category). As this example illustrates, identifying the assumptions that guide individuals' reasoning serves as a means of characterizing progress in understanding.^{16,17} By using a method such as think-aloud interviewing to identify all assumptions about molecular and spectral features that guide individuals' thinking during spectral interpretation, a model can be developed that describes how conceptual understanding may evolve with increasing expertise in this practice.

3.3.3 Mapping the evolution of information processing

To provide a complete map of development, this study also aimed to describe how information processing during spectral interpretation changes with increasing expertise. NMR spectra and structural formulae are the most common visualizations used by synthetic chemists, making spectral interpretation an inherently visual task.¹⁸ Individuals with expertise in spectral interpretation thus also possess expertise in comprehending

these related visualizations. Multiple cognitive theories propose that experts process information contained within visualizations differently than non-experts.^{19–22} The information-reduction hypothesis suggests that experts use their developed knowledge structure to disregard irrelevant information within visualizations and selectively focus on task-relevant features, thus optimizing their processing of information.¹⁹ Further, the theory of long-term working memory posits that experts rapidly store information in their long-term memory and easily retrieve it as they complete tasks using their working memory, or their temporary, limited storage of information they are immediately processing.²² As a result of this rapid encoding and retrieval, they spend less time processing information contained within visualizations than novices. In the context of spectral interpretation, experts should thus be able to disregard uninformative molecular and spectral features and focus only on those relevant to their analysis. They should also be able to readily access knowledge about these features and efficiently use this understanding as they make judgments and decisions. As research demonstrates, novices would likely struggle to identify relevant features, access knowledge about these features, and even infer meaning from organic structural formulae.^{9,23} Cognitive processes such as attention and memory are the mental operations involved in information processing.²⁴ Investigating differences in cognitive processes between undergraduate and doctoral students using a method such as eye-tracking thus serves as an essential step toward modeling how expertise in this practice develops.

3.4 Expertise analysis

This study investigated the development of expertise in ¹H NMR spectral interpretation among undergraduate and doctoral chemistry students. Specifically, it characterized the assumptions guiding undergraduate and doctoral students' thinking during the interpretation of ¹H NMR spectra and complementary IR spectra, how these assumptions vary in sophistication, and how cognitive processes involved in the interpretation of these data differ between undergraduate and doctoral participants. Eighteen undergraduate and seven doctoral chemistry students from a research-intensive university determined the success of three syntheses using spectral data corresponding to the product of each synthesis while having their eye movements tracked. Participants were provided with a

structural formula corresponding to the target of synthesis, in addition to an ^1H NMR spectrum and complementary IR spectrum corresponding to the product of synthesis (Figure 3.1, Scheme 3.2). Tasks included a variety of spectral and molecular features, as well as ^1H NMR and IR reference material (Figures 3.4-3.7, Supporting Information). Complementary IR spectra were included to increase the authenticity of each task while also providing spectral data accessible to both undergraduate and doctoral populations. Participants then completed cued, retrospective think-aloud (RTA) interviews where they verbalized all thoughts involved in interpretation while watching a recording of their eye movements. Interviews were qualitatively coded for the assumptions guiding participants' thinking. Qualitative coding involved assigning short, summative labels (i.e., codes) to portions of transcribed verbal data - a systematic, analytical process that afforded a

Chemists conducted a series of reactions to synthesize N-(2-hydroxyethyl)-propanamide. The chemists then analyzed the final product spectroscopically to determine if the synthesis was successful. Based on the spectroscopic data of the final product below (IR and ^1H NMR spectra), did the chemists successfully synthesize N-(2-hydroxyethyl)-propanamide?

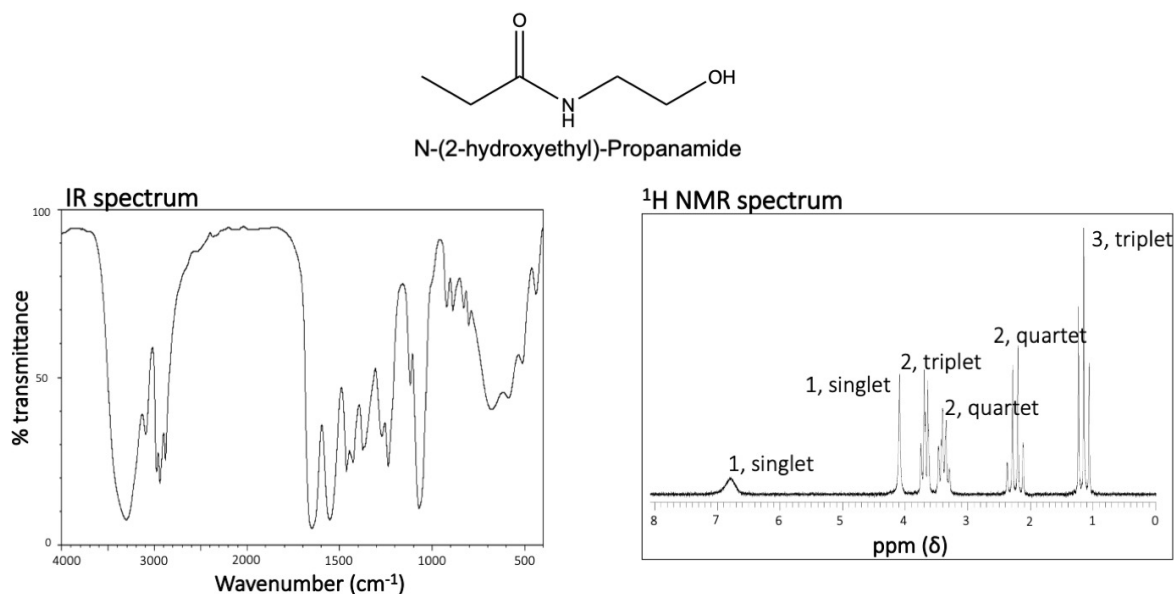
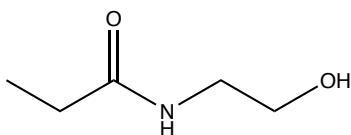


Figure 3.1. Spectral interpretation task (Synthesis 1) asking participants to determine if N-(2-hydroxyethyl)-propanamide was successfully synthesized using the provided IR spectrum and ^1H NMR spectrum. Text and images in this figure have been enlarged to improve readability; original tasks are provided in the Supporting Information. Spectra used in the study were obtained from the Spectral Database for Organic Compounds and are reproduced herein with permission from SDBSWeb.

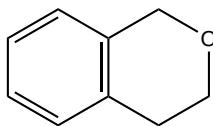
Scheme 3.2. Compounds identified as targets of synthesis in each interpretation task.

(a) Synthesis 1



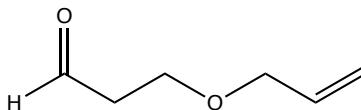
- N-(2-hydroxyethyl)-propanamide identified as target, with spectra corresponding to target
-

(b) Synthesis 2



- Isochroman identified as target, with spectra corresponding to target
-

(c) Synthesis 3



- 3-(allyloxy)propanal identified as target, with spectra corresponding to carboxylic acid derivative 3-(allyloxy)propionic acid

reliable and valid inference of participants' assumptions.²⁵ Themes among all coded assumptions were then identified, and assumptions within themes were categorized into levels of sophistication.²⁶ Identifying themes involved systematically grouping similar codes to identify patterns in participants' assumptions, thus providing a comprehensive summary of their thinking.²⁷ Eye-tracking data were quantitatively analyzed to identify differences in eye movements, a common measure of cognitive processes, between undergraduate and doctoral participants.²⁸

3.5 Results and discussion

3.5.1 Chemical assumptions

Qualitative analysis of the assumptions guiding participants' thinking revealed five areas of understanding necessary for interpreting spectra, in addition to clear progress in understanding across these areas (Figure 3.2). In general, this progress corresponded to increasingly sophisticated knowledge of experimental and implicit chemical variables. Doctoral participants relied almost exclusively on more sophisticated assumptions in each area, suggesting that the identified progression of assumptions validly reflects how

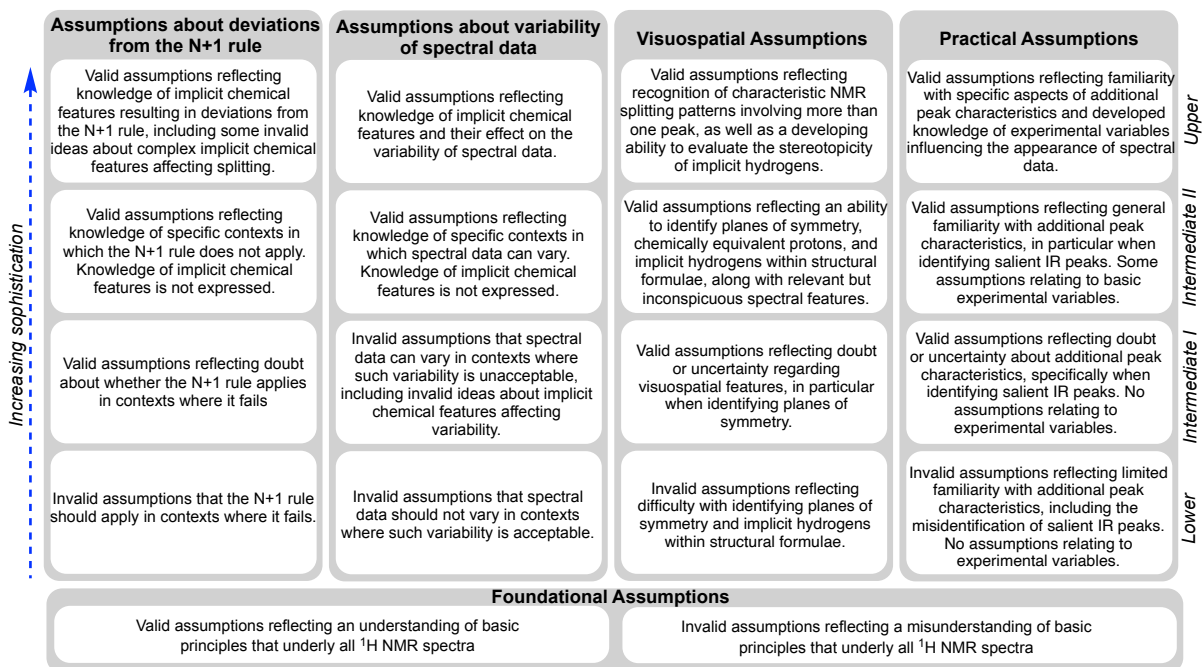


Figure 3.2. Five areas of understanding and description of increasing levels of sophistication.

understanding evolves with increasing expertise (Table 3.1). These findings indicate that the organic chemistry community can support its newest members' development of expertise in part by helping them cultivate sophisticated understanding across these five areas.

An understanding of basic principles that underly all ¹H NMR spectra comprised the foundation of this progression. Assumptions in this area reflected either an understanding or misunderstanding of fundamental principles such as shielding and basic application of the N+1 rule. Building on this foundation were areas corresponding to understanding of (1) deviations from the N+1 rule, (2) acceptable variation in spectral data such as absorption frequency, resolution, and signal intensity, (3) visuospatial aspects of structural formulae and spectra, and (4) practical considerations such as the influence of experimental variables on spectral appearance, as well as additional peak characteristics beyond N+1 rule deviations and acceptable variability in data.

Progress in the understanding of N+1 rule deviations reflected increasing knowledge of contexts in which this heuristic does not apply and other implicit chemical

Table 3.1. Number of undergraduate and doctoral participants with assumptions contributing to corresponding areas and levels.

	Foundational Assumptions		Assumptions about deviations from the N+1 rule		Assumptions about variability of spectral data		Visuospatial Assumptions		Practical Assumptions	
	Number of participants									
Sophistication Level	Ug	PhD	Ug	PhD	Ug	PhD	Ug	PhD	Ug	PhD
Upper	n/a	n/a	7	4	14	4	0	4	1	3
Intermediate II	n/a	n/a	16	7	15	2	7	4	18	7
Intermediate I	n/a	n/a	11	0	1	0	2	0	6	0
Lower	n/a	n/a	12	0	11	0	8	2	8	1
Understanding	18	7	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Misunderstanding	6	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

features affecting splitting. The least sophisticated assumptions in this area included rule-based, invalid ideas such as the splitting of signals corresponding to hydroxyl and vinylic hydrogen atoms should be accordance with the N+1 rule (Scheme 3.1, a and c). Assumptions of intermediate sophistication included ideas that such deviations were acceptable, albeit these ideas still appeared to be verbalisms as they focused on context and not underlying physical principles. The most sophisticated assumptions included ideas about implicit variables such as proton exchange and chemical inequivalence due to limited rotation around double bonds, including how these variables affect splitting. Notably, some of these ideas were invalid; this finding underscores that inaccurate, advanced ideas can still serve as a productive step toward developing expertise. Similarly, progress in the understanding of acceptable variation in spectral data reflected increasing knowledge of contexts in which variability is acceptable and other implicit chemical features resulting in variability. The least sophisticated assumptions included invalid ideas such as signals corresponding to aryl hydrogen atoms should be completed resolved (Scheme 3.2, b). Assumptions of intermediate sophistication included ideas that variability is acceptable, though these ideas either focused exclusively on context or reflected a misunderstanding of introductory-level physical principles (e.g., electronegativity). The most sophisticated assumptions related to implicit features affecting variability, such as the presence of nearby electronegative atoms and their effect

on shielding and, ultimately, chemical shift.

Progress in the understanding of visuospatial aspects corresponded to an increasing ability to recognize and reason about visuospatial features within provided structural formulae and spectra. Visuospatial information is both visual in nature and exhibits spatial properties involving the representation of space and relationships between entities within a given space.²⁹ The least sophisticated assumptions in this area included invalid ideas about the number of implicit hydrogen atoms represented in structural formulae, as well as the idea that isochroman contains a symmetry plane perpendicular to the plane of the molecule (Scheme 2, b). Assumptions of intermediate sophistication reflected a developing ability to identify visuospatial features such as symmetry and relevant yet poorly resolved signals. The most sophisticated assumptions went beyond this ability and involved recognizing characteristic splitting patterns like that corresponding to the ethyl group in (2-hydroxyethyl)-propanamide and reasoning about the stereotopicity of implicit hydrogen atoms (Scheme 2, a). Lastly, progress in the understanding of practical considerations reflected increasing familiarity with additional peak characteristics such as characteristic absorption frequency and intensity, as well as experimental variables and their effect on the appearance of spectra. The least sophisticated assumptions reflected a lack of familiarity with such variables. Assumptions of intermediate sophistication reflected a developing familiarity, like the idea that signals corresponding to solvent may appear in spectra. The most sophisticated assumptions reflected knowledge of more advanced experimental variables that influence the appearance of spectra, such as the effect of temperature on amide bond rotation and in turn the number of resulting signals. These assumptions also reflected existing knowledge of specific absorption frequencies not provided in the reference material. The five identified areas of understanding, their corresponding levels, and exemplar assumptions contributing to each area and level are described in additional detail in the Supporting Information (Tables 3.3-3.7).

3.5.2 Cognitive processes

Quantitative analysis of eye movements revealed several differences in cognitive processes between undergraduate and doctoral participants. These differences aligned

Table 3.2. Mann-Whitney U test comparisons of fixation counts between undergraduate (n=18) and doctoral (n=7) participants. *Corresponds to significance at the $p < 0.05$ level, ** at the < 0.01 level, and *** at the < 0.001 level. Medium effect sizes ranged from 0.3-0.5 and large effect sizes were > 0.5 .

	AOI	U Statistic	p-value	Effect size (r)	Ug Median	PhD Median
Synthesis 1	Fingerprint region	29.5	0.045*	0.407	5.5	3.0
	NH NMR Peak	27.5	0.034*	0.430	24.5	13.0
	Wavenumber Axis	27.0	0.031*	0.436	12.0	2.0
	IR Table	27.5	0.034*	0.430	52.5	2.0
Synthesis 2	Fingerprint Region	28.0	0.036*	0.426	8.0	3.0
	PPM Axis	27.0	0.031*	0.437	10.0	5.0
	Wavenumber Axis	29.0	0.042*	0.413	11.5	2.0
	IR Table	20.0	0.010*	0.521	71.0	1.0
	NMR Table	2.5	0.0003***	0.734	72.5	0.0
Synthesis 3	Fingerprint region	20.5	0.011*	0.516	13.0	6.0
	Molecule	25.0	0.021*	0.460	157.0	85.0
	Wavenumber Axis	26.0	0.027*	0.449	17.0	4.0
	IR Table	26.0	0.027*	0.448	88.0	13.0
	NMR Table	14.0	0.003**	0.594	86.5	9.0

closely with doctoral participants' nearly exclusive use of more sophisticated assumptions (Table 3.1). Collectively, these results suggest that undergraduates engaged in uninformed, bidirectional processing of all information. In contrast, doctoral participants relied on their more sophisticated understanding of experimental and implicit chemical variables to engage in informed, unidirectional processing of relevant information. These findings ultimately suggest that in addition to helping its newest members cultivate sophisticated understanding, the organic chemistry community can support the development of expertise by encouraging informed, unidirectional processing of relevant spectral features.

Visual attention. Fixation counts of undergraduate and doctoral participants were compared to identify differences in visual attention, a measure of perceived importance of information, between the two groups.³⁰ Mann-Whitney U test comparisons of undergraduate and doctoral participants' fixation counts revealed significant differences for several areas of interest (AOIs) (Table 3.2). All significant differences exhibited p-values less than 0.05 and medium to large effect sizes, indicating substantive differences in visual attention between the groups (Table 3.2).³¹ Overall, undergraduates fixated significantly more on reference tables, absorption frequency axes, and spectral features

that deviated from the N+1 rule when compared to doctoral participants. When combined with findings on participants' assumptions, these differences suggest that doctoral participants were able to use their more sophisticated understanding to selectively focus on relevant spectral features. In contrast, undergraduates' relatively limited understanding resulted in an uninformed focus on both relevant and irrelevant information.

Undergraduates' greater number of fixations on reference tables and absorption frequency axes aligns with their use of less sophisticated assumptions about the variability of spectral data (Tables 3.1-3.2). For these individuals holding assumptions about limited or context-specific variability, reference material and absorption frequencies likely served as important information and received more visual attention. Conversely, doctoral participants allocated less visual attention to this information while relying on more sophisticated assumptions about implicit features affecting variability (Tables 3.1-3.2). This more sophisticated understanding likely allowed doctoral participants to efficiently focus their visual attention on information they deemed relevant for decision making. Undergraduates' greater number of fixations on reference material also aligns with their use of less sophisticated practical assumptions, which reflect limited familiarity with peak characteristics such as absorption frequency and intensity. Their greater number of fixations on the fingerprint region of each IR spectrum further reflects these assumptions, which unlike more sophisticated practical assumptions fail to acknowledge that this region provides variably useful information. Conversely, doctoral participants fixated less on this information while relying on more sophisticated practical assumptions, suggesting that their existing familiarity with this information allowed them to focus their visual attention more selectively (Tables 3.1-3.2).

Further, undergraduates' greater number of fixations on the NH NMR peak in Synthesis 1 aligns with their use of less sophisticated assumptions about deviations from the N+1 rule (Tables 3.1-3.2). This NMR peak likely served as important information and in turn received greater visual attention due to undergraduates' relatively limited knowledge of implicit variables affecting splitting. In contrast, doctoral participants fixated less on this peak while relying on more sophisticated assumptions about implicit features

affecting splitting (Tables 3.1-3.2). Their more sophisticated understanding likely allowed them to process this information efficiently and then move on to additional task-relevant features.

Processing strategies. Undergraduate and doctoral participants' fixation transitions were compared to provide insight into differences in cognitive processing strategies between the two groups, in particular differences in the perceived complexity or importance of related information and the refreshment of working memory.³² Mann-Whitney U test comparisons of transition counts between undergraduate and doctoral participants revealed several significant differences. All significant differences exhibited p-values less than 0.05 and medium to large effect sizes, indicating substantive differences in processing strategies (Tables 3.8-3.11, Supporting Information).³¹ These differences are visualized in chord diagrams in Figure 3.3.

Across tasks, undergraduates made significantly more transitions involving reference tables and frequency axes, again suggesting that their limited understanding of acceptable variability and practical considerations inhibited more efficient processing demonstrated by doctoral participants. Nearly half of significantly different transitions (22 of 45) involved undergraduates moving more between reference tables and other AOIs, including the molecule, spectral peaks, and frequency axes. Further, several significantly different transitions (4 of 45) involved undergraduates moving more between frequencies axes and spectral peaks. For individuals with limited knowledge of characteristic absorption frequencies and the appropriate contexts in which spectral data can vary, reference material and frequency axes, along with corresponding spectral and molecular features, would expectedly serve as important related information. Further, clear bidirectional transitions involving reference tables and frequency axes suggest that undergraduates found this information not only more important but less familiar. Research demonstrates that bidirectional transitions, or back-and-forth movement between two AOIs, reflect a need to refresh working memory.³³ Notably, nearly half of the transitions involving reference tables (14 of 22) were bidirectional (Figure 3.3). For example, in Synthesis 3 undergraduates transitioned more both from the molecule to the IR table and from the IR table to the molecule (Figure 3.3). Half of transitions between spectral peaks

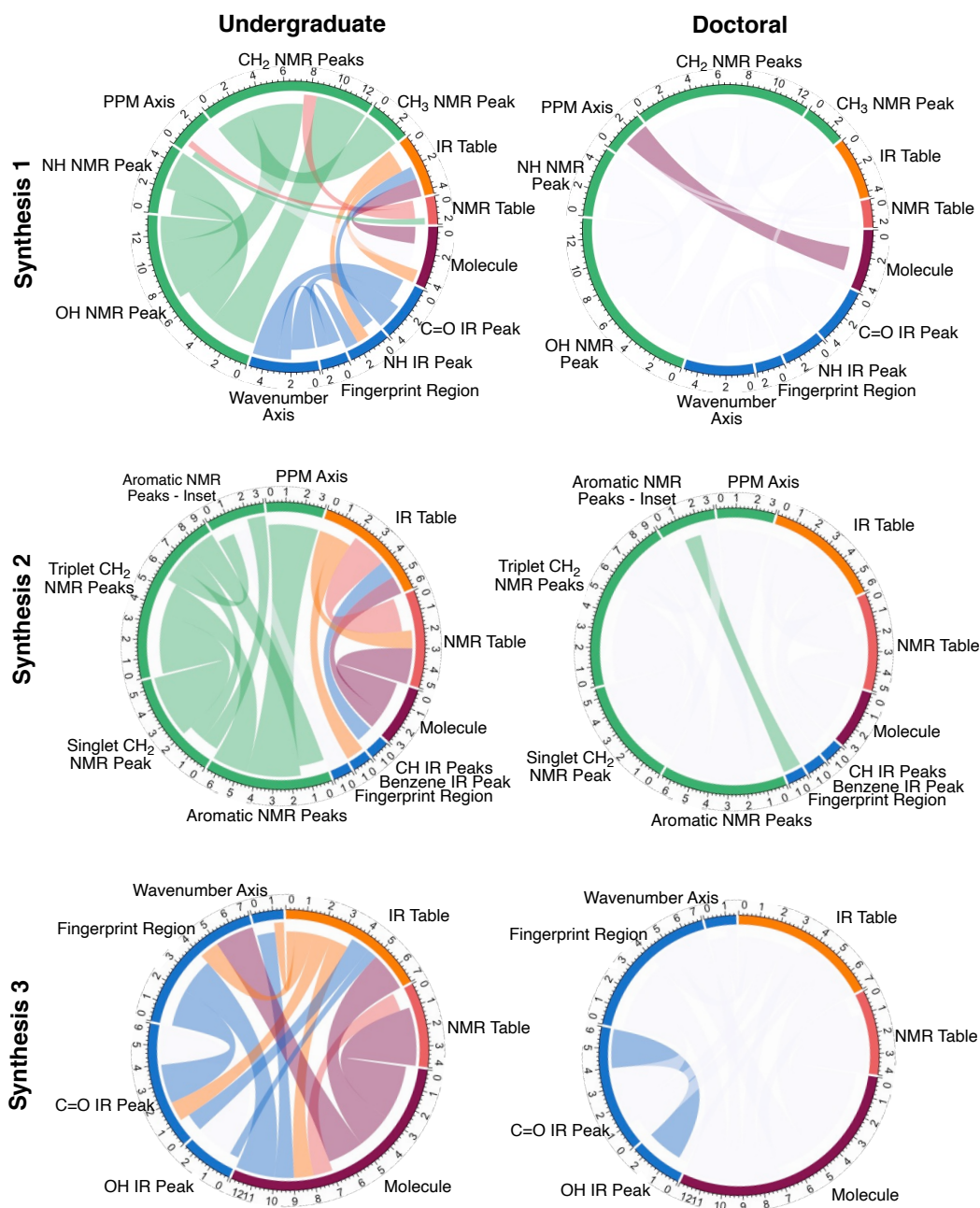


Figure 3.3. Chord diagrams of significantly different transition counts between undergraduate and doctoral participants for Syntheses 1-3. Transitions made more by undergraduate participants (N=18) are depicted on the left, and transitions made more by doctoral participants (n=7) are on the right. The color of each chord corresponds to the AOI group from which the transition originates (e.g., green corresponds to transitions beginning in NMR spectra), and the shorter end of each chord corresponds to the AOI in which the transition begins. The width of each chord corresponds to the difference in median number of transitions between undergraduate and doctoral participants.

and frequency axes (2 of 4) were also bidirectional. The absence of significant bidirectional transitions among doctoral participants suggests they have an existing familiarity with this information; this existing familiarity is thus likely necessary for efficient,

expert-like information processing.

Further, several significantly different transitions point toward a difference in the overall interpretation approach between the two groups. This difference in approach is reflected predominantly by the second most prevalent class of significantly different transitions, which involved movement between spectral peaks (15 out of 45). The majority of these transitions (13 out of 15) were made more by undergraduates and involved movement between peaks within NMR spectra in Syntheses 1 and 2 and between peaks within IR spectra in Syntheses 1 and 3 (Figure 3.3). The bidirectional nature of these transitions, which typically reflect the refreshment of working memory, combined with the proximity, similarity, and one-dimensional spatial relationship of the stimuli, suggest undergraduates were engaging in searching behavior. Doctoral participants, on the other hand, made more unidirectional transitions between NMR and IR peaks, specifically from the NMR peak inset to the fingerprint region in Synthesis 2 (Figure 3.3). They also made significantly more unidirectional transitions from the molecule to ppm axis in Synthesis 1 (1 of 45). These differences suggest that undergraduates searched within spectra to identify peaks corresponding to molecular features, whereas doctoral participants knew where a peak should appear on a spectrum and looked to confirm its presence in its characteristic chemical shift region. They also suggest that doctoral participants looked between spectra for complementary pieces of data, whereas undergraduates analyzed one spectrum at a time. This finding aligns with that of other another study demonstrating that experts in chemistry are able to coordinate information across representations more easily than novices.³⁴ The remaining significantly different transitions not described herein (3 of 45) involved transitions between the molecule and other spectral features. These transitions support the above findings and are described in detail in the Supporting Information.

3.6 Conclusions

Findings from this study provide multiple insights into how individuals develop expertise in interpreting NMR spectra, including how the organic chemistry community may help cultivate such proficiency among its newest members. Advisors and mentors, those who

interact with these members most often, can play an essential role in their development by encouraging informed unidirectional processing of spectra. Such processing involves predicting expected NMR peaks using structural formulae and subsequently confirming their appearance using corresponding NMR spectra, as well as predicting and confirming complementary features across spectra. Advisors and mentors may also encourage these members to preliminarily evaluate experimental and implicit chemical variables outlined in four of the five areas of understanding and predict how these variables will affect the appearance of spectra. Such guidance will serve to shift novices away from the uninformed bidirectional processing of all (i.e., relevant and irrelevant) information toward expert-like thinking. Graduate-level instructors, those who play a more formal role in educating incoming graduate students, could then design supplementary instruction around the five identified areas of understanding and incorporate activities that encourage informed processing and the evaluation of these variables. Undergraduate-level instructors could also adopt a similar instructional model to help further ensure that incoming graduate students and individuals entering industry directly from undergraduate institutions are prepared to engage in this practice. A detailed description of a potential instructional model is provided for graduate and undergraduate-level instructors in the Supporting Information.

3.7 Experimental section

3.7.1 Sample and setting

Study participants included 18 undergraduates and seven doctoral students enrolled at a large Midwestern university in Spring 2018. This sample size is comparable to that of other studies employing similar methods.^{11,35} Undergraduate participants were recruited from a second-semester organic chemistry laboratory course via email and in-class announcements, and doctoral participants were recruited via email and announcements in the staff meeting of this course. One undergraduate participant was recruited via snowball sampling, though they had completed the laboratory course in the semester prior.³⁶ All individuals consented to participate in the study, and approval from the university's institutional review board was obtained.

Of the 18 undergraduate participants, there were nine females and nine males. The second-semester organic chemistry course from which these participants were recruited covered ^1H NMR and IR spectroscopy in detail. The course consisted of a weekly one-hour laboratory lecture, during which instructors covered concepts relevant to ^1H NMR and IR spectral interpretation throughout the semester. Students were also required to utilize ^1H NMR and IR spectroscopy during their weekly 4-hour lab section, pre- and post-laboratory assignments, and 1-hour quizzes administered during the laboratory lecture. For additional practice with spectral interpretation, students were provided with a coursepack containing problems necessitating ^1H NMR and IR spectral interpretation. Connor *et al.* provides a detailed account of spectroscopy concepts covered in the course.⁹

Of the seven doctoral participants, there were three females and four males. All doctoral participants were in either the first or second year of their program. They were recruited from a range of subdisciplines, including inorganic chemistry, organic chemistry, and chemical biology. All doctoral participants had research experience with ^1H NMR or IR spectroscopy prior to the study, and the majority (5 out of 7) had taken a graduate-level spectroscopy course within the last semester. All doctoral participants had also served as graduate student instructors in the laboratory course from which undergraduates were recruited.

3.7.2 Data collection

This study employed a mixed-methods approach with concurrent design, meaning that both qualitative and quantitative data were collected concurrently to provide a rich, more complete characterization of the phenomenon under investigation.³⁶ Each study participant took part in one 30 to 60-minute session in which they completed three spectral interpretation tasks while having their eye movements tracked. Following the completion of each task, individuals then participated in a cued RTA interview in which they watched a recording of their eye movements and verbalized in as much detail as possible what they were looking at and thinking about. Information on participants' prior coursework and research experience involving ^1H NMR and IR spectroscopy was collected prior to the

start of each session.

Eye tracking as a research tool. Eye tracking was used to collect quantitative data on participants' cognitive processes during spectral interpretation. Eye tracking involves measuring individuals' eye movements as they engage in visual tasks³⁷, and it has emerged as a powerful research tool for investigating cognitive processes involved in the comprehension of visualizations.³⁸ Two basic assumptions link eye movements with cognitive processes: the eye-mind assumption and the immediacy assumption. The eye-mind assumption states that individuals process information on which their eyes focus, and the immediacy assumption states that this processing occurs immediately.³⁷ Eye fixations, or the focusing of eye movements on visual stimuli, are thus of interest to researchers given that they provide insight into the information an individual processes in any given moment.³⁰ Given the inherently visual nature of spectral interpretation, eye tracking was a particularly well-suited tool for investigating the cognitive processes involved in this practice.

Eye fixations provide insight into a range of cognitive processes. Specifically, fixation counts provide a measure of how frequently an individual processes information, with research demonstrating that fixation counts provide an indication of the perceived importance of such information.³⁹ Further, the order in which individuals fixate on different information provides insight into patterns of visualization and processing strategies used to interpret the sum of a visual stimulus.^{40,41} The movement from one piece of information to another is referred to as a fixation transition. Research on fixation transitions in problem-solving contexts has demonstrated that the number of transitions is directly proportional to dwell time, a measure which reflects the perceived complexity or importance of related information.^{28,32,33} This research has also demonstrated a connection between the number of bidirectional transitions and the need to refresh working memory.³³

Cued RTA interviewing as a research tool. Cued RTA interviewing was used to collect qualitative data on participants' assumptions during spectral interpretation. This interviewing technique is often employed following eye tracking to further investigate

participants' thinking.⁴² During a cued RTA interview, a participant watches a recording of their eye movements overlaid on a recently completed visual task while simultaneously narrating everything they looked at and thought about. This interviewing technique elicits similar problem-solving process information when compared to other approaches like concurrent think-aloud interviewing.⁴³ It also allows participants to work in silence as they complete cognitively demanding tasks and thus reduce their cognitive load, or the amount of working memory storage they are using.

Description of interpretation tasks. All spectral interpretation tasks were of an identical format and included a prompt explaining that chemists attempted to synthesize a given compound, for which the structural formula was provided (Figure 3.1, Scheme 3.2, and Figures 3.4-3.7 in Supporting Information). The prompt went on to state that chemists analyzed their product spectroscopically to determine the success of the synthesis. Participants were then asked to evaluate the success of each synthesis using the provided spectroscopic data (¹H NMR and IR spectra). Response options were multiple choice and included “yes, the product was synthesized,” “no, the product was not synthesized,” and “not enough information to tell.” This prompt was developed using literature on common problem-types encountered by practicing organic chemists.⁴⁴ Spectra used in the study were obtained from the Spectral Database for Organic Compounds and are reproduced herein with permission from SDBSWeb.⁴⁵ To reduce undergraduate participants' cognitive load and allow for their completion of relatively complex tasks, ¹H NMR peaks were labelled with integration values and multiplicities. ¹H NMR and IR reference tables were also included with each task (Figure 3.7, Supporting Information). Prior to the study, a faculty member with over ten years of teaching experience in ¹H NMR spectroscopy was interviewed to identify spectral features that often create difficulty for undergraduates. Tasks were designed to include a variety of these features in order to elicit a range of chemical assumptions and processing strategies. These features included signals resulting from nonstandard coupling in Synthesis 1, overlapping signals in Synthesis 2, signals resulting from second order coupling in Synthesis 3, and the far downfield signal corresponding to the carboxylic acid hydrogen atom in Synthesis 3 (Figures 3.4-3.7, Supporting Information). The later feature

was included after the interviewee identified aldehyde hydrogen atoms and their far downfield shift due to anisotropy of the carbonyl group as creating difficulty for undergraduates; we thus anticipated that a signal even further downfield would compound this difficulty and probe participants' knowledge of acceptable variability in chemical shift.

Procedure. Eye movements were collected using a Tobii Pro X3-120 remote eye tracking system with a sampling rate of 120 Hz. Tobii systems provide a measure of eye movements by illuminating participants' pupils with near-infrared light and using image sensors to measure the pupils and reflection of light off the cornea.⁴⁶ This reflection can then be used with the system's image-processing algorithm to provide an estimation of participants' gaze on a visual stimulus. The eye tracking system was attached to a 24-inch HP monitor with a resolution of 1920 x 1200 pixels, and all interpretation tasks and subsequent eye movement recordings were displayed on the monitor. Participants completed a nine-point manual calibration prior to the start of each interpretation task to ensure the accuracy of eye movement recordings. During eye tracking, participants were allowed to move freely between pages containing the prompt and reference material. No time limit was imposed. Once participants had finished each task, they could progress to a selection page and indicate their response.

Cued RTA interviews were conducted using the Tobii Studio 3.4.8 RTA feature.⁴⁶ This tool provides simultaneous audio recording and playback of eye movement recordings. Participants were able to pause the playback of eye movement recordings when they required additional time to verbalize their thoughts, and the interviewer was able to pause this recording to ask probing questions about participants' thinking. Data collected from RTA interviews included audio-visual recordings of participants' eye movement recordings overlaid with their verbalized thoughts.

Distribution of responses. The distribution of responses to interpretation tasks suggests that doctoral participants (n=7) possessed relatively more expertise compared to undergraduates (n=18), providing empirical support for the comparison of their cognitive processes. For Syntheses 1 and 2, two-thirds (n=12) of undergraduates correctly indicated that the syntheses were successful, whereas all doctoral students

correctly responded to these tasks. A similar number of undergraduates (n=13) correctly indicated that Synthesis 3 was unsuccessful, while five doctoral participants provided this response. This distribution further suggests that while doctoral participants outperformed undergraduate participants, tasks were not too difficult for the undergraduate population or too simple for the doctoral population. Both groups were thus able to engage with these tasks, suggesting that findings may inform the design of instruction that is accessible to both undergraduate and doctoral students and effectively promotes expertise.

3.7.3 Qualitative analysis of RTA interviews

RTA interviews were analyzed qualitatively to characterize the assumptions that guided participants' reasoning during spectral interpretation. For this analysis, RTA interview responses were transcribed verbatim, and audio-visual recordings were used to clarify ambiguous references to spectral data. The first author then inductively coded all interviews for valid chemical assumptions, or scientifically accurate ideas about the nature and behavior of chemical and spectral features. This process involved generating descriptive and *in vivo* codes corresponding to specific ideas, followed by the refinement of codes and definitions to combine similar ideas.⁴⁷ To establish interrater reliability (IRR) with such a large number of codes, the first author then used the initial list of codes and definitions to generate codebooks specific to each interpretation task. Codes relating to task-specific molecular and spectral features were included in each respective codebook (e.g., assumptions about the NH NMR peak in Synthesis 1 were only included in the codebook specific to this task), and codes relating to multiple tasks were included in all codebooks. The first and second authors then independently coded 10% of RTA interview responses to Synthesis 1 using the task's codebook. IRR was calculated for this data using the Fuzzy kappa statistic, a derivative of Cohen's kappa that allows for the coding of a single unit using multiple codes.⁴⁸ A Fuzzy kappa statistic of 0.69 was obtained, indicating moderate agreement.⁴⁹ Code definitions were then modified, and exclusion lists were developed to improve reliability. For codes appearing in all codebooks, all definitions were uniformly modified. The first and second author then independently coded another 10% of RTA responses to Synthesis 1, and a Fuzzy kappa statistic of 0.82 was obtained, indicating strong agreement.⁴⁹ The first and second author then independently coded

10% of RTA responses to Synthesis 2 and 3 using the modified codebooks, and Fuzzy kappa statistics of 0.88 and 0.85 were obtained, respectively, both of which indicate strong agreement. The second author then deductively coded all responses using the modified codebooks.

As part of a previous investigation of undergraduate participants' invalid chemical assumptions, the first author inductively coded all undergraduate RTA responses for scientifically inaccurate ideas about chemical and spectral features.⁹ This inductive coding resulted in an initial codebook containing codes, definitions, and exclusion lists. For this previous investigation, the first and third authors then independently and deductively coded all undergraduate RTA responses using this codebook. These authors then discussed and revised codes until 100% consensus was obtained. For the current investigation, the first author deductively coded doctoral participants' RTA responses using the revised codebook from the previous investigation, as well as inductively coded to identify invalid assumptions specific to doctoral participants' interviews; old and new codes were then combined to generate a new codebook. The first and second authors then independently coded 15% of doctoral participants' RTA responses using this final codebook, and a Fuzzy kappa statistic of 1.0 was obtained, indicating perfect agreement. The first author then deductively coded all doctoral interview responses using this codebook. The attainment of perfect agreement without the revision of old codes and definitions from the previous investigation suggests that such codes were straightforward and reliably applied; an IRR coefficient was therefore not retrospectively calculated for these previously reported findings. Further, Connor *et al.* provide additional details about the reliability of this coded data.⁹

Following coding and the establishment of IRR, the first author then identified themes among codes using constant comparative analysis.²⁶ The identification of themes was facilitated by the regular writing of analytic memos and discussion of emerging themes with the corresponding author. Following thematic analysis, assumptions corresponding to each theme were then organized into levels of sophistication through a secondary constant comparative analysis to provide insight into how they may change with the development of expertise.

3.7.4 Quantitative analysis of eye-tracking data

Eye-tracking data were quantitatively analyzed by first defining AOIs around features within each task (Figure 3.4-3.7, Supporting Information). Defining AOIs allows for quantification of the number and duration of fixations on information contained within each region, as well as the number and order of fixation transitions between information in different regions. The size and location of AOIs were informed by the research questions and resolution of the eye-tracker. After AOIs were defined, raw gaze data were then converted to fixation data using Tobii Studio 3.4.8 software and a standard fixation threshold of 100 ms.^{11,35} Fixation data included a chronological sequence of all fixations on AOIs and their corresponding duration for each participant and task. With 25 participants each having completed three tasks, fixation data included 75 chronological fixation sequences.

Fixation counts. The Pearson product-moment correlation between fixation count and total fixation duration on AOIs was first evaluated to determine if one or both metrics should be used for subsequent analysis. When these metrics are highly correlated, fixation count provides an inferential measure of perceived importance. A low to moderate correlation in the form of high fixation count but low fixation duration would indicate predominantly searching behavior, meaning that fixation count could not be interpreted as a measure of perceived importance across AOIs.³⁰ Both metrics would then need to be evaluated for each AOI to determine which information, if any, is perceived as important. Fixation count and total fixation duration were highly correlated across all interpretation tasks ($r = 0.94$, $p < 0.001$), allowing subsequent analysis to focus only on fixation count.³⁰ Undergraduate and doctoral participants' fixation counts on each AOI were then compared using Mann-Whitney U tests to identify differences in visual attention between the two groups. This nonparametric test was employed given that the sample size restricted use of its parametric counterpart.⁵⁰ Effect sizes (r) were evaluated *post hoc* for all significant differences.⁵¹ Small effect sizes ranged from 0.1-0.3, medium ranged from 0.3-0.5, and large were > 0.5 .³¹

Fixation transitions. To compare fixation transitions between undergraduate and

doctoral participants, each fixation sequence was first transformed into a collapsed fixation sequence using the open-source software *eyePatterns*.⁵² A collapsed fixation sequence facilitates the identification of patterns by combining multiple sequential fixations within a single AOI into one fixation. Each collapsed fixation sequence was then used to generate a corresponding transition matrix containing observed frequencies of all possible transitions between AOIs. Undergraduate and doctoral participants' transition counts for each possible transition were then compared using Mann-Whitney U tests to identify any differences between the two groups. The median number of transitions was used as a measure of central tendency for all significant differences, and effect sizes (r) for these differences were evaluated *post hoc* using the previously described cutoff criteria. Subsequent analysis focused on significantly different transitions with nonzero differences in median rather than mean (45 of 53); the median is the appropriate measure of central tendency when using this nonparametric test, so these larger differences could be more reliably ascribed meaning. Significance was set at 0.05 for all statistical analyses, all of which were completed using the R Stats Package in RStudio.⁵³ Chord diagrams visualizing significantly different transitions were generated using the R package *circlize*.⁵⁴

3.7.5 Limitations

This study had inherent limitations that serve as important considerations during the design of instruction and curricular materials. Foremost, the interpretation tasks used during data collection may not have elicited all assumptions and cognitive processes involved in the interpretation of more authentic spectra. To make these tasks accessible to the undergraduate population, labelled integration values and multiplicities were included on all NMR spectra. Spectra were also free from peaks resulting from solvent or impurities to further reduce these participants' cognitive load and allow for task completion. Notably, undergraduate and doctoral participants still relied on assumptions involving the evaluation of multiplicity or solvent peaks despite these design considerations, with multiple doctoral participants even noting that multiplicity labels only corresponded to the apparent splitting of given peaks. Tasks therefore effectively elicited some assumptions and cognitive processes involved in the evaluation of this information;

however, this characterization may not be complete. Instructors should be aware of this limitation when using findings to inform instruction involving the interpretation of more authentic tasks.

Relatedly, IR spectra were provided as a complementary data source in this investigation because undergraduate participants regularly interpreted this form of data in the laboratory course from which they were recruited, thus making interpretation tasks accessible to both undergraduate and doctoral populations. However, practicing organic chemists no longer depend on IR spectroscopy to gain significant structural insight, instead opting for characterization techniques such as high resolution mass spectrometry. The degree to which findings transfer to contexts in which ^1H NMR is used in combination with more common characterization methods is thus uncertain, another limitation that instructors should consider when designing instruction. However, most areas of understanding contained assumptions relating to both NMR and IR spectra, suggesting that qualitative findings in part transfer to other characterization methods. Quantitative findings would also be expected to transfer, as using a different form of complementary data would likely not have altered undergraduates' search within NMR spectra and doctoral participants' search across spectra and prediction of expected NMR signals. Functional groups may have received less visual attention if complementary data were instead provided as a mass spectrum, though this potential difference does not threaten overall findings of the study.

Further, eye-tracking methodology necessitated the comparison of undergraduate and doctoral participants' eye movements, as considerable differences in expertise are typically needed to elicit measurable differences in such data.³⁸ While the comparison of undergraduate and doctoral participants' eye movements provides useful insight into how information processing changes with increasing expertise, it restricts the mapping of this progression to two levels. Moreover, cognitive processes are unfortunately difficult to characterize given their tacit nature; therefore, this limitation may simply be inherent to the domain of interest rather than the research tool of choice. Relatedly, undergraduate and doctoral students were selected as study participants to investigate changes in expertise, yet doctoral students do not necessarily possess more sophisticated

understanding or information processing ability. However, doctoral participants' almost exclusive use of more sophisticated assumptions, combined with their academic and research experience with spectral interpretation, provide support for this comparison. Statements comparing undergraduate and doctoral participants are therefore not intended to be absolute, but rather general to describe how expertise may broadly develop.

Additional limitations regarding eye-tracking methodology involved the close proximity of a subset of AOIs within NMR spectra, e.g., the OH singlet AOI and CH triplet and quartet AOI in Synthesis 1 (Figure 3.4, Supporting Information). While participants completed a calibration of eye movements prior to the start of each task, this calibration may have drifted slightly during task completion. Eye fixations recorded near the edge in one of these AOIs may have therefore been made in an adjacent AOI. For a subset of fixations on these AOIs, it is therefore difficult to be certain of which NMR signals participants fixated on. However, quantitative analysis focused on identifying differences in eye movements between undergraduate and doctoral participants. By focusing on differences in eye movements we are able to largely control for this source of error, as this drift would have occurred across groups. Further, findings indicate that undergraduates made similarly more transitions between both distant and closely-spaced AOIs in NMR spectra, suggesting recorded fixation locations were accurate and that the error associated with fixations near the edge of closely-spaced AOIs is minimal. This limitation therefore does not threaten the quantitative findings.

3.8 Supporting Information

Table 3.3. Exemplar foundational assumptions, including levels of sophistication and the number of undergraduate and doctoral participants holding assumptions from given levels.

Level	N _{Ug}	N _{PhD}	Exemplar assumptions
Understanding	18	7	<p>Integration values in an ¹H NMR spectrum correspond to the number of equivalent protons generating a peak</p> <p>¹H NMR and IR peaks will appear in a characteristic region given the type of proton or bond, respectively</p> <p>The N+1 rule can be applied to protons on atoms adjacent to an alkyl group to determine the splitting of the corresponding peak in an ¹H NMR spectrum</p> <p>The number of equivalent proton groups in a molecule corresponds to the number of peaks on an ¹H NMR spectrum</p>
Misunderstanding	6	0	<p>The N+1 rule can be applied to protons on (rather than adjacent to) an atom to determine the splitting of the corresponding peak</p> <p>The splitting of a peak is determined by using the absolute number of adjacent protons (N) rather than the absolute number plus one (N+1)</p> <p>De-shielding of protons causes the corresponding peak to shift upfield on an NMR spectrum rather than downfield</p>

Table 3.4. Exemplar assumptions about deviations from the N+1 rule, including levels of sophistication and the number of undergraduate and doctoral participants holding assumptions from given levels.

Level	NUg	NPhD	Exemplar assumptions
Upper	7	4	<p>Chemical properties of the amide bond (e.g., slow rotation) will influence the appearance of NMR signal(s) corresponding to adjacent methylene protons in Synthesis 1</p> <p>Terminal vinyl protons in Synthesis 3 are chemically inequivalent and thus do not obey the N+1 rule</p> <p>Vinyl protons undergo complex splitting, where the corresponding peaks labelled as apparent doublets in Synthesis 3 are actually doublets of doublets</p> <p>The NMR peak corresponding to the OH proton in Synthesis 1 will not be split due to hydrogen bonding (invalid)</p> <p>High temperatures will increase the diastereotopic character of methylene hydrogens adjacent to the NH group in Synthesis 1 (invalid)</p> <p>The temperature-dependent chemical inequivalence of methylene protons in Synthesis 1 means that the corresponding apparent quartet is actually the result of second-order splitting (invalid)</p>
Intermediate II	16	7	<p>NH or OH protons in Synthesis 1 can appear as singlets</p> <p>NH and OH protons can result in unexpected splitting of signals corresponding to adjacent methylene protons</p> <p>Aromatic protons can appear as multiplets rather than resolved, separate peaks</p> <p>Vinyl protons do not obey the N+1 rule</p>
Intermediate I	11	0	<p>The N+1 rule might not apply to NH or OH protons in Synthesis 1</p> <p>Vinyl hydrogens in Synthesis 3 might not obey the N+1 rule</p> <p>Either only the NH or OH proton affect splitting in Synthesis 1</p>
Lower	12	0	<p>NMR peaks corresponding to NH or OH protons in Synthesis 1 should not appear as singlets given that they have adjacent protons</p> <p>Terminal vinylic protons in Synthesis 3 should appear as a single doublet and not two apparent doublets</p>

Table 3.5. Exemplar assumptions about the variability of spectral data, including levels of sophistication and the number of undergraduate and doctoral participants holding assumptions from given levels.

Level	NU	NPh	Exemplar assumptions
Upper	14	4	<p>The NMR peak potentially corresponding to the aldehyde proton in Synthesis 3 should not be so far downfield since other electronegative atoms in the molecule would not result in such significant variability</p> <p>The NMR peak potentially corresponding to the aldehyde proton in Synthesis 3 is far downfield because it actually corresponds to a carboxylic acid proton</p> <p>The multiplet in Synthesis 2 corresponding to aromatic protons actually contains a doublet, with the electronegative oxygen atom in the molecule resulting in the resolution of this signal</p>
Intermediate II	15	2	<p>NMR peaks corresponding to aromatic protons are sometimes not distinguishable</p> <p>The chemical shift of OH protons is highly variable</p> <p>Peaks corresponding to OH protons may be missing from the NMR spectrum</p> <p>IR peaks can overlap</p> <p>Chemical shift values do not have to exactly match reference material</p>
Intermediate I	1	0	<p>The peak potentially corresponding to the aldehyde proton in Synthesis 3 could appear far downfield due to the electronegativity of the oxygen atom in the molecule</p>
Lower	11	0	<p>Chemical shift values should exactly match reference material</p> <p>IR peaks should be prominent and definite (i.e., non-overlapping) if functional groups are present</p> <p>Aromatic protons in Synthesis 2 should have distinct, separate corresponding NMR peaks</p>

Table 3.6. Exemplar visuospatial assumptions, including levels of sophistication and the number of undergraduate and doctoral participants holding assumptions from given levels.

Level	N	NPh	Exemplar assumptions
Upper	0	4	<p>The triplet-quartet pattern in Synthesis 1 is characteristic of an ethyl group</p> <p>The lone pair on the nitrogen atom in Synthesis 1 results in an asymmetric center, which then affects the stereotopicity of the implicit, adjacent methylene hydrogen atoms</p>
Intermediate II	7	4	<p>The structural formula in Synthesis 2 does not contain a plane of symmetry</p> <p>aromatic protons in Synthesis 2 are all inequivalent</p> <p>The signal potentially corresponding to the aldehyde proton in Synthesis 3 should be split by adjacent, implicit protons</p> <p>The multiplet corresponding to aromatic protons in Synthesis 2 contains a doublet</p> <p>The labeled apparent doublets in Synthesis 3 resemble two doublets of doublets upon close inspection of the spectrum's inset</p>
Intermediate I	2	0	<p>The structural formula of isochroman in Synthesis 2 might contain a plane of symmetry</p>
Lower	8	2	<p>The structural formula of isochroman in Synthesis 2 contains a plane of symmetry</p> <p>Only one hydrogen atom is attached to a methylene carbon</p>

Table 3.7. Exemplar practical assumptions, including levels of sophistication and the number of undergraduate and doctoral participants holding assumptions from given levels.

Level	N	NPh	Exemplar assumptions
Upper	1	3	<p>The carbonyl group will have different characteristic IR absorption frequencies depending on the type of compound in which it is a constituent (e.g., carboxylic acid, amide, ketone, etc.)</p> <p>The IR spectrum in Synthesis 2 should contain separate peaks near 3000 cm⁻¹ for sp³ C-H and sp² C-H stretches</p> <p>The degree of rotation around the amide bond in Synthesis 1 will vary with temperature and affect observed splitting within the NMR spectrum</p> <p>2D NMR experiments would facilitate the assignment of peaks resulting from second-order splitting in Synthesis 2</p>
Intermediate II	18	7	<p>The NMR peak with a chemical shift of 11 ppm in Synthesis 3, combined with the strong, broad IR peak near 3000 cm⁻¹, indicates the present of a carboxylic acid</p> <p>IR peaks correspond to the stretching of a bond rather than simply the presence of bonds or functional groups</p> <p>The fingerprint region of an IR spectrum provides variably useful information</p> <p>Peaks corresponding to solvent can appear in spectra</p> <p>Sample concentration can affect NMR signal intensity</p>
Intermediate I	6	0	<p>The IR peak near 3000 cm⁻¹ in Synthesis 3 was broader than expected for a CH group</p>
Lower	8	0	<p>The strong, broad IR peak near 3000 cm⁻¹ in Synthesis 3 corresponded to the CH functional group</p>

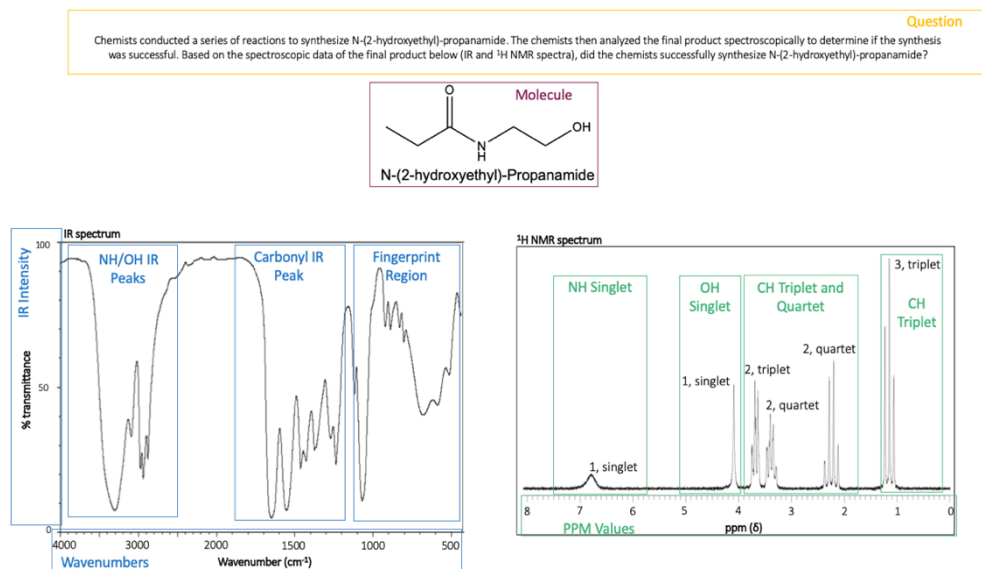


Figure 3.4. Spectral interpretation task (Synthesis 1) for which participants were asked to determine if N-(2-hydroxyethyl)-propanamide was successfully synthesized using the provided IR and ¹H NMR spectra. The provided spectra correctly represent the compound, so the correct answer for this task was “Yes, the synthesis was successful.” Participants’ eye movements were tracked and categorized according to several AOIS, represented in the figure by colored boxes. Colored labels within the boxes signify the AOI names.

Question

Chemists conducted a series of reactions to synthesize isochroman. The chemists then analyzed the final product spectroscopically to determine if the synthesis was successful. Based on the spectroscopic data of the final product below (IR and ^1H NMR spectra), did the chemists successfully synthesize isochroman?

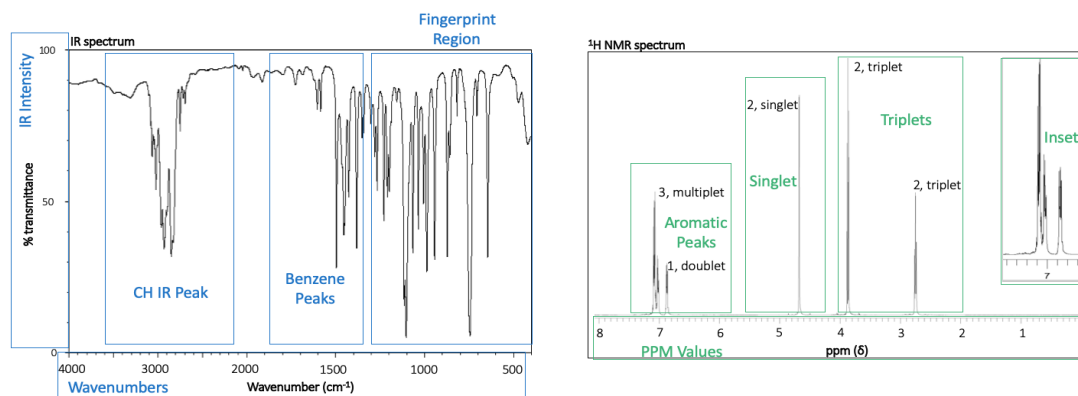
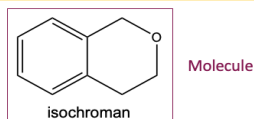


Figure 3.5. Spectral interpretation task (Synthesis 2) for which participants were asked to determine if isochroman was successfully synthesized using the IR and ^1H NMR spectra. Since the provided spectra correspond to isochroman, the correct answer for this task was “Yes, the synthesis was successful.” Colored boxes represent AOIs created for the task. Colored labels within the boxes signify the AOI names.

Question

Chemists conducted a series of reactions to synthesize 3-(allyloxy)propanal. The chemists then analyzed the final product spectroscopically to determine if the synthesis was successful. Based on the spectroscopic data of the final product below (IR and ^1H NMR spectra), did the chemists successfully synthesize 3-(allyloxy)propanal?

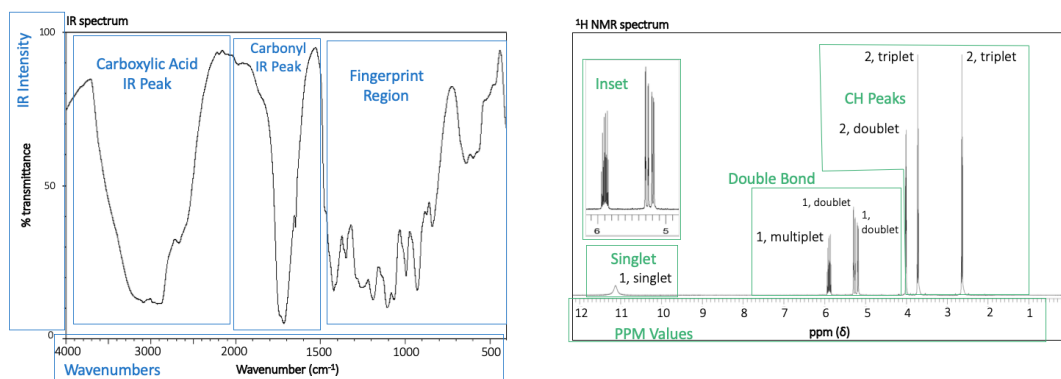
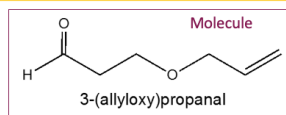


Figure 3.6. Spectral interpretation task (Synthesis 3) for which participants were asked to determine if 3-(allyloxy)propanal was successfully synthesized using the IR and ^1H NMR spectra. The spectra correspond to 3-(allyloxy)propionic acid, a carboxylic acid derivative of the provided molecule, so the correct answer for this task was “No, the synthesis was not successful.” Colored boxes represent AOIs created for the task, and AOI names appear in the accompanying colored text present either within or directly adjacent to the AOI boxes.

Reference Material

Infrared Spectroscopy: Typical IR absorption ranges			¹ H NMR Spectroscopy: Typical proton chemical shift values			
Type of Bond	Wavenumber (cm ⁻¹)	Intensity	Type of Hydrogen Atom	δ [*] (ppm)	Type of Hydrogen Atom	δ [*] (ppm)
N-H	3500-3300	medium, broad	RCH ₃	0.9		7.3
C≡C	2260-2100	medium to weak	RCH ₂ R	1.3		9.7
C=C	1680-1600	medium	cyclic	1.5	RNH ₂	1-3
	~1600 and ~1500-1430	strong to weak	R ₃ CH	1.5 - 2.0		3-5
C=O	1780-1650	strong		1.8		5-9
C-Cl	800-600	strong		2.0 - 2.3	ROH	1-5
C-O	1250-1050	strong		2.3		4-7
O-H (alcohol)	3650-3200	strong, broad	RC≡CH	2.5		10-13
O-H (carboxylic acid)	3300-2500	strong, very broad	RNHCH ₃	2-3		
C-H	3300-2700	medium	RCH ₂ X (X = Cl, Br, I)	3.5		
				3.8		
			R ₂ C=CH ₂	5.0		
			R ₂ C=CHR	5.3		

IR Table

NMR Table

Figure 3.7. IR and NMR tables provided as the second page in each of the interpretation tasks. While completing each task, participants were able to freely move between the two pages. Colored boxes represent AOIs, and AOI names appear in the accompanying colored text present adjacent to the AOI boxes. For this page of the task, AOIs were the same across all three syntheses.

Table 3.8. Areas of interest in each synthesis and corresponding letter codes used to report Mann-Whitney U test comparisons in Tables 3.9-3.11.

Areas of Interest			Corresponding Letter Code
Synthesis 1	Synthesis 2	Synthesis 3	
Question	Question	Question	A
Molecule	Molecule	Molecule	B
NH NMR Peak	CH IR Peaks	OH IR Peak	C
C=O IR Peak	Aromatic NMR Peaks	Singlet OH NMR Peak	D
NH IR Peak	Benzene IR Peak	C=O IR Peak	E
Fingerprint Region	Fingerprint Region	Fingerprint Region	F
OH NMR Peak	Singlet CH ₂ NMR Peak	Vinyl NMR Peaks	G
CH ₂ NMR Peaks	Triplet CH ₂ NMR Peaks	CH ₂ NMR Peaks	H
CH ₃ NMR Peaks	Aromatic NMR Peaks - Inset	Vinyl NMR Peaks - Inset	I
Wavenumber axis	Wavenumber Axis	Wavenumber Axis	J
PPM Axis	IR Intensity	IR Intensity	K
IR Intensity	PPM Axis	PPM Axis	L
IR Table	IR Table	IR Table	M
NMR Table	NMR Table	NMR Table	N

Table 3.9. Mann-Whitney U test comparisons of transition counts between undergraduate (n=18) and doctoral (n=7) participants for Synthesis 1, including the U statistic, p-value, and effect size for significant transitions. Letters in rows and columns correspond to areas of interest in Table S6. Rows correspond to the starting point of transitions, and columns correspond to the end point of transitions. *Corresponds to significance at the $p < 0.05$ level, **corresponds to significance at the $p < 0.01$ level, ***corresponds to significance at the $p < 0.001$ level, and NS corresponds to not significant. Medium effect sizes ranged from 0.3-0.5 and large effect sizes were > 0.5 .

		U													
AOI	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
A	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
B	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	23.5	ns	108.5	ns	
C	ns	ns	ns	ns	ns	ns	ns	ns	ns	103.5	ns	ns	ns	ns	
D	ns	ns	ns	ns	ns	ns	98.5	ns	ns	ns	ns	ns	ns	ns	
E	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	96.5	ns	
F	ns	ns	ns	ns	101.5	ns	ns	ns	ns	101.5	ns	ns	ns	ns	
G	ns	ns	ns	97	ns	ns	ns	100.5	ns	ns	ns	ns	ns	ns	
H	ns	ns	ns	ns	ns	ns	104	ns	104	ns	ns	ns	ns	ns	
I	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
J	ns	ns	99	ns	ns	91	ns	ns	ns	ns	ns	ns	ns	ns	
K	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	94.5	
L	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
M	36	98	ns	ns	98.5	ns	ns	ns	ns	ns	ns	ns	ns	ns	
N	ns	ns	ns	ns	ns	ns	ns	100	ns	ns	94.5	ns	ns	ns	
		p-value													
AOI	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
A	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
B	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.012*	ns	0.004**	ns	
C	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.013*	ns	ns	ns	ns	
D	ns	ns	ns	ns	ns	ns	0.032*	ns	ns	ns	ns	ns	ns	ns	
E	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.024*	ns	
F	ns	ns	ns	ns	0.016*	ns	ns	ns	ns	0.010*	ns	ns	ns	ns	
G	ns	ns	ns	0.039*	ns	ns	ns	0.024*	ns	ns	ns	ns	ns	ns	
H	ns	ns	ns	ns	ns	ns	0.014*	ns	0.013*	ns	ns	ns	ns	ns	
I	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
J	ns	ns	0.026*	ns	ns	0.042*	ns	ns	ns	ns	ns	ns	ns	ns	
K	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.029*	
L	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
M	0.004**	0.018*	ns	ns	0.026*	ns	ns	ns	ns	ns	ns	ns	ns	ns	
N	ns	ns	ns	ns	ns	ns	ns	0.018*	ns	ns	0.028*	ns	ns	ns	
		Effect size (r)													
AOI	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
A	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
B	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.510	ns	0.586	ns	
C	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.501	ns	ns	ns	ns	
D	ns	ns	ns	ns	ns	ns	0.435	ns	ns	ns	ns	ns	ns	ns	
E	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.459	ns	
F	ns	ns	ns	ns	0.488	ns	ns	ns	ns	0.520	ns	ns	ns	ns	
G	ns	ns	ns	0.418	ns	ns	ns	0.456	ns	ns	ns	ns	ns	ns	
H	ns	ns	ns	ns	ns	ns	0.499	ns	0.502	ns	ns	ns	ns	ns	
I	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
J	ns	ns	0.451	ns	ns	0.413	ns	ns	ns	ns	ns	ns	ns	ns	
K	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.445	
L	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
M	0.580	0.500	ns	ns	0.450	ns	ns	ns	ns	ns	ns	ns	ns	ns	
N	ns	ns	ns	ns	ns	ns	ns	0.479	ns	ns	0.447	ns	ns	ns	

Table 3.10. Mann-Whitney U test comparisons of transition counts between undergraduate (n=18) and doctoral (n=7) participants for Synthesis 2, including the U statistic, p-value, and effect size for significant transitions. Letters in rows and columns correspond to areas of interest in Table S6. Rows correspond to the starting point of transitions, and columns correspond to the end point of transitions. *Corresponds to significance at the $p < 0.05$ level, **corresponds to significance at the $p < 0.01$ level, ***corresponds to significance at the $p < 0.001$ level, and ns corresponds to not significant. Medium effect sizes ranged from 0.3-0.5 and large effect sizes were > 0.5 .

		U													
AOI	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
A	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
B	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	102.5	107.5	
C	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	97	ns	
D	ns	ns	ns	ns	ns	ns	ns	98	ns	ns	ns	ns	ns	ns	
E	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
F	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
G	ns	ns	ns	ns	ns	ns	ns	96.5	89.5	ns	ns	ns	ns	ns	
H	ns	ns	ns	96	ns	ns	98	ns	ns	ns	ns	ns	ns	ns	
I	ns	ns	ns	ns	ns	27	ns	101	ns	ns	ns	ns	ns	ns	
J	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	91	ns	
K	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
L	ns	ns	ns	107	ns	45	ns	ns	ns	ns	ns	ns	ns	ns	
M	ns	ns	ns	ns	95	ns	ns	ns	ns	ns	ns	ns	ns	117.5	
N	ns	ns	ns	ns	ns	ns	ns	91	ns	ns	ns	ns	112	ns	

		p-value													
AOI	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
A	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
B	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.011*	0.005**	
C	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.029*	ns	
D	ns	ns	ns	ns	ns	ns	ns	0.037*	ns	ns	ns	ns	ns	ns	
E	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
F	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
G	ns	ns	ns	ns	ns	ns	ns	0.049*	0.043*	ns	ns	ns	ns	ns	
H	ns	ns	ns	0.043*	ns	ns	0.033*	ns	ns	ns	ns	ns	ns	ns	
I	ns	ns	ns	ns	ns	0.0007***	ns	0.017*	ns	ns	ns	ns	ns	ns	
J	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.044*	ns	
K	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
L	ns	ns	ns	0.007**	ns	0.024	ns	ns	ns	ns	ns	ns	ns	ns	
M	ns	ns	ns	ns	0.043*	ns	ns	ns	ns	ns	ns	ns	ns	0.0006***	
N	ns	ns	ns	ns	ns	ns	ns	0.042*	ns	ns	ns	ns	0.002**	ns	

		Effect Size (r)													
AOI	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
A	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
B	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.510	0.563	
C	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.443	ns	
D	ns	ns	ns	ns	ns	ns	ns	0.440	ns	ns	ns	ns	ns	ns	
E	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
F	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
G	ns	ns	ns	ns	ns	ns	ns	0.411	0.368	ns	ns	ns	ns	ns	
H	ns	ns	ns	0.411	ns	ns	0.432	ns	ns	ns	ns	ns	ns	ns	
I	ns	ns	ns	ns	ns	0.686	ns	0.484	ns	ns	ns	ns	ns	ns	
J	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.044	ns	
K	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
L	ns	ns	ns	0.543	ns	0.463	ns	ns	ns	ns	ns	ns	ns	ns	
M	ns	ns	ns	ns	0.411	ns	ns	ns	ns	ns	ns	ns	ns	0.691	

Table 3.11. Mann-Whitney U test comparisons of transition counts between undergraduate (n=18) and doctoral (n=7) participants for Synthesis 3, including the U statistic, p-value, and effect size for significant transitions. Letters in rows and columns correspond to areas of interest in Table S6. Rows correspond to the starting point of transitions, and columns correspond to the end point of transitions. *Corresponds to significance at the $p < 0.05$ level, **corresponds to significance at the $p < 0.01$ level, ***corresponds to significance at the $p < 0.001$ level, and ns corresponds to not significant. Medium effect sizes ranged from 0.3-0.5 and large effect sizes were > 0.5

		U													
AOI	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
A	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
B	ns	ns	ns	ns	ns	104	ns	ns	ns	ns	ns	ns	104	104	
C	ns	ns	ns	ns	32	ns	ns	ns	39.5	ns	ns	ns	94.5	ns	
D	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
E	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	101.5	ns	
F	ns	101.5	ns	ns	ns	106	ns	ns	ns	ns	ns	ns	ns	ns	
G	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
H	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
I	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	91	
J	ns	112	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
K	ns	ns	45	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
L	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
M	ns	99	ns	ns	96	97.5	ns	ns	ns	94.5	ns	ns	ns	ns	
N	ns	98	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
		p-value													
AOI	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
A	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
B	ns	ns	ns	ns	ns	0.011*	ns	ns	ns	ns	ns	ns	0.011*	0.012*	
C	ns	ns	ns	ns	0.046*	ns	ns	ns	0.028*	ns	ns	ns	0.027*	ns	
D	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
E	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.010*	ns	
F	ns	0.015*	ns	ns	ns	0.009**	ns	ns	ns	ns	ns	ns	ns	ns	
G	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
H	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
I	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.043*	
J	ns	0.002**	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
K	ns	ns	0.02431*	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
L	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
M	ns	0.021*	ns	ns	0.033*	0.030*	ns	ns	ns	0.029*	ns	ns	ns	ns	
N	ns	0.018*	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
		Effect Size (r)													
AOI	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
A	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
B	ns	ns	ns	ns	ns	0.516	ns	ns	ns	ns	ns	ns	0.513	0.510	
C	ns	ns	ns	ns	0.405	ns	ns	ns	0.448	ns	ns	ns	0.448	ns	
D	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
E	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.520	ns	
F	ns	0.491	ns	ns	ns	0.531	ns	ns	ns	ns	ns	ns	ns	ns	
G	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
H	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
I	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	0.412	
J	ns	0.632	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
K	ns	ns	0.4633	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
L	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	
M	ns	0.469	ns	ns	0.433	0.440	ns	ns	ns	0.445	ns	ns	ns	ns	
N	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	

3.8.1 Additional description of areas of understanding, levels, and contributing assumptions

Through the coding of undergraduate and doctoral students' RTA interview responses, we identified a total of 83 distinct chemical assumptions. Among these 83 assumptions, we identified five themes that more comprehensively describe the assumptions that guided participants' thinking during spectral interpretation. These themes represent areas of understanding that are necessary for interpreting NMR spectra. Further, the majority of these areas demonstrated four levels of sophistication. Notably, undergraduate participants used assumptions across all levels of sophistication, whereas doctoral participants used assumptions almost exclusively from the uppermost levels (Tables 3.3-3.7). This finding suggests that the identified progression of assumptions validly reflects how understanding evolves with increasing expertise. The areas of understanding and their corresponding levels are described in detail below.

Foundational assumptions. Responses from all undergraduate and doctoral participants included foundational assumptions, or ideas relating to basic fundamental principles that underly spectral data. These assumptions were termed foundational given that they appear to serve as a groundwork on which other classes of assumptions build. Foundational assumptions accounted for 34% of all coded assumptions and reflected either an understanding or misunderstanding of fundamental principles, thus demonstrating two levels of sophistication (Table 3.3). They included ideas such as the N+1 rule can be applied to protons on atoms adjacent to an alkyl group to determine the splitting of the corresponding peak in an ^1H NMR spectrum. Foundational assumptions that reflected a misunderstanding of fundamental principles appeared in responses from several undergraduate participants ($n=6$), though they accounted for only 1% of all coded assumptions. They included ideas such as the N+1 rule can be applied to protons on (rather than adjacent to) an atom to determine the splitting of the corresponding peak.

Assumptions about deviations from the N+1 rule. All interpretation tasks included in this study incorporated at least one molecular feature for which the N+1 rule cannot be applied to correctly predict the multiplicity of a corresponding NMR peak.

Building off of foundational assumptions about *how* to apply the N+1 rule were assumptions relating to such contexts where the N+1 rule cannot be applied to successfully determine multiplicity. This second class of assumptions reflected ideas about the acceptability of such deviations from the N+1 rule. Such ideas accounted for 15% of all coded assumptions (Table 3.4). Successive levels of sophistication for this area contained assumptions that reflected increasing knowledge of (1) contexts in which the N+1 rule does not apply and (2) other implicit chemical features affecting splitting. Doctoral participants relied on assumptions exclusively from the two uppermost levels of this area, suggesting that developing expertise corresponds to increasing understanding of these factors (Table 3.4). This progression aligns with Cartrette and Bodner's finding that individuals more successful at interpreting spectral data were those who demonstrated flexible use of the N+1 rule.⁵

The lower level of sophistication corresponded to limited knowledge of these factors and contained invalid assumptions that the N+1 rule should apply in contexts where the rule fails. These assumptions appeared exclusively in the responses of undergraduate participants (n=12) and included ideas such as NMR peaks corresponding to NH or OH protons in Synthesis 1 should not appear as singlets given that they have adjacent protons. The next level corresponded to developing knowledge of appropriate contexts in which the N+1 rule does not apply. This level contained valid assumptions reflecting doubt or uncertainty regarding whether the N+1 rule applies in contexts where the rule fails. These assumptions also appeared exclusively in responses from undergraduate participants (n=11) and included ideas such as the N+1 rule *might* not apply to NH or OH protons in Synthesis 1. Following the first intermediate level, the second intermediate level contained valid assumptions reflecting knowledge of specific contexts in which the N+1 rule does not apply, though knowledge of implicit chemical features affecting splitting was still not expressed. These assumptions appeared in responses from the majority of both undergraduate (n=16) and doctoral (n=7) participants and included ideas such as that NH or OH protons in Synthesis 1 can appear as singlets.

The upper level of sophistication contained primarily valid assumptions relating to implicit chemical features affecting splitting and thus went beyond identifying specific,

valid contexts in which the N+1 rule does not apply. These assumptions appeared in responses from approximately half of both undergraduate (n=7) and doctoral (n=4) participants. Valid assumptions in this level included ideas such as the chemical properties of the amide bond (e.g., slow rotation) will influence the appearance of NMR signal(s) corresponding to adjacent methylene protons in Synthesis 1. The upper level also contained three invalid assumptions relating to implicit chemical features; while these ideas were invalid, they still focused on implicit features. Their use can thus be viewed as a productive step toward engaging in sophisticated thinking. Among these ideas was the belief that the NMR peak corresponding to the OH proton in Synthesis 1 will not be split due to hydrogen bonding.

Assumptions about the variability of spectral data. Building off of foundational assumptions that peaks will appear in a characteristic region on a spectrum were assumptions about the extent to which spectral data can demonstrate variability in resolution, absorption frequency, and signal intensity. This class of assumptions accounted for 16% of all codes. Successive levels of sophistication for this area corresponded to increasing knowledge of (1) contexts in which variability of spectral data is acceptable and (2) other implicit chemical features resulting in variability. Doctoral participants used assumptions exclusively from the two uppermost levels of this area as well, suggesting that developing expertise also corresponds to increasing understanding of these factors (Table 3.5).

The lower level of this area corresponded to limited knowledge of these factors and included invalid assumptions that spectral data should not vary in contexts where such variability is acceptable. Like lower-level assumptions in the previous area, these assumptions also appeared exclusively in responses from undergraduate participants (n=11). Included in this level were ideas such as chemical shift values should exactly match reference material and IR peaks should be prominent and definite (i.e., non-overlapping) if functional groups are present. The next level corresponded to developing knowledge of appropriate contexts and other implicit chemical features resulting in variability. Within this level was one invalid assumption reflecting the idea that the peak potentially corresponding to the aldehyde proton in Synthesis 3 could appear far

downfield due to the electronegativity of the oxygen atom in the molecule. This assumption appeared in the response of one undergraduate participant. The second intermediate level corresponded to knowledge of appropriate contexts in which spectral data can vary, though knowledge of other implicit features was not expressed. These valid assumptions appeared in responses from both undergraduate (n=15) and doctoral (n=2) participants. Included in this level were ideas such as the chemical shift of OH protons is highly variable.

The upper level of this area corresponded to knowledge of implicit chemical features and their effect on the variability of spectral data. Valid assumptions in this level appeared in responses from several undergraduate (n=14) and doctoral (n=4) participants. The most common of these assumptions was that the NMR peak potentially corresponding to the aldehyde proton in Synthesis 3 should not be so far downfield since other electronegative atoms in the molecule would not result in such significant variability in chemical shift. Other participants correctly assumed that this peak is far downfield because it actually corresponds to a carboxylic acid proton.

Visuospatial assumptions. Participants' reasoning was guided by a number of visuospatial assumptions. As noted in the primary text, visuospatial information is visual in nature and exhibits spatial properties involving the representation of space and relationships between entities within a given space.⁵⁵ Chemistry education researchers expand on this definition and refer to visuospatial thinking as the thought processes involved in the identification of spatial features, as well as the generation and recognition of structural formulae and other symbols.⁵⁵ Structural formulae and NMR spectra are two of the most common visualizations used by practicing chemists.¹⁸ For this study, visuospatial assumptions included ideas about spatial features within provided structural formulae and spectra, as well as the recognition of structural formulae. These assumptions accounted for 6% of all coded assumptions. Successive levels of sophistication reflected an increasing ability to reason about spatial features within provided structural formulae and spectra, as well as to recognize and correctly reason with structural formulae. Doctoral participants used assumptions predominantly from the

two uppermost levels of this area, suggesting that increasing understanding of visuospatial features is another essential aspect of developing expertise (Table 3.6).

The lower level of this area contained invalid assumptions reflecting limited visuospatial ability, specifically difficulty with identifying planes of symmetry and implicit hydrogens within structural formulae. These assumptions appeared predominantly in responses from undergraduate participants (n=8), though they also appeared in responses from some doctoral participants (n=2). They included ideas such as the structural formula of isochroman in Synthesis 2 contains a plane of symmetry. The next level corresponded to developing visuospatial ability; within this level was one valid assumption reflecting doubt or uncertainty regarding whether the structural formula in Synthesis 2 was symmetric. This assumption appeared exclusively in the responses of undergraduate participants (n=2). The second intermediate level contained valid assumptions reflecting general visuospatial ability, including an ability to identify planes of symmetry, chemically equivalent protons, and implicit hydrogens within structural formulae, in addition to relevant but inconspicuous spectral features. These assumptions appeared in responses from both undergraduate (n=7) and doctoral (n=4) participants. Assumptions reflecting an ability to recognize and correctly reason with structural formulae included ideas such as the structural formula in Synthesis 2 does not contain a plane of symmetry. Assumptions relating to spatial features within spectra involved participants identifying relevant yet poorly resolved spectral features. They included the valid idea that the multiplet corresponding to aromatic protons in Synthesis 2 contains a doublet.

The upper level of this area corresponded to sophisticated visuospatial ability, with assumptions reflecting recognition of characteristic NMR splitting patterns involving more than one peak and a developing ability to identify asymmetric centers and the stereotopicity of implicit hydrogens in structural formulae. These assumptions appeared exclusively in the responses of doctoral participants (n=4). Several participants (n=4) recognized the characteristic triplet-quartet pattern of an ethyl group in Synthesis 1. Further, one participant incorrectly assumed the lone pair on the nitrogen atom in

Synthesis 1 results in an asymmetric center, which then affects the stereotopicity of the implicit, adjacent methylene hydrogen atoms.

Practical assumptions. Practical assumptions encompassed knowledge of additional peak characteristics beyond N+1 rule deviations and acceptable variability in resolution, absorption frequency, and signal intensity. They also reflected knowledge of experimental variables that influence the appearance of spectral data. These ideas were termed practical given that they reflect knowledge which could be cultivated through practical experience interpreting spectra. They accounted for 29% of all codes. Successive levels for this area corresponded to increasing familiarity with additional peak characteristics such as characteristic absorption frequency and intensity, as well as experimental variables. Again, doctoral participants used assumptions almost exclusively from the two uppermost levels, suggesting that increasing understanding of these factors contributes to increasing expertise (Table 3.7).

Lower-level assumptions reflected a lack of familiarity with these factors and appeared almost exclusively in responses from undergraduate participants (n=8), with one doctoral participant exhibiting such thinking. Invalid assumptions comprising this level most commonly involved participants misidentifying characteristic IR peaks as those appearing in the same wavenumber region but corresponding to functional groups with different characteristic absorption intensities. The most common assumption involved undergraduate participants' (n=5) belief that the strong, broad IR peak near 3000 cm^{-1} in Synthesis 3 corresponded to the CH functional group. Assumptions in this level also reflected a lack of familiarity with characteristic chemical shift values, including the idea that signals corresponding to aldehyde protons typically appear from 10-12 ppm. The above level corresponded to developing familiarity with additional peak characteristics, though knowledge of experimental variables was still limited or not expressed. These assumptions also appeared exclusively in the responses of undergraduate participants (n=6). They included valid ideas such as the IR peak near 3000 cm^{-1} in Synthesis 3 was broader than expected for a CH group. The second intermediate level reflected general familiarity with additional peak characteristics and some knowledge of experimental variables that influence the appearance of spectral data. Valid assumptions comprising

this level appeared in the responses of all undergraduate (n=18) and doctoral (n=7) participants. Several assumptions in this level involved participants correctly correlating IR peaks with particular functional groups or bonds. Valid assumptions relating to experimental variables included ideas such as peaks corresponding to solvent can appear in spectra.

The upper level of this area corresponded to familiarity with specific aspects of additional peak characteristics, as well as developed knowledge of experimental variables influencing the appearance of spectra. Valid assumptions comprising this level appeared predominantly in the responses of doctoral participants (n=3), with one appearing in the response of an undergraduate participant. Assumptions reflecting familiarity with specific aspects of additional peak characteristics included ideas such as the IR spectrum in Synthesis 2 should contain separate peaks near 3000 cm^{-1} for sp^3 C-H and sp^2 C-H stretches. This specific information was not provided in the reference material, and its consideration mirrors other studies' finding that individuals more successful at spectral interpretation adopt a dynamic rather than static view of bonding.^{35,56} Assumptions reflecting developed knowledge of experimental variables included ideas that the degree of rotation around the amide bond in Synthesis 1 will vary with temperature and affect observed splitting within the NMR spectrum.

3.8.2 Description of transitions not described in the text

Significantly different transitions not described in the text (3 out of 45) involved undergraduate and doctoral participants moving more between the molecule and spectral features, including spectral peaks and frequency axes. While these differences do not collectively suggest a single difference in cognitive processes, they do individually support other findings. For instance, undergraduate participants made significantly more transitions between the fingerprint region and the molecule in Synthesis 3, which aligns with their increased visual attention on the fingerprint region in other tasks. Further, undergraduates made more transitions from the wavenumber axis to the molecule (Syn. 3, Figure 3.3), whereas doctoral participants made more transitions from the molecule to the ppm axis

(Syn. 1, Figure 3.3); this finding further supports the notion that doctoral participants first looked to the molecule to determine the peak they would expect, whereas undergraduates used an alternative strategy.

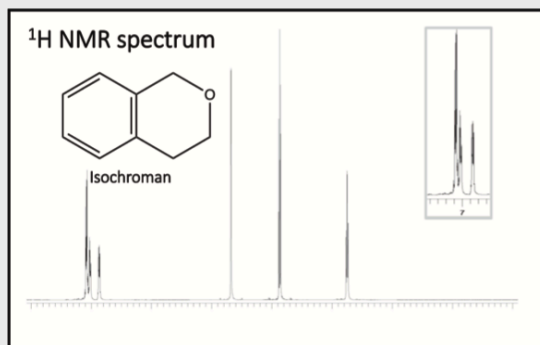
3.8.3 Instructional model for graduate and undergraduate-level instructors

The analysis of assumptions provides instructors with a map for cultivating students' conceptual understanding for NMR spectral interpretation. This map suggests that instruction and curricular materials should be designed around the five identified areas of understanding, with an initial focus on basic principles. Subsequent focus on the four areas building upon this foundation should then ideally involve scaffolded interpretation activities that separately encourage the evaluation of implicit molecular features affecting splitting, implicit molecular features affecting the variability of data, implicit visuospatial features of spectra and structural formulae, characteristic spectral features such as chemical shift and absorption intensity, and the effect of experimental variables on spectral appearance. For instance, instructors could incorporate activities that require students to predict how differences in the electronegativity of nearby atoms will affect protons' chemical shift, as well as activities that require students to explain how implicit chemical features (e.g., proton exchange) affect observed splitting patterns. Activities that integrate the evaluation of these components (e.g., complete structure elucidation) can then be incorporated once students demonstrate understanding in each individual area.

This analysis also provides instructors with insight into the design of appropriate formative assessments, or evaluations of learning throughout a course. Given that limited understanding in a single area can derail students' interpretation efforts, the design of assessments that individually evaluate understanding in each dimension would provide instructors with a means of identifying specific areas of both understanding and difficulty. For example, instructors could incorporate multi-tiered test questions that require students to separately evaluate molecular symmetry and splitting. They could also design open-ended questions that require students to explain how specific structural features result in spectral peaks. Such assessments would allow instructors to determine if students are struggling with visuospatial aspects or if they should instead focus additional

A chemist conducted a series of reactions to synthesize isochroman and analyzed the product spectroscopically. Based on the data provided, determine whether the synthesis was successful.

A. Use the data provided to explain your answer.



B. Refer to specific structural features of isochroman that support your explanation.

Figure 3.8. Open-ended assessment that requires students to explain how specific structural features result in spectral peaks.

instruction on basic principles underlying the N+1 rule and implicit features resulting in deviations. An example of an open-ended assessment is provided in Figure 3.8.

Differences in information processing between the two groups have additional implications for the design of instruction and curricular materials. These findings indicate that undergraduates worked to process information that conflicted with their limited understanding and familiarity, unlike doctoral participants who demonstrated an existing familiarity and focused their attention more efficiently. These findings underscore the importance of cultivating understanding across the five identified areas; an existing familiarity with characteristic chemical shift values, acceptable variability in chemical shift, and acceptable deviations from the N+1 rule will allow undergraduates to efficiently focus their attention on information relevant to decision making. These results further suggest that visualizations used in instruction and curricular materials should only contain such information, however minor, after cultivating relevant understanding. Such visualizations will help condition novice interpreters to process spectral information in a holistic and unbiased manner, shifting them from processing that disproportionately focuses on unexpected or unfamiliar features. Differences in information processing between groups

lastly suggest that instruction should include interpretation activities that encourage the search for complementary pieces of data between spectra and the prediction of expected spectral peaks. Such activities have potential for shifting novice interpreters from searching behavior toward more informed and efficient information processing.

3.9 Acknowledgements

This work is supported by the University of Michigan Department of Chemistry.

3.10 References

- (1) Schummer, J. The Impact of Instrumentation on Chemical Species Identity: From Chemical Substances to Molecular Species. In *From Classical to Modern Chemistry: The Instrumental Revolution*; Morris, P., Ed.; The Royal Society of Chemistry: Cambridge, 2002; pp 188–211.
- (2) Groleau, R. R.; Chapman, R. S. L.; Ley-Smith, H.; Liu, L.; James, T. D.; Bull, S. D. A Three-Component Derivatization Protocol for Determining the Enantiopurity of Sulfinamides by ^1H and ^{19}F NMR Spectroscopy. *J. Org. Chem.* **2020**, *85* (2), 1208–1215. <https://doi.org/10.1021/acs.joc.9b02473>.
- (3) Chrominski, M.; Baranowski, M. R.; Chmielinski, S.; Kowalska, J.; Jemielity, J. Synthesis of Trifluoromethylated Purine Ribonucleotides and Their Evaluation as ^{19}F NMR Probes. *J. Org. Chem.* **2020**, *85* (5), 3440–3453. <https://doi.org/10.1021/acs.joc.9b03198>.
- (4) Jennings, J. J.; Wigman, B. W.; Armstrong, B. M.; Franz, A. K. NMR Quantification of the Effects of Ligands and Counterions on Lewis Acid Catalysis. *J. Org. Chem.* **2019**, *84* (24), 15845–15853. <https://doi.org/10.1021/acs.joc.9b02107>.
- (5) Cartrette, D. P.; Bodner, G. M. Non-Mathematical Problem Solving in Organic Chemistry. *J. Res. Sci. Teach.* **2010**, *47* (6), 643–660. <https://doi.org/10.1002/tea.20306>.
- (6) Hoye, T. R.; Hanson, P. R.; Vyvyan, J. R. A Practical Guide to First-Order Multiplet Analysis in ^1H NMR Spectroscopy. *J. Org. Chem.* **1994**, *59* (15), 4096–4103. <https://doi.org/10.1021/jo00094a018>.
- (7) Hoye, T. R.; Zhao, H. A Method for Easily Determining Coupling Constant Values: An Addendum to “a Practical Guide to First-Order Multiplet Analysis in ^1H NMR Spectroscopy.” *J. Org. Chem.* **2002**, *67* (12), 4014–4016. <https://doi.org/10.1021/jo001139v>.

- (8) Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. NMR Chemical Shifts of Common Laboratory Solvents as Trace Impurities. *J. Org. Chem.* **1997**, *62* (21), 7512–7515. <https://doi.org/10.1021/jo971176v>.
- (9) 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>.
- (10) Stowe, R. L.; Cooper, M. M. Arguing from Spectroscopic Evidence. *J. Chem. Educ.* **2019**, *96* (10), 2072–2085. <https://doi.org/10.1021/acs.jchemed.9b00550>.
- (11) Topczewski, J. J.; Topczewski, A. M.; Tang, H.; Kendhammer, L. K.; Pienta, N. J. NMR Spectra through the Eyes of a Student: Eye Tracking Applied to NMR Items. *J. Chem. Educ.* **2017**, *94* (1), 29–37. <https://doi.org/10.1021/acs.jchemed.6b00528>.
- (12) 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>.
- (13) Margolis, E. A Reassessment of the Shift from the Classical Theory of Concepts to Prototype Theory. *Cognition* **1994**, *51* (1), 73–89. [https://doi.org/10.1016/0010-0277\(94\)90009-4](https://doi.org/10.1016/0010-0277(94)90009-4).
- (14) Gelman, S. A. Learning from Others: Children's Construction of Concepts. *Annu. Rev. Psychol.* **2009**, *60*, 115–140. <https://doi.org/10.1146/annurev.psych.59.103006.093659>.
- (15) Murphy, G. L.; Medin, D. L. The Role of Theories in Conceptual Coherence. *Psychol. Rev.* **1985**, *92* (3), 289–316. <https://doi.org/10.1037/0033-295X.92.3.289>.
- (16) 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>.
- (17) Sevian, H.; Talanquer, V. Rethinking Chemistry: A Learning Progression on Chemical Thinking. *Chem. Educ. Res. Pract.* **2014**, *15*, 10–23. <https://doi.org/10.1039/c3rp00111c>.
- (18) Kozma, R.; Russell, J. Students Becoming Chemists: Developing Representational Competence. In *Visualization in Science Education*; Gilbert, J. K., Ed.; Springer: Dordrecht, 2005; pp 121–146. https://doi.org/10.1007/1-4020-3613-2_4.
- (19) Haider, H.; Frensch, P. A. Eye Movement during Skill Acquisition: More Evidence for the Information-Reduction Hypothesis. *J. Exp. Psychol. Learn. Mem. Cogn.* **1999**, *25* (1), 172–190. <https://doi.org/10.1037/0278-7393.25.1.172>.

- (20) Cowan, N. The Magical Number 4 in Short-Term Memory: A Reconsideration of Mental Storage Capacity. *Behav. Brain Sci.* **2001**, *24* (1), 87–114. <https://doi.org/10.1017/S0140525X01003922>.
- (21) Kundel, H. L.; Nodine, C. F.; Conant, E. F.; Weinstein, S. P. Holistic Component of Image Perception in Mammogram Interpretation: Gaze-Tracking Study. *Radiology* **2007**, *242* (2), 396–402.
- (22) Ericsson, K. A.; Kintsch, W. Long-Term Working Memory. *Psychol. Rev.* **1995**, *102* (2), 211–245.
- (23) 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>.
- (24) Smith, A. D.; Kelly, A. Cognitive Processes. In *The Encyclopedia of Adulthood and Aging*; Whitbourne, S. K., Ed.; John Wiley & Sons, Inc., 2016; pp 1–4. https://doi.org/10.1007/978-3-319-58598-7_4.
- (25) Saldana, J. *The Coding Manual for Qualitative Researchers*; Sage Publications: Los Angeles, 2016.
- (26) Creswell, J. W.; Poth, C. N. *Qualitative Inquiry and Research Design*, 4th ed.; Sage Publications: Los Angeles, 2018.
- (27) Braun, V.; Clarke, V. Using Thematic Analysis in Psychology. *Qual. Res. Psychol.* **2006**, *3*, 77–101.
- (28) Cullipher, S.; Hansen, S. J. R.; Vandenplas, J. R. Eye Tracking as a Research Tool: An Introduction. In *Eye Tracking for the Chemistry Education Researcher*; American Chemical Society: Washington, DC, 2018; pp 1–9.
- (29) Harle, M.; Towns, M. A Review of Spatial Ability Literature, Its Connection to Chemistry, and Implications for Instruction. *J. Chem. Educ.* **2011**, *88* (3), 351–360. <https://doi.org/10.1021/ed900003n>.
- (30) Cullipher, S.; Vandenplas, J. R. Using Fixations to Measure Attention. In *Eye Tracking for the Chemistry Education Researcher*; American Chemical Society: Washington, DC, 2018; pp 53–72. <https://doi.org/10.1021/bk-2018-1292.ch004>.
- (31) Cohen, J. A Power Primer. *Quant. Methods Psychol.* **1992**, *112* (1), 155–159. <https://doi.org/10.1016/j.jorganchem.2011.01.025>.
- (32) Day, E. L.; Tang, H.; Kendhammer, L. K.; Pienta, N. J. Sequence Analysis: Use of Scanpath Patterns for Analysis of Students' Problem-Solving Strategies. *ACS Symp. Ser.* **2018**, *1292*, 73–97. <https://doi.org/10.1021/bk-2018-1292.ch005>.

- (33) Holmqvist, K.; Nystrom, M.; Andersson, R.; Dewhurst, R.; Jarodzka, H.; van de Weijer, J. *A Comprehensive Guide to Methods and Measures*; Oxford University Press, 2011.
- (34) Kozma, R. The Material Features of Multiple Representations and Their Cognitive and Social Affordances for Science Understanding. *Learn. Instr.* **2003**, *13* (2), 205–226. [https://doi.org/10.1016/s0959-4752\(02\)00021-x](https://doi.org/10.1016/s0959-4752(02)00021-x).
- (35) Cullipher, S.; Sevian, H. Atoms versus Bonds: How Students Look at Spectra. *J. Chem. Educ.* **2015**, *92* (12), 1996–2005. <https://doi.org/10.1021/acs.jchemed.5b00529>.
- (36) Cohen, L.; Manion, L.; Morrison, K. *Research Methods in Education*, 7th ed.; Routledge: London, 2011.
- (37) Just, M. A.; Carpenter, P. A. A Theory of Reading: From Eye Fixations to Comprehension. *Psychol. Rev.* **1980**, *87* (4), 329–354.
- (38) Gegenfurtner, A.; Lehtinen, E.; Säljö, R. Expertise Differences in the Comprehension of Visualizations: A Meta-Analysis of Eye-Tracking Research in Professional Domains. *Educ Psychol Rev* **2011**, *23*, 523–552. <https://doi.org/10.1007/s10648-011-9174-7>.
- (39) Green, H. J.; Lemaire, P.; Dufau, S. Eye Movement Correlates of Younger and Older Adults' Strategies for Complex Addition. *Acta Psychol. (Amst)*. **2007**, *125* (3), 257–278. <https://doi.org/10.1016/j.actpsy.2006.08.001>.
- (40) Rayner, K. Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychol. Bull.* **1998**, *124* (3), 372–422. <https://doi.org/10.1037/0033-2909.124.3.372>.
- (41) Grant, E. R.; Spivey, M. J. Eye Movements and Problem Solving. *Am. Psychol. Soc.* **2003**, *14* (5), 462–466. <https://doi.org/10.1111/1467-9280.02454>.
- (42) Hyrskykari, A. Gaze Path Stimulation in Retrospective Think-Aloud. *J. Eye Mov. Res.* **2008**, *2* (4), 1–18. <https://doi.org/10.16910/jemr.2.4.5>.
- (43) van Gog, T.; Paas, F.; van Merriënboer, J. J. G. G.; Witte, P. Uncovering the Problem-Solving Process: Cued Retrospective Reporting versus Concurrent and Retrospective Reporting. *J. Exp. Psychol. Appl.* **2005**, *11* (4), 237–244. <https://doi.org/10.1037/1076-898X.11.4.237>.
- (44) 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. *Chem. Educ. Res. Pract.* **2012**, *13* (3), 179–185. <https://doi.org/10.1039/C2RP90001G>.

- (45) National Institute of Advanced Industrial Science and Technology. SDBSWeb <http://sdb.s.db.aist.go.jp> (accessed May 4, 2020).
- (46) Tobii Technology. Eye Tracking for Research <https://www.tobiiipro.com/> (accessed Feb 26, 2020).
- (47) Miles, M. B.; Huberman, A. M. *Qualitative Data Analysis*, 2nd ed.; Sage Publications: Thousand Oaks, 1994.
- (48) 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>.
- (49) McHugh, M. L. Lessons in Biostatistics Interrater Reliability : The Kappa Statistic. *Biochem. Medica* **2012**, *22* (3), 276–282.
- (50) Sheskin, D. J. *Handbook of Parametric and Nonparametric Statistical Procedures*, 5th ed.; CRC Press: Boca Raton, 2011.
- (51) Tomczak, M.; Tomczak, E. The Need to Report Effect Size Estimates Revisited. An Overview of Some Recommended Measures of Effect Size. *Trends Sport Sci.* **2014**, *1* (21), 19–25.
- (52) West, J. M.; Haake, A. R.; Rozanski, E. P.; Karn, K. S. EyePatterns: Software for Identifying Patterns and Similarities Across Fixation Sequences. In *Proceedings of the 2006 Symposium on Eye Tracking Research & Applications*; Association for Computing Machinery: San Diego, CA, 2006; pp 149–154.
- (53) R Core Team. R: A Language and Environment for Statistical Computing. Vienna 2019.
- (54) Gu, Z.; Gu, L.; Eils, R.; Schlesner, M.; Brors, B. Genome Analysis Circlize Implements and Enhances Circular Visualization in R. **2014**, *30* (19), 2811–2812. <https://doi.org/10.1093/bioinformatics/btu393>.
- (55) Harle, M.; Towns, M. A Review of Spatial Ability Literature, Its Connection to Chemistry, and Implications for Instruction. *J. Chem. Educ.* **2011**, *88* (3), 351–360. <https://doi.org/10.1021/ed900003n>.
- (56) Wright, L.; Oliver-Hoyo, M. Student Assumptions and Mental Models Encountered in IR Spectroscopy Instruction. *Chem. Educ. Res. Pract.* **2019**, *16*, 143. <https://doi.org/10.1039/C9RP00113A>.

Chapter 4

Teaching assistants' Topic-Specific Pedagogical Content Knowledge in ^1H NMR Spectroscopy

4.1 Initial remarks

Findings presented in Chapters 2 and 3 serve as empirical evidence of how students develop expertise in ^1H NMR spectral interpretation. These findings provide instructors with significant insight into designing instructional materials and activities that foster relevant expertise. Moreover, experienced instructors also have their own insight into effectively teaching this practice that studies on students' learning are not designed to elicit. For instance, instructors often have knowledge of the order in which content in a curriculum is presented and how this ordering may either facilitate or hinder the learning of new content. Further, research demonstrates that instructors' knowledge for teaching positively correlates with both instruction quality and student learning outcomes.¹⁻³ In order to design the most effective instruction, research that characterizes instructors' knowledge for teaching ^1H NMR spectroscopy is thus also needed.

This chapter corresponds to the first of two studies on teaching ^1H NMR spectroscopy. Specifically, this chapter details an investigation of teaching assistants' (TAs') knowledge for teaching ^1H NMR spectroscopy. This study aimed to characterize TAs' insight into teaching this topic, as well as to determine how these instructors develop such knowledge. Findings will facilitate the design of instructor education materials that cultivate this knowledge for teaching, as well as point toward alternate avenues for developing such knowledge. TAs regularly instruct undergraduates as they learn to interpret ^1H NMR spectra, so findings have the potential to broadly impact instruction. Pedagogical content knowledge (PCK) was used as a framework to investigate knowledge for teaching, as it encompasses knowledge of students' understanding, the

curriculum, instructional strategies, and ways of representing content that best facilitate learning.³ Characterizing PCK in ^1H NMR spectroscopy thus served as a means of characterizing knowledge for teaching this practice.

Findings from this study suggest that not only do the cognitive barriers identified in Chapter 2 constrain undergraduates' thinking, but that overcoming these barriers may serve as the most challenging aspect of developing expertise in ^1H NMR spectral interpretation. When asked to describe what makes this practice most challenging for students, experienced TAs often indicated undergraduates' rigidity regarding chemical shift reference values and signal resolution. Further, when asked to describe what students found difficult about interpreting particular ^1H NMR spectra, they typically pointed toward spectral features with either chemical shift or resolution deviating from rule-based expectations. In response to these difficulties, they reported providing students with explanations of implicit chemical variables resulting in such deviations. Some experienced TAs went on to note that the nature of instruction may promote these rule-based overgeneralizations of learned principles. Conversely, less experienced TAs failed to acknowledge these sources of difficulty while providing instruction that promoted an algorithmic approach to interpreting spectra. These findings suggest that instructor education materials should focus on these cognitive barriers, as well as provide evidence-based instructional strategies that promote analytical thinking. Such a focus would shift less experienced instructors away from instruction that fosters constrained thinking, as well as provide all instructors with knowledge of teaching strategies that encourage students' systematic evaluation of multiple implicit chemical variables.

Results further suggest that knowledge for teaching ^1H NMR spectroscopy is highly dependent upon the content knowledge required for specific problem types and subtopics. For instance, TAs who were able to determine the topicity of protons were also those who had greater knowledge for teaching how to identify diastereotopic protons. Likewise, TAs who were able to interpret complex ^1H NMR spectra were also those who had greater knowledge for teaching how to interpret spectra. Further, their knowledge for teaching still varied significantly across spectra, where one TA may have exhibited exemplary PCK with one spectral interpretation problem involving unexpected chemical

shifts but limited PCK on another involving molecular symmetry. These findings suggest that cultivating content knowledge for a wide range of NMR-specific problem types and subtopics, whether through instructor education materials or independent study, is essential for developing knowledge for teaching. Lastly, knowledge for teaching was dependent upon relative teaching experience in ^1H NMR spectroscopy, or one's amount of teaching experience in ^1H NMR spectroscopy relative to their overall teaching experience. This result suggest that instructors may increase their knowledge for teaching by increasing their relative teaching experience in ^1H NMR spectroscopy, possibly due to the highly complex nature of this content and, in turn, the highly complex nature of the associated PCK.

This chapter first appeared as a research article in *Chemistry Education Research and Practice*, and the original publication and copyright information are provided below. The original publication was modified to adhere to Rackham dissertation formatting requirements, though no additional changes were made. Dr. Jordan Boothe assisted with the design of rubrics and scoring of questionnaire responses. All remaining work, including study design, data collection, additional qualitative data analysis, quantitative data analysis, and writing of the manuscript were completed independently by the author.

Original publication and copyright information

Reproduced from “Connor, M. C.; Shultz, G. V. Teaching Assistants’ Topic-Specific Pedagogical Content Knowledge in ^1H NMR Spectroscopy. *Chem. Educ. Res. Pract.* **2018**, 19 (3), 653–669. <https://doi.org/10.1039/c7rp00204a>” with permission from the Royal Society of Chemistry.

4.2 Abstract

Nuclear magnetic resonance (NMR) spectroscopy is an essential analytical tool in chemistry, and the technique is routinely included as a topic across the undergraduate chemistry curriculum. As a result of NMR's importance, classroom instruction of this topic has received considerable attention in chemistry education research. However, little is known about instructors' knowledge for teaching this topic. In order to better understand

this knowledge, we investigated topic-specific pedagogical content knowledge in ^1H NMR spectroscopy among 20 chemistry teaching assistants at a large Midwestern university in the United States. A questionnaire was developed to provide an inferential measure of content knowledge and topic-specific pedagogical content knowledge in ^1H NMR spectroscopy for participants with a range of teaching experience. Data from the questionnaire were analyzed qualitatively and quantized using a rubric. The quantitative data were transformed using the Rasch model and statistically analyzed. Results from these analyses indicate that pedagogical content knowledge increased with teaching experience in ^1H NMR spectroscopy, suggesting that knowledge for teaching this topic is developed through practice. Additionally, the development of pedagogical content knowledge was found to depend upon content knowledge required for specific NMR sub-topics and problems. This finding suggests that the ultimate “grain-size,” or domain-specificity, of pedagogical content knowledge may extend to the problem level. Results from this study have implications for how instructors may cultivate knowledge for teaching NMR spectroscopy, as well as for how pedagogical content knowledge may be more effectively incorporated into instructor training programs.

4.3 Introduction

Nuclear magnetic resonance (NMR) spectroscopy has developed into a versatile and powerful analytical tool in multiple scientific disciplines. As a result of NMR’s utility, the technique is included as a topic across the undergraduate chemistry curriculum. The American Chemical Society (ACS) makes the importance of this topic apparent in their requirements for approved undergraduate chemistry programs, with an NMR spectrometer listed as the only mandatory instrumentation.⁵ NMR spectroscopy is typically taught in introductory organic chemistry courses, where it is generally considered difficult to both teach and learn because it requires an understanding of complex concepts such as spin-spin coupling and chemical equivalency that are not inherent to other chemistry topics taught at the introductory level. Problem solving in the form of ^1H NMR spectral interpretation is also regularly included in the instruction of this topic given that ^1H NMR spectroscopy has found the widest application among chemists.⁶ The necessity

for problem-solving skills in addition to conceptual understanding further contributes to the difficulty of teaching and learning this topic.⁷

Classroom instruction plays an integral role in students' learning of basic principles of NMR spectroscopy and the practice of ¹H NMR spectral interpretation. Most organic chemistry textbooks provide general guidelines for ¹H NMR spectral interpretation, however these guidelines are insufficient for student learning.⁷ As a result of the importance of this topic, undergraduate classroom instruction of NMR spectroscopy has received significant attention in chemistry education research literature. The vast majority of this attention has focused on the development of NMR spectroscopy laboratory experiments and instructional scaffolding strategies.⁸⁻¹¹ In addition, empirical studies have characterized successful and unsuccessful problem-solving approaches among interpreters of ¹H NMR spectra in order to inform classroom instruction.^{12,13} However, little is known about instructors' knowledge for teaching ¹H NMR spectroscopy despite the abundance of curricular ideas for teaching this topic. An understanding of this knowledge, including how it develops, is needed in order to inform instructor education and in turn improve classroom instruction.

Pedagogical content knowledge (PCK) is an integral component of teaching knowledge. This theoretical construct was conceptualized by Shulman and defined as teachers' knowledge of the most useful and meaningful ways of transforming subject matter in order to make it comprehensible to learners.⁴ The most recent consensus model of PCK defines the construct as knowledge for teaching a particular topic, termed "reflection *on* action," combined with teachers' specific ways of acting on this knowledge, termed "reflection *in* action".¹⁴ A number of empirical studies have attempted to characterize and measure instructors' PCK as a means of understanding their knowledge for teaching.¹⁵ Recognition that this theoretical construct has value for implementation in instructor training programs has further contributed to the increase in the number of studies aimed at its characterization.^{3,16} Empirical studies have demonstrated that instructors' PCK can be improved through training programs and that instructors' PCK positively correlates with both instruction quality and student learning outcomes.¹⁻³ Characterizing instructors' PCK thus serves as a means of understanding knowledge for

teaching a topic and improving instructor education, instruction quality, and student learning outcomes.

This study aimed to investigate how teaching assistants (TAs) develop PCK in ^1H NMR spectroscopy and the nature of their PCK in order to gain insight into instructors' knowledge for teaching this topic. Such insight could then be used to improve instructor education, instruction quality, and student learning outcomes in this topic. More specifically, this study sought to address the following research questions:

1. How do TAs' content knowledge and teaching experience influence their development of PCK in ^1H NMR spectroscopy?
2. What is the nature of TAs' PCK in ^1H NMR spectroscopy?

4.4 Theoretical Framework and Background

4.4.1 Pedagogical Content Knowledge

Shulman originally described PCK as “subject matter knowledge *for teaching*,” or a teacher-specific type of knowledge that combines content and pedagogy and allows for the transformation of subject matter into a form comprehensible for learners.^{4,17} PCK is both personal and canonical; personal PCK develops through reflection on one's own practice, and canonical PCK forms through social means such as professional development.¹⁸ PCK can be examined at the discipline-specific, subject-specific, and topic-specific levels. A teacher with discipline-specific PCK would have an understanding of pedagogical concepts and strategies for teaching in a particular discipline (e.g., science, art, history or mathematics), whereas a teacher with subject-specific PCK would have this understanding for a particular subject within a discipline (e.g., the subject of chemistry in the science discipline). At the topic-specific level, the PCK of a teacher will differ for the particular topic being taught. For example, a chemistry teacher will have a different understanding and approach for how to best teach chemical equilibrium versus particle theory.¹⁹ Moreover, identifying the ultimate “grain-size,” or domain-specificity, of PCK presents a challenge to researchers; Shulman made this uncertainty evident by

questioning in his opening address at the PCK Summit if this domain encompasses a discipline, a field of practice, specific topics, or even certain problems within a discipline.²⁰

The most recent conceptualization of PCK emphasizes the construct's topic-specific nature.¹⁴ Instructors with topic-specific PCK (TS-PCK) are those who can appropriately transform their content knowledge in a topic into a form comprehensible for learners using five components of their teaching knowledge: (1) learners' prior knowledge, including misconceptions; (2) curricular saliency (i.e. the specific content in a curriculum, the sequence in which it is presented, prerequisite knowledge, and the importance of teaching the content); (3) what makes the topic easy or difficult to understand; (4) representations, including powerful examples and analogies; and (5) conceptual teaching strategies.³ A common approach for evaluating PCK has been to identify components that form PCK and then view the construct as an amalgamation of those components.²¹

This study examines PCK at the topic-specific level. The topic-specific nature of TS-PCK aligns with constructivist theories of learning, which describe the growth of instructors' understanding as a construction process specific to the content, students, and context in which the content is taught.²² This alignment indicates that TS-PCK is an appropriate framework for evaluating how instructors develop teaching expertise for a particular topic. The model of TS-PCK also places emphasis on both content knowledge in a topic and the components that transform it, allowing for the development of teaching expertise to be more directly modeled.

4.4.2 Content knowledge

Shulman identified content knowledge (CK), often referred to as subject matter knowledge, as one of several knowledge bases for teaching.⁴ CK is necessary but not independently sufficient for forming PCK, however the exact relationship between CK and PCK remains unclear.^{15,23} CK of participants thus merits evaluation. CK for this study is aligned with Ausubel's theory of meaningful learning, which assumes that in order for meaningful learning to occur, the learner must have relevant prior knowledge, the new

knowledge must relate to this prior knowledge in a meaningful way, and the learner must choose to integrate this new knowledge into existing prior knowledge.²⁴

4.4.3 Research on TS-PCK in chemistry topics

Recent studies on PCK have contributed to an understanding of instructors' TS-PCK for chemistry topics that contain a central problem-solving component, much like NMR spectroscopy. Problem solving is conceptualized in this study as completing a task that contains unfamiliar aspects and extends beyond a routine exercise for the problem solver.²⁵ Successful problem solving thus requires an integration of conceptual understanding rather than merely the application of an algorithm. In their investigation of TS-PCK relating to the mole, Rollnick *et al.* found that teachers often promoted algorithmic approaches to mathematical problem solving rather than conceptual understanding.²⁶ Malcolm also suggested that instructors were quite capable of providing teaching strategies for stoichiometry problems potentially because of a reliance on algorithms rather than developing conceptual understanding relating to this topic.²⁷ It is unclear if this reliance on algorithms while teaching is specific to stoichiometry and the mole or if it extends to other chemistry topics in which problem solving is a central component, namely NMR spectroscopy. Although NMR spectral interpretation is a type of nonmathematical problem solving involving the determination of spatial relationships,¹² instructional strategies that rely on algorithms may also extend to this topic. This would be a problematic teaching strategy among instructors given that a single ^1H NMR spectrum has the potential to provide an abundance of information, yet there is no straightforward algorithm or general procedure for spectral interpretation.⁷

A number of empirical studies on chemistry instructors' PCK have also contributed to a developing understanding of how to cultivate knowledge for teaching chemistry topics. Mavhunga demonstrated that explicitly discussing a topic through the five components of TS-PCK and engaging with concepts in the topic improves teachers' TS-PCK in that topic.²⁸ The ability to use the five components of TS-PCK is then transferrable to other topics, but these components must be accompanied by successful engagement with the concepts of the new topic if TS-PCK in this new topic is to improve. By identifying

the component most accessible to teachers, a possible leverage point can be established in this transfer process.²⁹ This finding aligns with that of Charalambous *et al.*,³⁰ who demonstrated that incorporating a high-leverage practice into a prospective teacher training program results in prospective teachers improving in this practice, given that they engage in active reflection. In an investigation of pre-service chemistry teachers' TS-PCK relating to the particulate nature of matter, knowledge of curricular saliency was identified as the most accessible component to teachers.²⁹ This finding was deemed encouraging given that it provides a potential means of “incrementally building topic-specific professional knowledge across core chemistry and physics topics.” However, in a similar study investigating the TS-PCK of “novice unqualified graduate science teachers” in the particulate nature of matter, learner’s prior knowledge was the most accessible component, whereas curricular saliency was much less accessible.³¹ Learner’s prior knowledge was also identified as the most accessible component in a study of pre-service teachers’ TS-PCK in stoichiometry, whereas curricular saliency was identified as the least accessible component.²⁷ These findings call into question whether the leverage point for transfer is potentially dependent upon the training of the instructor or the specific topic. Additional investigations of PCK at the topic level are needed to address this uncertainty.

4.5 Methods

A mixed methods approach was used to characterize TS-PCK in ^1H NMR spectroscopy among teaching assistants (TAs). As part of this approach, a questionnaire was designed to provide a measure of TAs’ CK and TS-PCK in ^1H NMR spectroscopy. Responses to the TS-PCK component of the questionnaire were quantitatively transformed using a rubric and analyzed using the Rasch model.³² Responses were also qualitatively analyzed to gain a richer understanding of TS-PCK in this topic. Questions were validated using the Rasch model, evaluation by external experts, and cognitive interviews.³³ The questionnaire was accompanied by a survey that assessed TA’s teaching experience, interest, and background information.

4.5.1 Participants

The study was performed at a large Midwestern university and consisted of 20 participants (16 graduate TAs enrolled in a doctoral program, two post-doctoral fellows, and two undergraduate TAs) with a range of teaching experience. Graduate and undergraduate TAs are formally referred to as graduate student instructors and undergraduate instructor assistants, respectively, at the university in which the study took place. TAs were selected as participants for two reasons. Firstly, TAs play a prominent role in undergraduate education at doctoral granting institutions.³⁴ Understanding TAs' knowledge for teaching is therefore critical for improving classroom instruction at large universities. Secondly, TAs commonly teach NMR spectroscopy at the institution in which the study took place. Participants with a range of teaching experience could thus be recruited in order to better understand how knowledge for teaching ¹H NMR spectroscopy develops.

TAs at this institution receive two days of formal training that primarily focuses on laboratory management. They also receive varying amounts of informal guidance, suggesting that participants likely possessed personal rather than canonical PCK.¹⁸ Faculty members with extensive teaching experience in NMR spectroscopy were not included as participants in this study population because they likely develop knowledge for teaching that is distinct from that of TAs. This difference is due to the situated nature of expertise. In line with socio-cultural views of teacher learning, an expert is one who engages fully in social practices specific to the area of expertise;³⁵ because TAs do not engage in the same social practices of teaching as faculty members (e.g., determining learning objectives, selecting content to present throughout a course, etc.), TAs will develop teaching expertise that differs from that of faculty members. While faculty members' knowledge for teaching NMR spectroscopy is of interest to the authors, it is beyond the scope of this study and will be the focus of future work.

NMR spectroscopy is taught in Organic Chemistry I lecture and Organic Chemistry II laboratory at this university. NMR spectroscopy is introduced in a limited fashion in Organic Chemistry I lecture, where only the concept of distinct chemical environments in

^{13}C and ^1H NMR spectroscopy is covered. TAs serve as discussion leaders for the recitation component of this lecture. NMR spectroscopy is taught more comprehensively in Organic Chemistry II laboratory, where spectral interpretation and more complex concepts of spin-spin coupling, chemical shift, and topicity effects are covered. Participants had a range of teaching experience in other chemistry courses. Two additional individuals not included in the 20 participants completed cognitive interviews during initial question piloting. All individuals voluntarily consented to participate in the study and IRB approval was obtained.

4.5.2 Questionnaire Design

A questionnaire was designed to provide an inferential measure of TAs' CK and TS-PCK in ^1H NMR spectroscopy. A questionnaire blueprint similar to that of Jüttner *et al.* was developed to assist with CK and PCK question design (Table 4.1).¹⁵ CK and PCK questions were written by adapting problems from an organic chemistry textbook and consulting with a faculty member who has over ten years of experience teaching organic chemistry.³⁶ Given that organic chemistry curricula do not vary significantly among instructors or institutions,³⁷ PCK questions thus had the potential to elicit insight into knowledge for teaching NMR spectroscopy that may be broadly applicable to instruction of this topic.

A second tier was included in PCK questions that asked TAs to identify whether their experience as a teacher, researcher, and/or student informed their response. Participants were instructed to select as many options as applicable. This second tier was included to determine if participants drew on direct teaching experience and not just reasoning from experience as a student. Responses to the second tier also provided insight into the collective research experience in NMR spectroscopy among the study population. The response frequencies to this second tier are provided in the Supporting Information (Table 4.12). The questionnaire was piloted with two content experts who were not part of the project team using cognitive interviews to investigate whether all the questions were interpreted as intended.³³ Content experts were graduate students with teaching experience in organic chemistry and research experience in organic chemistry

and chemistry education. The content experts then discussed all questions with one author, and revisions were made in accordance with their suggestions in order to improve question clarity. After the initial pilot, two study participants participated in cognitive interviews to evaluate the final version of the questionnaire. These interviews revealed that one final PCK question was not interpreted as intended; this PCK question was omitted from subsequent analysis.

The final version of the questionnaire contained five CK questions and four PCK questions. The complete set of CK and PCK questions are provided in the Supporting Information. CK questions were placed at the beginning of the questionnaire, followed by PCK questions; this was done to prevent TAs from using a proton chemical shift table provided in PCK questions to answer CK questions. Experts reported only slight difficulty in responding to the CK component during question piloting, so a 25-minute time limit was imposed with the final version to increase difficulty and reduce overall time for participation. No time limit was imposed on the PCK component so that TAs could provide responses that completely captured their thinking. Participants took approximately one hour to respond to the PCK component.

The questionnaire blueprint (Table 4.1) identifies the components of CK or PCK that each particular question targeted. CK questions were written to assess TAs' procedural and declarative knowledge of ^1H NMR spectroscopy.³⁸ Procedural CK questions required TAs to elucidate the molecular structure corresponding to a provided ^1H NMR spectrum and molecular formula. ^1H NMR spectra for the CK component were retrieved from the Spectral Database for Organic Compounds.³⁹ Declarative CK questions assessed TAs' knowledge of proton equivalency and spin-spin coupling. PCK questions were written to probe two components of PCK: 'what makes a topic difficult' and 'teaching strategies'.³ ^1H NMR spectra for the PCK component were generated using ChemDraw Professional 16.0.⁴⁰ For PCK questions that involved spectral interpretation, the faculty member consulted during initial question drafting provided perspectives on what spectral features most commonly make interpretation difficult for students. Although TAs and faculty members likely have distinct knowledge for teaching, insight into the *most*

Table 4.1. Questionnaire Blueprint.

Type of CK	CK Question
Procedural	CKQ1, Determine structure from spectrum. CKQ2, Determine structure from spectrum.
Declarative	CKQ3, Which sets of hydrogen atoms are equivalent? CKQ4, Which sets of molecules are distinguishable using ^1H NMR? CKQ5, How many ^1H NMR signals does the molecule produce?
Component of PCK	PCK Question
What makes a topic difficult	PCKQ1, What do students find most difficult about NMR spectral interpretation?
What makes a topic difficult and teaching strategies	PCKQ2, What would a student find difficult about determining a structure from this spectrum, and how would you help them interpret the spectrum? PCKQ3, Did the student correctly elucidate a structure from this spectrum, what (if anything) created difficulty, and how would you help them interpret the spectrum? PCKQ4, Did the student correctly determine equivalent hydrogen atoms, what (if anything) created difficulty, and how would you help them determine the correct answer?

common difficult features allowed for the design of questions that would most likely elicit TAs' PCK. The content experts also indicated that all PCK questions were appropriate for assessing TAs' knowledge for teaching.

Measuring particular components of PCK has been shown to provide a reliable measure of overall PCK,⁴¹ so the questionnaire was designed to probe only two of the five components of PCK. Researchers hold different conceptualizations of the components that contribute to PCK, however there is consensus that knowledge of students' understanding and knowledge of instructional strategies are integral to the construct.⁴² 'What makes a topic difficult' and 'teaching strategies' were selected because they align with this essential knowledge. In addition, the overall quality of PCK depends upon both the quality of individual components and the coherence among components, and knowledge of students' understanding and knowledge of instructional strategies are central in the integration of multiple PCK components.⁴³ By providing participants with questions that best allowed them to integrate multiple components of PCK, the quality of their PCK could be more effectively measured. Additional data for this study included

audiotaped cognitive interviews with two content experts and two study participants, as well as responses to survey questions that characterized TA teaching experience, interest in teaching, and additional background information.

4.5.3 Data analysis

PCK responses were scored on a 0-4 point scale using separate rubrics for each targeted PCK component (Table 4.2). Responses ranged from incorrect (0) to exemplary (4). The rubric used to score ‘what makes a topic difficult’ responses was similar to that of Mavhunga and Park and Oliver.^{21,44} Developing (3) and exemplary (4) responses were those that incorporated either one or two components of PCK, respectively. The total frequency of each PCK component in responses was also determined.

The rubric used to score ‘teaching strategies’ responses was similar to that of Hale *et al.*⁴⁵ Developing (3) responses incorporated either interactive teaching or the use of representations during explanations. “Interactive teaching” on this rubric was consistent

Table 4.2. PCK scoring rubrics.

Score	What makes a topic difficult	Teaching strategies
0 Incorrect	– Provides incorrect explanation	– Provides incorrect explanation
1 Limited	– Identifies difficult aspect – Provides no reasoning	– Provides problem-solving method – Does not relate method to problem
2 Basic	– Identifies difficult aspect – Provides broad and generic reasoning	– Provides problem-solving method – Relates method to problem
3 Developing	– Identifies difficult aspect – Provides reasoning relating to <u>one</u> PCK component: Learner’s prior knowledge Conceptual teaching strategies Representations (examples or models) Curricular saliency	– Uses interactive teaching, e.g., questioning to probe or promote students’ understanding or – Uses illustrations or models during explanations
4 Exemplary	– Identifies difficult aspect – Provides reasoning relating to <u>two or more</u> PCK components	– Recognizes learners’ prior knowledge

with the definition described by Chin.⁴⁶ Two coders discussed and revised operational definitions in the rubric until 96% agreement was reached. A third coder then scored a subset of questionnaire responses (15%) using the finalized rubric, and an acceptable Cohen's kappa (0.736) was achieved.⁴⁷ For PCK items that targeted both PCK components, 'teaching strategies' and 'what makes a topic difficult' scores were averaged. CK responses were scored based on a correct or incorrect basis. Exemplars corresponding to each rubric and PCK question are included in the Supporting Information (Tables 4.5-4.11).

Raw CK and TS-PCK scores were subjected to Rasch analysis using Winsteps software.⁴⁸ The Rasch model places person ability and item difficulty on the same scale in logit units. A logit is defined as the logarithmic transformation of the odds of success.³² The model sets the mean item difficulty to zero logit units, meaning that an item of average difficulty would have a logit unit equal to zero. The unidimensionality of the model provides an inferential measure of a person's overall ability relating to a single latent variable, in this instance, CK or TS-PCK.³² The Rasch model also provides a measure of the validity of the results through reliability estimates and model fit statistics.

4.6 Results

A questionnaire was developed to provide an inferential measure of teaching assistants' CK and TS-PCK in ¹H NMR spectroscopy. The results were validated through several processes: face validity of the questionnaire was established through consultation with external experts and construct validity through cognitive interviews and analysis using the Rasch model.^{49,50} The questionnaire was administered to 20 TAs, and their responses were evaluated both quantitatively and qualitatively.

4.6.1 Questionnaire development and Rasch model validity

The questionnaire for measuring CK and TS-PCK was validated in part using the Rasch model. Item reliability, a measure of the extent to which items represent a range of difficulty relating to a single variable, and person reliability, a measure of whether the questionnaire appropriately discriminates across the ability range of participants, were

used as reliability indices.³² Rasch measures produced acceptable person and item reliability indices for the PCK component, as well as an acceptable item reliability for the CK component. For the PCK component of the questionnaire, item reliability was 0.96 and person reliability was 0.78. In the case of the CK component, item reliability was 0.81 and person reliability was 0.47. Acceptable item and person reliability indices for the CK and PCK components were similar to those found by Mavhunga and Rollnick and Jüttner *et al.*^{3,15} Person reliability for the CK component was low but similar to that found by Hale *et al.*⁴⁵ This low reliability can be explained by the high CK of TAs. CK questions were derived from an undergraduate organic chemistry textbook,³⁶ so most TAs were able to perform well on these questions and, in effect, decrease the questionnaire's capability to discriminate among CK levels.

Fit statistics were used to assess the questionnaire's level of productive measurement. For all CK and PCK items, $MNSQ < 1.5$ and/or $t \leq |2|$, indicating that items were productive for measurement.⁴⁸ For all persons on the CK component, $MNSQ < 1.5$ and/or $t \leq |2|$, confirming that all participants fit the Rasch model. For 19 out of 20 person measures on the PCK component, $MNSQ < 1.5$ and/or $t \leq |2|$; this was acceptable given that 5% of people are expected to misfit the model by chance.⁵¹ These acceptable fit statistics further validate that the questionnaire reliably measures CK and TS-PCK.

4.6.2 Relationship between CK and PCK person measures

The relative placement of person ability and item difficulty determined through Rasch analysis can be depicted using an item-person map (Figures 4.1-4.2). A person's location on this map indicates the person's ability to correctly respond to questions of a given difficulty. For example, if a person and item have the same logit measure, then the person has a 50% chance of answering a similar item correctly. This person would have greater than a 50% chance of sufficiently answering items of lower difficulty (i.e., lower logit value) and lower than a 50% chance of correctly answering items of higher difficulty (i.e., higher logit value). CK and PCK person measures generated by the Rasch model may provide insight into how TS-PCK in ¹H NMR spectroscopy develops. According to the item-person

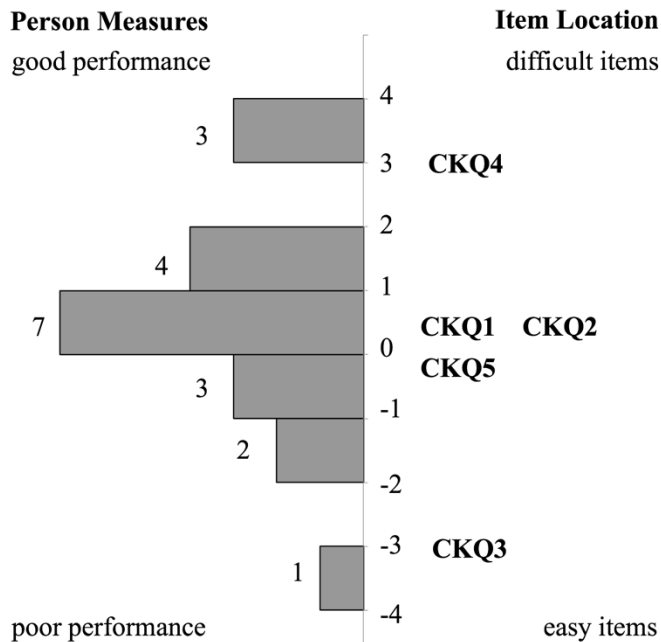


Figure 4.1. Item-person map of TA CK in ¹H NMR spectroscopy. Questions CKQ1 and CKQ2 assessed procedural knowledge, and questions CKQ3, CKQ4, and CKQ5 assessed declarative knowledge. Numbers to the right of the axis correspond to logit measures, and numbers to the left of each bar correspond to the number of participants falling within a given logit range.

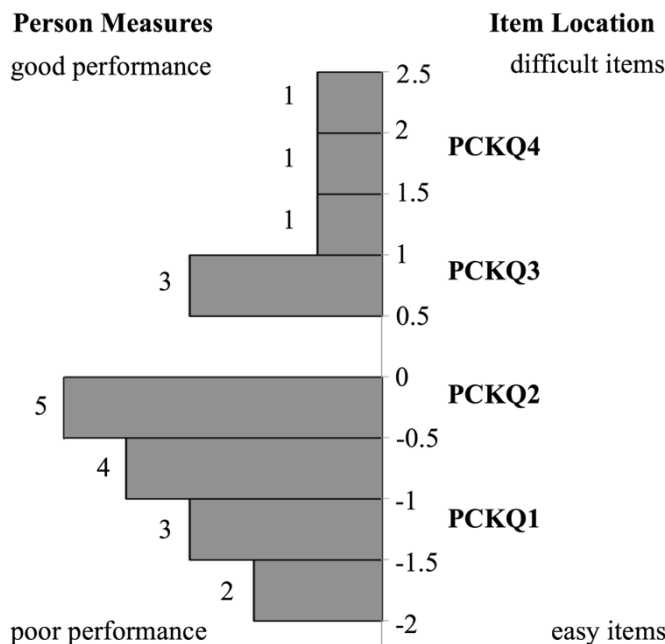


Figure 4.2. Item-person map of TA PCK in ¹H NMR spectroscopy. Question PCKQ1 targeted the ‘what makes a topic difficult’ PCK component, and questions PCKQ2, PCKQ3, and PCKQ4 targeted ‘what makes a topic difficult’ and ‘teaching strategies’ components. Numbers to the right of the axis correspond to logit measures, and numbers to the left of each bar correspond to the number of participants falling within a given logit range.

map for the CK component of the questionnaire (Figure 4.1), TAs had a relatively high

CK of ^1H NMR spectroscopy, with the mean person measure equal to 0.66 logits. TAs had comparatively low TS-PCK in ^1H NMR spectroscopy, with the mean person measure equal to -0.21 logits. A small number of TAs were able to score well above the -0.21 logit average measure, suggesting that these TAs may represent the upper anchor of TA teaching expertise for this topic. The relatively low level of TS-PCK compared to CK aligns with the notion that CK is necessary but not sufficient for the development of PCK.³¹ This finding was also similar to that of Hale *et al.*⁴⁵

CK and PCK person measures are depicted on a scatterplot in Figure 4.3. A strong and significant positive correlation was found between CK and PCK (Table 4.3), which also aligns with the notion that CK is necessary for the development of PCK. This result is consistent with those of Hale and Jüttner and further validates that the questionnaire measured what was conceptualized as TS-PCK.^{15,45} Data points in the lower right quadrant correspond to TAs with high CK scores but low PCK scores; this observation supports the general agreement that CK is necessary but not sufficient for the development of PCK. One data point unexpectedly populates the upper left quadrant corresponding to relatively low CK and

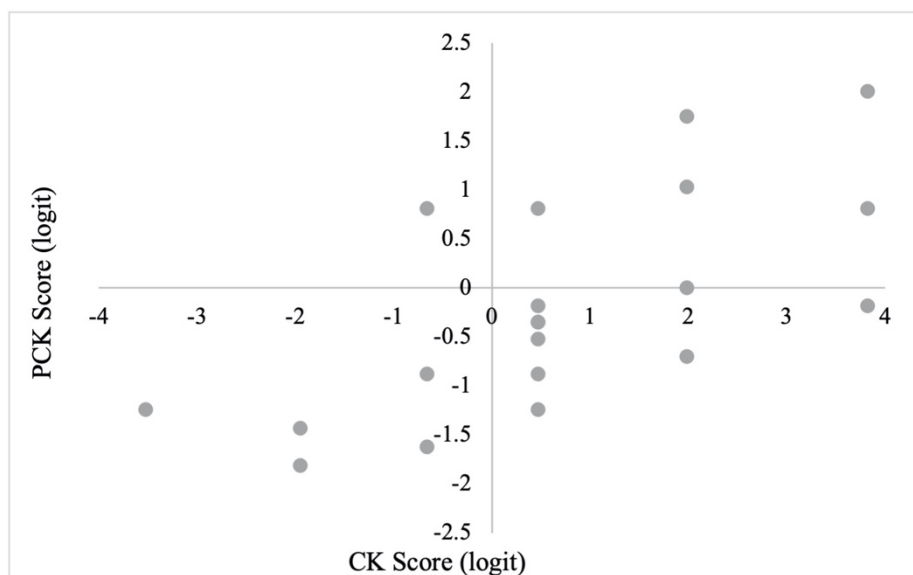


Figure 4.3. Scatterplot of participants' CK and PCK measures ($r=0.670$, $p<0.01$). Higher CK logit measures indicate a greater ability to correctly respond to CK questions, and higher PCK logit measures indicate a greater ability to adequately respond to PCK questions.

Table 4.3. Pearson correlations between CK, PCK, and TA characteristics. * indicates p (two-tailed) < 0.05, ** indicates p (two-tailed) < 0.01.

Indices	Content knowledge	Pedagogical content knowledge
Content knowledge	1	0.670**
Pedagogical content knowledge	0.670**	1
Number of chemistry courses taught	-0.153	-0.308
Number of terms teaching organic II lab	-0.075	-0.073
Relative teaching experience (Number of terms teaching Organic II lab divided by number of chemistry courses taught)	0.278	0.490*
Terms since last teaching organic II lab	-0.397	-0.432
Teaching interest	0.367	0.424
Organic sub-discipline	0.148	-0.217

high PCK. This could be due to this participant not performing well in timed test-taking situations.

4.6.3 Dependence of PCK on CK

The Rasch model provided an ordering of CK and PCK questions from least to most difficult that may also provide insight into how TAs develop TS-PCK in ^1H NMR spectroscopy. This ordering, termed a difficulty hierarchy,⁵² is depicted in each respective item-person map (Figures 4.1-4.2). A comparison of CK and PCK difficulty hierarchies reveals that CK questions that targeted procedural knowledge align in difficulty with PCK questions that targeted procedural knowledge. For example, CKQ1 and CKQ2 (Figure 4.1) and PCKQ2 and PCKQ3 (Figure 4.2) were all of intermediate difficulty, and all assessed procedural knowledge involved in elucidating a structure from a spectrum. Additionally, CK questions that targeted declarative knowledge also align in difficulty with PCK questions that targeted declarative knowledge. CKQ5 and PCKQ4 were of greatest difficulty, and both assessed declarative knowledge relating to the determination of topicity in complex scenarios. The similarity in CK and PCK difficulty hierarchies suggests that TS-PCK depends on CK in that topic, which is consistent with PCK theory, but also that the development of PCK may be affected specifically by the declarative or procedural CK required for a particular sub-topic or problem. This alignment between CK and PCK

questions assessing similar procedural and declarative knowledge contrasts with the findings of Jüttner *et al.*,¹⁵ who found no alignment between CK and PCK questions that targeted declarative and procedural knowledge on an instrument designed to assess PCK in four biology topics. This alignment between declarative and procedural questions may therefore only be observable when examining PCK at the topic-specific level for a *single* topic.

4.6.4 Relationship between PCK and teaching experience

A significant positive correlation was found between PCK and relative teaching experience (Table 4.3). We define relative teaching experience as the number of times a TA taught Organic Chemistry II laboratory relative to the total number of chemistry courses the TA taught. For example, if a TA taught Organic Chemistry II laboratory once and General Chemistry laboratory twice, relative teaching experience would equal 0.33. As previously noted, NMR spectroscopy is only taught in Organic Chemistry I lecture and Organic Chemistry II laboratory at the university in which the study took place, and NMR spectroscopy is introduced in a very limited manner in Organic Chemistry I lecture, where only the concept of distinct chemical environments in ^{13}C and ^1H NMR spectroscopy is covered. NMR spectroscopy is taught more comprehensively in Organic Chemistry II laboratory, and this is the only course in which TAs would teach spectral interpretation and the inherent concepts of spin-spin coupling and chemical shift that must be incorporated into problem solving. The relationship between PCK and relative teaching experience suggests that TAs develop TS-PCK in ^1H NMR spectroscopy by teaching chemistry courses in which NMR is explicitly taught. This further supports the consensus that PCK is topic-specific in nature and further validates that the questionnaire measures TS-PCK.¹⁴

No significant relationship was found between PCK and total chemistry teaching experience or teaching experience in Organic Chemistry II laboratory. This finding was not consistent with that of Hale *et al.* who found that overall teaching experience was correlated for TAs with TS-PCK in thin layer chromatography.⁴⁵ However, this result may be attributed to the sample size of the study; within this small sample size were

participants who had previously taught Organic Chemistry II laboratory but had also taught many other courses in which ^1H NMR spectroscopy is not explicitly taught. In some cases, TAs had experience teaching Organic Chemistry II laboratory, but a substantial time lapse had occurred between that experience and participation in the study. A potential inability among these participants to recall knowledge for teaching this particular topic may have contributed to the insignificant relationship between PCK and overall teaching experience. Further, chromatography is a general approach that TA's may have encountered in other courses and contexts, and such experience may be more easily translated to specific chromatographic techniques such as TLC. Chromatography is also conceptually more tractable than NMR. Finally, we also found no significant relationship between PCK and teaching interest, terms since last teaching, or research sub-discipline.

4.6.5 Identifying a leverage point for transfer of TS-PCK

Analysis of responses to PCK questions may provide insight into the nature of TS-PCK in ^1H NMR spectroscopy. TAs that demonstrated either “developing” or “exemplary” PCK on ‘what makes a topic difficult’ questions primarily did so by incorporating an understanding of learners’ prior knowledge into their responses. The frequencies of PCK components incorporated in “developing” or “exemplary” ‘what makes a topic difficult’ responses are depicted in Table 4.4. An understanding of curricular saliency, representations, and conceptual teaching strategies were present in responses to ‘what makes a topic difficult’ questions, however their frequencies were much lower (Table 4.4). The prevalence of this incorporation suggests that TAs found learners’ prior knowledge to be a relatively accessible component of PCK as opposed to other components. Rollnick and Mavhunga suggested that by identifying the component most accessible to instructors, a possible leverage point can be established in order to transfer the ability to

Table 4.4. Frequencies of PCK components incorporated into “developing” or “exemplary” ‘what makes a topic difficult’ responses

PCK Component	Learners’ prior knowledge	Curricular saliency	Representations	Conceptual teaching strategies
Frequency	24	4	4	1

use the five components of PCK to other topics.²⁹ The accessibility of learners' prior knowledge contrasts with the findings of Rollnick and Mavhunga, who identified curricular saliency as being most accessible to pre-service chemistry teachers when investigating TS-PCK in the particulate nature of matter. However, it is consistent with those of Malcolm (2015), who identified learners' prior knowledge as most accessible to pre-service teachers when investigating TS-PCK in stoichiometry. This alignment suggests that the potential anchoring component may depend upon the particular topic. Given that both stoichiometry and NMR include elements of problem solving, this difference may be attributed to the particular nature of the topics. Another possibility may be the nature of the study population. Pitjeng identified learners' prior knowledge as the most accessible component among "novice unqualified graduate science teachers" developing TS-PCK in the particulate nature of matter.³¹ This agreement suggests that learners' prior knowledge may be the most accessible component of PCK among instructors without extensive professional development in teaching, implying that the anchoring component may also depend on instructors' training in addition to the particular topic.

4.6.6 Teaching strategies of TAs

For 'teaching strategies' PCK questions, TAs most commonly described interactive teaching strategies involving questioning to probe or promote students' understanding and the use of drawings or models during explanations.⁴⁶ Of the 60 'teaching strategies' responses, 23 employed interactive techniques or the use of representations during explanations. The frequencies of responses that described a teaching strategy related to the problem (11), a strategy unrelated to the problem (9), or a strategy based on incorrect CK (11) were much lower in comparison. The use of algorithms did not dominate TAs' teaching strategies, with only seven out of 60 responses providing a simplified step-by-step problem-solving approach. This result contrasts with findings of other investigations on TS-PCK in chemistry topics with a central problem-solving component.^{26,27}

4.6.7 Problem-specificity of PCK

Analysis of PCKQ2 and PCKQ3 may provide insight into both the “grain-size” of PCK and the nature of TAs’ TS-PCK in ^1H NMR spectroscopy. These questions were of similar problem type (i.e., determine ‘what makes a topic difficult’ and ‘teaching strategies’ for a problem where a student must determine a structure from a spectrum), yet they had different difficulty levels as depicted in the PCK item-person map (Figure 4.2). For the less difficult item (PCKQ2), TAs readily identified a difficult feature and incorporated an understanding of learners’ prior knowledge into their responses. Thirteen out of 20 TAs identified a difficult feature on PCKQ2. TAs that identified a difficult feature mostly attributed difficulty to students not understanding that peaks in different chemical environments may potentially overlap. These observations align with what an experienced instructor of the course described as student difficulties specific to this problem during initial PCK item drafting. Many TAs also noted students’ rigidity and heavy reliance on the proton chemical shift table as a heuristic. Participants who demonstrated an “exemplary” response typically recognized this misconception and also incorporated another component of PCK. In the following response, this second component was an understanding of curricular saliency, in particular knowledge of the content presented in the curriculum:

*“A large difficulty for the student lies in the integration of the aromatic hydrogens. **We teach them that they are groups of different hydrogens depending on their aromatic position**; however, it is now slightly confusing because those are now all grouped in one peak. An additional difficulty may lie in the reasoning for the downshift of the peak labeled ‘2, quartet’ as it does not exactly align with the NMR [chemical shift] chart.” –Participant One*

The majority of participants also demonstrated either “developing” or “exemplary” ‘teaching strategies’ responses on PCKQ2. Seven out of 20 TAs demonstrated a “developing” transformation of CK to PCK by integrating interactive teaching into their teaching strategy, and four out of 20 TAs demonstrated an “exemplary” transformation of CK to PCK by incorporating an understanding of learners’ prior knowledge into their ‘teaching strategy’ response:

*“The peak at 2.75 [ppm] would most likely be causing issues. The student probably sees a peak at 7.2 [ppm] and looks at the table and assigns it as an aromatic proton. The student looks at the 1.2 ppm peak and assigns it as alkyl. But the peak at 2.75 [ppm] doesn’t match anything on the table. They would tell me it is in the range for an alcohol peak or an amine, but the structure doesn’t contain O or N. **I would first ask them to describe what kinds of protons account for the chemical shifts in the spectra. They would give me the answer [above], then I would talk them through how inductive effects can affect chemical shift.** Similar to how an electronegative atom de-shields protons adjacent to it, the aromatic ring can be thought of as electronegative, and pulls electrons away from adjacent protons. This means that even though the benzylic protons are alkyl, they can appear at a higher ppm than normal.” –Participant Two*

TAs with low PCK scores often advocated for use of the proton chemical shift table as a heuristic and routinely suggested an algorithmic approach to problem solving. These TAs did not demonstrate knowledge of students’ rigidity while problem solving:

“1. Calculate degree of unsaturation 2. Using chemical shift table + integration, identify likely functional groups containing [hydrogen atoms]. 3. Begin drawing possible structures, keep formula, symmetry (# of signals), + degree of unsaturation in mind. 4. Identify correct structure from your possibilities, making sure splitting agrees with assignment.”- Participant Three

PCKQ3 was more difficult for TAs. Of the 20 responses to PCKQ3, only five identified a difficult feature. This problem involved elucidating the structure of 2,4-hexanedione, the methylene group of which has a chemical shift much further downfield than a chemical shift table would indicate. This problem-specific difficult feature was also identified by an experienced instructor of the course during PCK item drafting. No TA identified this particular feature, and TAs that did identify a feature attributed student difficulty to the student not understanding splitting patterns. Cognitive interviews suggested that among TAs who did not identify a challenging aspect, a difficult feature was not apparent. Failure to identify a feature was therefore not due to a misreading of the question or fatigue. Only one TA incorporated an understanding of learners’ prior knowledge into a teaching strategy. Notably, this was one of the four TAs who did so on PCKQ2:

“The student knew that the structure contained an ethyl group, but they should have realized that the structure they drew cannot account for the two singlets. In this case, I would suggest a guess and check approach. Just starting drawing structures, keeping the ethyl group but moving the other carbons around to get different connectivity. Keep

changing the structure until all the NMR features are accounted for. I find that to be generally a good approach for me, as well as the students I have taught.” – Participant Two

The number of TAs who incorporated interactive teaching into their teaching strategy remained relatively unchanged from PCKQ2, with 11 out of 20 TAs able to make this incorporation on PCKQ3 compared to seven out of 20 on PCKQ2. Out of the four TAs who were able to incorporate learners' prior knowledge into their teaching strategy on PCKQ2 but not PCKQ3, three still described interactive teaching strategies on PCKQ3. This relatively unchanged number of TAs using interactive teaching strategies, combined with the large decrease in the number of TAs able to identify a difficult feature on PCKQ3, suggests that TAs' use of interactive teaching strategies is not strongly influenced by their understanding of what makes a problem difficult. These results also suggest that as TAs struggle to identify what makes a particular problem difficult, they also struggle to incorporate learners' prior knowledge into their teaching strategy for that problem. The difference in ability to 1) recognize what would make different problems of an identical type (i.e., determine structure from spectrum) difficult for a student and 2) provide exemplary teaching strategies for these problems implies that TAs' PCK may be specific to certain problems in addition to certain topics.

4.7 Discussion

A questionnaire was developed to provide an inferential measure of CK and TS-PCK in ^1H NMR spectroscopy. The questionnaire was administered to TAs with a range of teaching experience, and their responses provided a means of understanding how TAs develop TS-PCK in ^1H NMR spectroscopy, the nature of this TS-PCK, and the nature of PCK more broadly. Multiple findings emerged from the analysis of questionnaire responses. First, CK in NMR spectroscopy was significantly correlated with PCK in this topic, further supporting the notion that CK is necessary for the development of PCK. In addition, CK and PCK questions had similar difficulty hierarchies, with the most difficult question on both components targeting similar declarative knowledge and questions of intermediate difficulty on both components targeting similar procedural knowledge. This

alignment suggests that the development of PCK may also depend upon the declarative or procedural CK required for a specific sub-topic or problem.

Second, TAs' TS-PCK in ^1H NMR spectroscopy was significantly correlated with relative teaching experience. This correlation suggests that knowledge for teaching ^1H NMR spectroscopy develops through practice teaching this topic, and that teaching additional courses that do not involve NMR spectroscopy may not contribute to the formation of TS-PCK in this topic. This correlation further supports the general agreement that PCK is topic-specific in nature. Practice teaching this topic may provide TAs with greater opportunity to engage in pedagogical reasoning,⁴ a cyclic process that involves comprehension and transformation of the subject matter, instruction, evaluation, and reflection. Engagement in this process through practice teaching may then serve as the means by which TAs develop TS-PCK.

A third finding is that while TAs' TS-PCK in ^1H NMR spectroscopy was relatively low, they were not reliant on algorithmic approaches to problem solving and routinely described interactive teaching strategies. This result suggests that algorithmic approaches to problem solving may not be a significant component of practices for teaching NMR spectroscopy. Additionally, TAs most commonly demonstrated an understanding of learners' prior knowledge relative to other components of PCK in their responses to 'what makes a topic difficult' questions. Given that TAs found this component of PCK relatively accessible when teaching NMR spectroscopy, this component may be a possible leverage point that allows TAs to transfer their ability to use PCK components to other topics. The accessibility of learners' prior knowledge also suggests that discussing NMR spectroscopy through this component of PCK may serve as a way to initially cultivate knowledge for teaching this topic.

Lastly, TAs demonstrated a considerable difference in their ability to provide sufficient responses for PCKQ2 and PCKQ3, both questions that assessed procedural knowledge and involved elucidating a structure from a spectrum. The difference in ability to 1) recognize what would make different problems of an identical type difficult for

students and 2) provide exemplary teaching strategies for each problem suggests that PCK may be specific to certain problems in addition to certain topics.

4.8 Conclusions

The questionnaire reported here provided an inferential measure of CK and PCK in ^1H NMR spectroscopy, and it provided a means of investigating knowledge for teaching ^1H NMR spectroscopy. Several significant findings emerged from this investigation. TAs with greater relative teaching experience in NMR spectroscopy were found to have higher levels of PCK in this topic, reinforcing the topic-specific nature of PCK. Results from the quantitative and qualitative analysis of responses to the questionnaire also provide evidence that the development of PCK is dependent upon CK required for a specific sub-topic or problem and that TAs develop PCK for specific types of problems in addition to specific topics. These results suggest that the domain-specificity of PCK may extend to the problem level.

4.9 Limitations

PCK is difficult to measure because it exists both internally in the mind of a teacher and as an external construct. 'Materials-based items,' or questions that incorporate materials used in the classroom (e.g., worksheets completed by students or lesson plans), have been shown to reliably assess PCK.⁵³ The majority of short-answer PCK questions included in this questionnaire are a type of 'material-based item' and are thus appropriate for evaluating TAs' PCK. However, this evaluation is still limited given that short answer responses only reflect thinking or teaching approaches that TAs choose to report and not necessarily what they might think or do in the classroom. To add to this limitation, scorers were required to make inferences about the intended meaning of TAs' sometimes ambiguous responses. In these instances, responses were regularly given the higher potential score. Additionally, the relatively small sample size of the study was also a limitation that may have resulted in the insignificant correlations between PCK and total chemistry teaching experience or teaching experience in Organic Chemistry II laboratory. However, as indicated by the person reliability index for the PCK component of the

questionnaire, participants demonstrated a range of teaching ability; the study population was therefore representative in that regard.

4.10 Implications

A number of studies have attempted to characterize PCK in order to incorporate the construct into instructor training programs and in turn improve instructor education and student learning outcomes.^{3,54} This study provided additional insight into the nature of PCK that may facilitate such an incorporation; our findings suggest that PCK may extend to the problem or sub-topic level, implying that instructors may first develop knowledge for teaching problems or sub-topics and this knowledge may then contribute to their knowledge for teaching a topic. In order to facilitate instructors' development of knowledge for teaching, training programs may therefore need to initially focus on the development of knowledge for teaching problems or sub-topics before aiming to cultivate knowledge for teaching a topic. In addition, the accessibility of learners' prior knowledge among participants suggests that this component of PCK may serve as a leverage point to begin building knowledge for teaching among either TAs or instructors of ^1H NMR spectroscopy. Lastly, our results suggest that knowledge for teaching ^1H NMR spectroscopy develops through practice and that teaching additional courses in which this topic is not included may fail to contribute to the development of TS-PCK. This finding implies that TAs should regularly teach the same course whenever possible in order to improve classroom instruction of ^1H NMR spectroscopy and other topics, though this may pose a challenge to chemistry departments at doctoral granting institutions.

4.11 Supporting information

4.11.1 Exemplars and second tier response frequencies

Table 4.5. PCKQ1 exemplars: What aspects of NMR spectral interpretation are most challenging for introductory organic chemistry students?

What makes a topic difficult score		Participant response
1 Limited	<ul style="list-style-type: none">– Identifies difficult aspect– Provides no reasoning	<p>“Stereochemistry – especially equivalent hydrogens on rotatable bonds versus non-rotatable bonds.”</p> <p>“Gathering all the information necessary from what is given. What I mean by this is getting students to realize the spectrum not only gives them information but also the molecular formula in terms of conjugation. This also goes into looking at the number of [hydrogen atoms], the splitting, as well as the shift.”</p>
2 Basic	<ul style="list-style-type: none">– Identifies difficult aspect– Provides broad and generic reasoning	<p>“In my experience, most students had a hard time interpreting results/spectral data that didn’t follow exactly the parameters they were taught – many relied heavily on the spectral tables in their notebooks, using it as a crutch. They could not fathom that a methyl group might be higher than 1.6 ppm if it was between an oxygen [atom] and a double bond.”</p>
3 Developing	<ul style="list-style-type: none">– Identifies difficult aspect– Provides reasoning relating to <u>one</u> PCK component:<ul style="list-style-type: none">Learner’s prior knowledgeConceptual teaching strategiesRepresentations (examples or models)Curricular saliency	<p>“The most challenging aspects for introductory organic chemistry students, I have found have been the concept of chemical shifts and how it is additive and not A=B (look at chart and find exact answer) every time... Students struggle to grasp that they may have to apply what they know about electronegativity and hybridization to figuring out how chemical shifts can be altered/shifted differently than the values given to them in the chart. Another aspect is the concept of [stereochemistry]/topicity which is taught briefly but not very emphasized. Students often have trouble visualizing in 3D which leads to confusion about how [stereochemistry] can affect equivalence of [hydrogen atoms].”</p>
4 Exemplary	<ul style="list-style-type: none">– Identifies difficult aspect– Provides reasoning relating to <u>two or more</u> PCK components	

Table 4.6. PCKQ2a exemplars: A student brings the ^1H NMR spectrum and chemical shift table depicted below to your office hours. The student says that he or she is having trouble figuring out the corresponding molecular structure. Identify and describe particular features (if any) in the ^1H NMR spectrum that may be creating difficulty for the student.

What makes a topic difficult score		Participant response
1 Limited	<ul style="list-style-type: none"> – Identifies difficult aspect – Provides no reasoning 	“The broad, downfield multiplet compared to the distinct quartet, triplet.”
2 Basic	<ul style="list-style-type: none"> – Identifies difficult aspect – Provides broad and generic reasoning 	“I think the initial challenge is that there are only three different proton signals even though there are 10 [hydrogen atoms].”
3 Developing	<ul style="list-style-type: none"> – Identifies difficult aspect – Provides reasoning relating to <u>one</u> PCK component: <ul style="list-style-type: none"> Learner’s prior knowledge Conceptual teaching strategies Representations (examples or models) Curricular saliency 	“A large difficulty for the student lies in the integration of the aromatic hydrogens. We teach them that they are groups of different hydrogens depending on their aromatic position; however, it is now slightly confusing because those are now all grouped in 1 peak.”
4 Exemplary	<ul style="list-style-type: none"> – Identifies difficult aspect – Provides reasoning relating to <u>two or more</u> PCK components 	No examples

Table 4.7. PCKQ2b exemplars: As though you are talking with the student, explain how he or she should go about interpreting the spectrum below.

Teaching strategies score		Participant response
1 Limited	<ul style="list-style-type: none"> – Provides problem-solving method – Does not relate method to problem 	<p>“1. Calculate degree of unsaturation 2. Using chemical shift table + integration, identify likely functional groups containing [hydrogen atoms]. 3. Begin drawing possible structures, keep formula, symmetry (# of signals), + degree of unsaturation in mind. 4. Identify correct structure from your possibilities, making sure splitting agrees with assignment.”</p>
2 Basic	<ul style="list-style-type: none"> – Provides problem-solving method – Relates method to problem 	<p>“First we should be able to see we have [a] benzene ring in our structure (5 proton[s], shift ~7.3 [ppm]). Then we see the other peaks have two [hydrogen atoms] and three [hydrogen atoms]. If we see the shift of methyl attached to benzene, it’s ~2.3 [ppm] and methyl attached to alkyl it’s ~ 0.9 [ppm]. Therefore we can assume that 2 [hydrogen atoms are] attached to carbon next to benzene, and 3 [hydrogen atoms are] attached to carbon next to alkyl. The quartet makes sense for [the two hydrogen atoms] and triplet makes sense for [the three hydrogen atoms] in this structure.”</p>
3 Developing	<ul style="list-style-type: none"> – Uses interactive teaching, e.g., questioning to probe or promote students’ understanding <li style="text-align: center;">or – Uses illustrations or models during explanations 	<p>“I would start by telling them to figure out what types of protons may be present (i.e. aromatic, alkane, alkene, etc.) I would then ask what the triplet, quartet, and multiplets mean in terms of what other protons are around each set of protons. From there, they would hopefully connect that the t and q are next to each other + the multiplet isn’t interacting w/ these problems. After that, they could piece together there is an aromatic ring w/ an alkane chain off it.”</p>
4 Exemplary	<ul style="list-style-type: none"> – Recognizes learners’ prior knowledge 	<p>“I’d tell the student to first identify the degree of unsaturation. This will suggest a benzene ring. Then identify the diagnostic ethyl. They should then arrive at ethyl benzene. We could then try to rationalize the ‘single’ aromatic peak. Since they observe that it is a multiplet, then they can rationalize that it is clearly not a single peak, but likely multiple peaks that come out very near one another w/ splitting. This peak is likely what they expect, a combination of a doublet: 2 x triplet. So why are they on top of one another? I would then explain that the electron donorability of an ethyl is comparable to a [hydrogen atom]. Therefore, the [hydrogen] atoms on the ring are very minimally effected by the electronic change, and therefore their peaks don’t shift appreciably. I may then find an example spectrum of maybe aniline or something and show that a greater electronic change will make the resonance move apart.”</p>

Table 4.8. PCKQ3a exemplars: You are grading problem #1 on a student’s quiz. For problem #1, was this student able to provide the correct answer? What (if any) features contributed to difficulty?

What makes a topic difficult score		Participant response
1 Limited	<ul style="list-style-type: none"> – Identifies difficult aspect – Provides no reasoning 	“It is not the correct structure. The difficulty lies within the integration and splitting patterns.”
2 Basic	<ul style="list-style-type: none"> – Identifies difficult aspect – Provides broad and generic reasoning 	“I would expect that the presence of 2 singlets would be confusing because that adds to the complexity of the molecule + making sure there are protons not near these problems.”
3 Developing	<ul style="list-style-type: none"> – Identifies difficult aspect – Provides reasoning relating to <u>one</u> PCK component: <ul style="list-style-type: none"> Learner’s prior knowledge Conceptual teaching strategies Representations (examples or models) Curricular saliency 	“This student did very well but got the answer wrong because they did not interpret the splitting pattern correctly. They understood that there are 2 degrees of unsaturation and two methyl [groups] but did not understand they the ketones must be placed to generate the two singlet peaks.”
4 Exemplary	<ul style="list-style-type: none"> – Identifies difficult aspect – Provides reasoning relating to <u>two or more</u> PCK components 	“The student most likely had trouble recognizing their symmetry mistake because of the...ketone groups pointing different ways. They also probably have trouble visualizing where hydrogen [atoms] go so I would encourage them to allow them in. They totally missed the singlets so it is possible they do not understand what coupling means so that should also be explained.”

Table 4.9. PCKQ3b exemplars: If incorrect, how would you help this student correctly interpret this spectrum?

Teaching strategies score	Participant response
1 Limited – Provides problem-solving method – Does not relate method to problem	“To help, I would help them figure out which protons are interacting with each other. I would then have them figure out what the ppm shift means for the amount of deshielding present for each proton. From there, I would help them piece the molecule together based on these observations.”
2 Basic – Provides problem-solving method – Relates method to problem	“I would tell them to adjust the placement of the oxygen (carbonyl) to achieve different variants of splitting patterns and observe if the ppm ranges make sense based on the functional groups provided.”
3 Developing – Uses interactive teaching, e.g., questioning to probe or promote students’ understanding or – Uses illustrations or models during explanations	“Since they are so close to the correct answer, I would ask them how they could modify this molecule to generate singlet peaks. By having them identify the protons that cause the triplet and quartet peaks, they could hopefully see that half the compound is correct. By guess and check they should be able to figure out where to position the ketone correctly.”
4 Exemplary – Recognizes learners’ prior knowledge	“The student knew that the structure contained an ethyl group, but they should have realized that the structure they drew cannot account for the two singlets. In this case, I would suggest a guess and check approach. Just start drawing structures, keeping the ethyl group but moving the other carbons around to get different connectivity. Keep changing the structure until all the NMR features are accounted for. I find that to be generally a good approach for me, as well as the students I have taught.”

Table 4.10. PCKQ4a exemplars: You are now grading problem #2 on a student’s quiz. For problem #2, was this student able to provide the correct answer? What (if any) features contributed to difficulty?

What makes a topic difficult score	Participant response
1 Limited – Identifies difficult aspect – Provides no reasoning	“Bromine creates diastereotopic protons.”
2 Basic – Identifies difficult aspect – Provides broad and generic reasoning	“They are right for the most part except for the starred [hydrogen atoms]. These are diastereotopic and will show up differently on an NMR spectrum.”
3 Developing – Identifies difficult aspect – Provides reasoning relating to <u>one</u> PCK component: Learner’s prior knowledge Conceptual teaching strategies Representations (examples or models) Curricular saliency	“The greatest difficulty here is the appearance of pseudo-symmetrical qualities in the molecule. The student does not fully understand that the Bromine disrupts the symmetry as it generates a stereocenter.”
4 Exemplary – Identifies difficult aspect – Provides reasoning relating to <u>two or more</u> PCK components	No examples

Table 4.11. PCKQ4b exemplars: If incorrect, how would you help this student determine the correct answer?

Teaching strategies score		Participant response
1 Limited	<ul style="list-style-type: none"> – Provides problem-solving method – Does not relate method to problem 	“I would encourage students to find centers of chirality first (in these types of questions) and then determine whether this might affect the chemical environments of protons that are near (1 carbon away from) the stereocenter.”
2 Basic	<ul style="list-style-type: none"> – Provides problem-solving method – Relates method to problem 	“You are correct in saying that the -CH ₃ methyl groups are different from one another. However, the -CH ₂ group is not only 1 group of hydrogens. Those two hydrogens are distinct from the other. The explanation for this relies on the deuterium test. If we deuterated one of those hydrogen [atoms], then a stereocenter is formed; however, there is another stereocenter in the molecule. Therefore, we just formed a diastereomer. If you deuterated the other hydrogen, then you form the opposite diastereomer. Seeing as diastereomers form with the deuterium test, these two hydrogens are different and will be individual ‘groups.’” (No illustrations included in response)
3 Developing	<ul style="list-style-type: none"> – Uses interactive teaching, e.g., questioning to probe or promote students’ understanding or – Uses illustrations or models during explanations 	[Participant included drawing of two Newman projections of 2-bromobutane with deuterium individually substituted for each diastereotopic hydrogen atom] “Ask the student if any chiral centers are present. ‘How do we check if protons adjacent to chiral centers are equivalent?’ Have the student draw both or draw one myself (student draw the other) of structure-projections above. Explain H _A is between Br + H while H _B is between Br + CH ₃ .”
4 Exemplary	<ul style="list-style-type: none"> – Recognizes learners’ prior knowledge 	“When I was a [TA for Organic Chemistry II lab], we taught students that chirality would make two or more protons on an adjacent carbon or adjacent carbons inequivalent. I would also stress in class that with really good NMRs, you may be able to see proton differences on further carbons but we will only grade nearest neighbors as different. I might use a model kit to explain that a certain “face” of the molecule will have protons interacting with the bromine or proton on the chiral center similar to how we explain rings to students.”

Table 4.12. Frequency of reported experiences contributing to TAs’ PCK responses

PCK Question	Prior teaching experience in organic chemistry course	Prior teaching experience in another chemistry course	Experience as a student	Research experience with NMR spectroscopy	Observing an experienced teacher	Other	Mean logit score
PCKQ1	16	0	14	10	5	2	-1.25
PCKQ2	15	1	16	10	5	0	-0.41
PCKQ3	15	1	14	11	6	0	0.28
PCKQ4	15	1	13	6	6	0	1.39

4.11.2 CK component of questionnaire

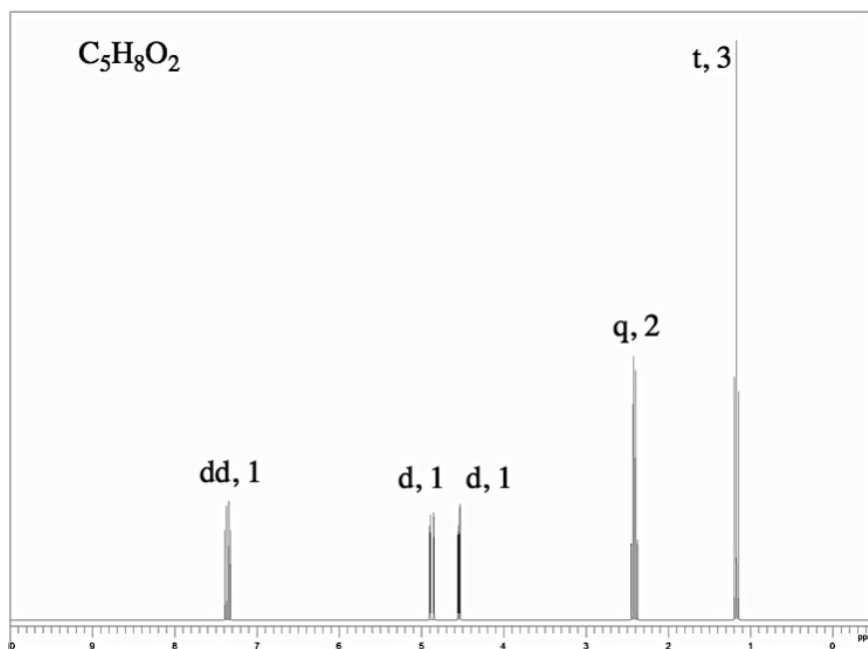


Figure 4.4. CKQ1: Determine the molecular structure corresponding to the given molecular formula and ^1H NMR spectrum.

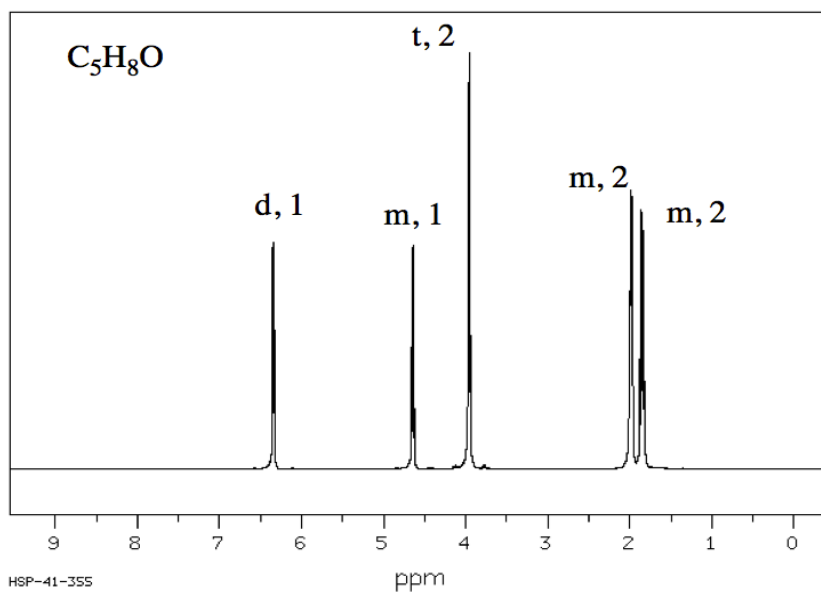


Figure 4.5. CKQ2: Determine the molecular structure corresponding to the given molecular formula and ^1H NMR spectrum.

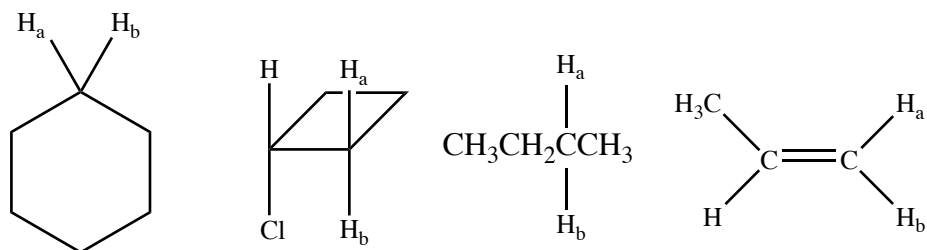


Figure 4.6. CKQ3: In which molecules are H_a and H_b equivalent?

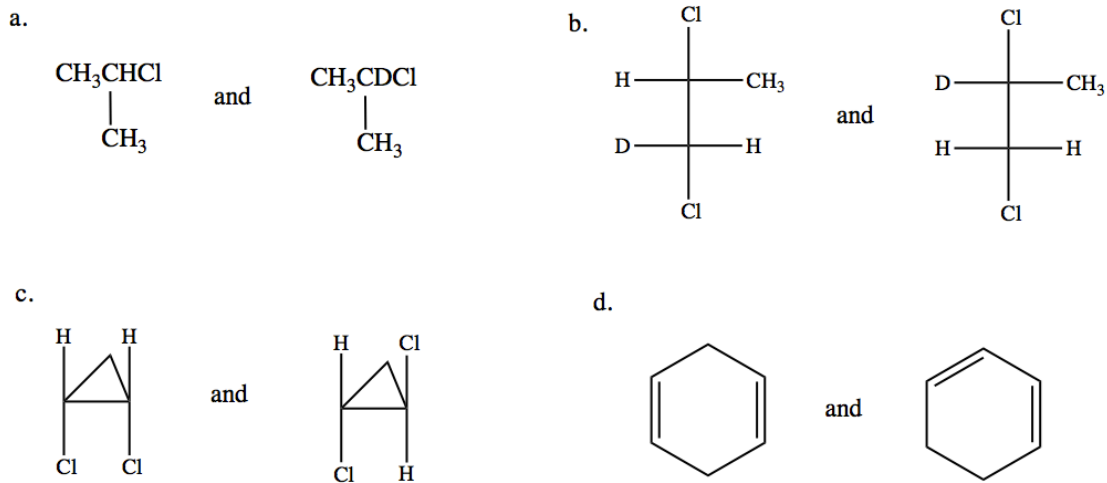


Figure 4.7. CKQ4: Circle all molecule pairs that are distinguishable using ^1H NMR spectroscopy.

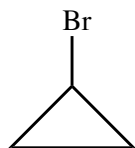


Figure 4.8. CKQ5: How many ^1H NMR signals are produced by the following compound?

4.11.3 PCK component of questionnaire

What aspects of NMR spectral interpretation are most challenging for introductory organic chemistry students?
Please be as specific as possible and provide a response that is complete and concise.

Figure 4.9. PCKQ1, which assessed knowledge of what makes this topic difficult.

A student brings the ^1H NMR spectrum and chemical shift table depicted below to your office hours. The student says that s/he is having trouble figuring out the corresponding molecular structure.

a) Identify and describe particular features (if any) in the ^1H NMR spectrum that may be creating difficulty for the student.

b) As though you are talking with the student, explain how s/he should go about interpreting the spectrum below. Please provide a sufficiently detailed response that completely captures your thinking. Extra space is provided on the next page.

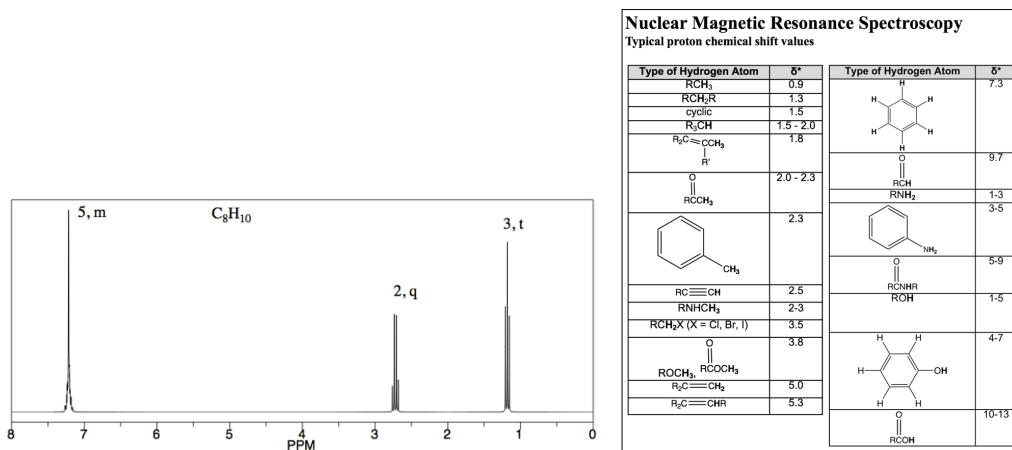


Figure 4.10. PCKQ2, which assessed what makes this topic difficult and teaching strategies.

You are grading problem #1 on a student's quiz (see below).

a) For problem #1, was this student able to provide the correct answer? What (if any) features contributed to difficulty?

b) If incorrect, how would you help this student correctly interpret this spectrum?

Please provide a sufficiently detailed response that completely captures your thinking. Extra space is provided on the next page.

Problem #1

Determine the molecular structure corresponding to the given molecular formula and ^1H NMR spectrum.

Proton chemical shift table provided to Organic Chemistry II laboratory students for problem solving:

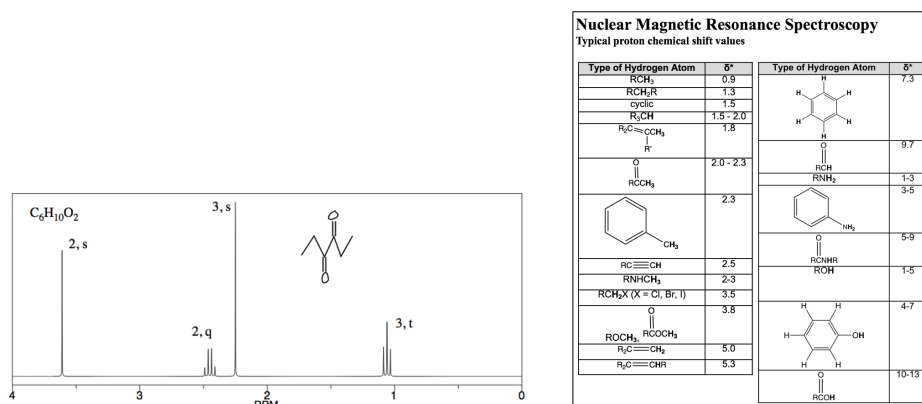


Figure 4.11. PCKQ3, which assessed what makes this topic difficult and teaching strategies.

You are now grading problem #2 on a student's quiz (see below).

a) For problem #2, was this student able to provide the correct answer? What (if any) features contributed to difficulty?

b) If incorrect, how would you help this student determine the correct answer?

Please provide a sufficiently detailed response that completely captures your thinking. Extra space is provided on the next page.

Problem #2

Circle all groups of equivalent hydrogen atoms.

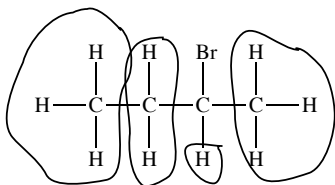


Figure 4.12. PCKQ4, which assessed what makes this topic difficult and teaching strategies.

4.12 References

- (1) Hill, H. C.; Rowan, B.; Ball, D. L. Effects of Teachers' Mathematical Knowledge for Teaching on Student Achievement. *Am. Educ. Res. J.* **2005**, *42* (2), 371–406. <https://doi.org/10.3102/00028312042002371>.
- (2) Park, S.; Jang, J. Y.; Chen, Y. C.; Jung, J. Is Pedagogical Content Knowledge (PCK) Necessary for Reformed Science Teaching?: Evidence from an Empirical Study. *Res. Sci. Educ.* **2011**, *41*, 245–260. <https://doi.org/10.1007/s11165-009-9163-8>.
- (3) Mavhunga, E.; Rollnick, M. Improving PCK of Chemical Equilibrium in Pre-Service Teachers Improving PCK of Chemical Equilibrium in Pre-Service Teachers. *African J. Res. Math. Sci. Technol. Educ.* **2013**, *17* (1–2), 113–125. <https://doi.org/10.1080/10288457.2013.828406>.
- (4) Shulman, L. Knowledge and Teaching: Foundations of the New Reform. *Harv. Educ. Rev.* **1987**, *57* (1), 1–23. <https://doi.org/10.17763/haer.57.1.j463w79r56455411>.
- (5) American Chemical Society Committee on Professional Training For Bachelor's Degree Programs. *Undergraduate Professional Education in Chemistry: ACS Guidelines and Evaluation Procedures for Bachelor's Degree Programs*; American Chemical Society: Washington, DC, 2015.
- (6) Gerothanassis, I. P.; Troganis, A.; Exarchou, V.; Barbarossou, K. Theme Issue 2 on Structural Concepts Contribution from Science (Invited Review Paper) Nuclear Magnetic Resonance (NMR) Spectroscopy: Basic Principles and Phenomena, and Their Applications to Chemistry, Biology, and Medicine. *Chem. Educ. Res. Pract. Eur. Educ. Res. Pr. Eur* **2002**, *3* (3), 229–252.
- (7) Angawi, R. F. Using a Problem Solving-Cooperative Learning Approach to Improve Students' Skills for Interpreting ¹H NMR Spectra of Unknown Compounds in an Organic Spectroscopy Course. *J. Chem. Educ.* **2014**, *91*, 823–829. <https://doi.org/10.1021/ed4004436>.
- (8) Lorigan, G. a; Minto, R. E.; Zhang, W. Teaching the Fundamentals of Pulsed NMR Spectroscopy in an Undergraduate Physical Chemistry Laboratory. *J. Chem. Educ.* **2001**, *78* (7), 956. <https://doi.org/10.1021/ed078p956>.
- (9) Simpson, A. J.; Mitchell, P. J.; Masoom, H.; Liaghati Mobarhan, Y.; Adamo, A.; Dicks, A. P. An Oil Spill in a Tube: An Accessible Approach for Teaching Environmental Nmr Spectroscopy. *J. Chem. Educ.* **2015**. <https://doi.org/10.1021/ed500364b>.
- (10) Wright, N. T. Heteronuclear Multidimensional Protein NMR in a Teaching

- Laboratory. *J. Chem. Educ.* **2016**. <https://doi.org/10.1021/acs.jchemed.5b00331>.
- (11) Winschel, G. A.; Everett, R. K.; Coppola, B. P.; Shultz, G. V.; Lonn, S. Using Jigsaw-Style Spectroscopy Problem-Solving to Elucidate Molecular Structure through Online Cooperative Learning. *J. Chem. Educ.* **2015**, *92*, 1188–1193. <https://doi.org/10.1021/acs.jchemed.5b00114>.
 - (12) Cartrette, D. P.; Bodner, G. M. Non-Mathematical Problem Solving in Organic Chemistry. *J. Res. Sci. Teach.* **2010**, *47* (6), 643–660. <https://doi.org/10.1002/tea.20306>.
 - (13) Topczewski, J. J.; Topczewski, A. M.; Tang, H.; Kendhammer, L. K.; Pienta, N. J. NMR Spectra through the Eyes of a Student: Eye Tracking Applied to NMR Items. *J. Chem. Educ.* **2017**, *94* (1), 29–37. <https://doi.org/10.1021/acs.jchemed.6b00528>.
 - (14) Gess-Newsome, J. A Model of Teacher Professional Knowledge and Skill Including PCK: Results of the Thinking from the PCK Summit. In *Re-examining Pedagogical Content Knowledge in Science Education*; Berry, A., Friedrichsen, P., Loughran, J., Eds.; Routledge: New York, 2015; pp 28–42.
 - (15) Jüttner, M.; Boone, W.; Park, S.; Neuhaus, B. J. Development and Use of a Test Instrument to Measure Biology Teachers' Content Knowledge (CK) and Pedagogical Content Knowledge (PCK). *Educ. Assessment, Eval. Account.* **2013**, *25*, 45–67. <https://doi.org/10.1007/s11092-013-9157-y>.
 - (16) Kind, V. Pedagogical Content Knowledge in Science Education: Perspectives and Potential for Progress. *Stud. Sci. Educ.* **2009**, *45* (2), 169–204. <https://doi.org/10.1080/03057260903142285>.
 - (17) Shulman, L. Those Who Understand: Knowledge Growth in Teaching. *Educ. Res.* **1986**, *15* (2), 4–14. <https://doi.org/10.3102/0013189X015002004>.
 - (18) Smith, P. S.; Banilower, E. R. Assessing PCK: A New Application of the Uncertainty Principle. In *Re-examining Pedagogical Content Knowledge in Science Education*; Berry, A., Friedrichsen, P. M., Loughran, J., Eds.; Routledge: New York, 2015; pp 88–103.
 - (19) Veal, W. R.; MaKinster, J. G. Pedagogical content knowledge taxonomies.
 - (20) Shulman, L. PCK: Its Genesis and Exodus. In *Re-examining Pedagogical Content Knowledge in Science Education*; Berry, A., Friedrichsen, P. M., Loughran, J., Eds.; Routledge: New York, 2015; pp 3–13.
 - (21) Park, S.; Oliver, J. S. Revisiting the Conceptualisation of Pedagogical Content Knowledge (PCK): PCK as a Conceptual Tool to Understand Teachers as Professionals. *Res. Sci. Educ.* **2008**. <https://doi.org/10.1007/s11165-007-9049-6>.

- (22) Cochran, K. F. Pedagogical Content Knowing: An Integrative Model for Teacher Preparation. *J. Teach. Educ.* **1993**, *44*, 263–272.
- (23) Rollnick, M.; Davidowitz, B.; Potgieter, M. Is Topic-Specific PCK Unique to Teachers? In *Cognitive and Affective Aspects in Science Education Research: Selected Papers from the ESERA 2015 Conference*; Hahl, K., Juuti, K., Lampiselkä, J., Uitto, A., Lavonen, J., Eds.; Springer International Publishing: Cham, 2017; pp 69–85.
- (24) Ausubel, D. P.; Novak, J. D.; Hanesian, H. *Educational Psychology: A Cognitive View*; Holt, Rinehart and Winston: Chicago, 1987.
- (25) Bodner, M. G.; Domin, D. S. Mental Models: The Role of Representations in Problem Solving in Chemistry. *Univ. Chem. Educ.* **2000**, *4* (1), 24–30.
- (26) Rollnick, M.; Bennett, J.; Rhemtula, M.; Dharsey, N.; Ndlovu, T. The Place of Subject Matter Knowledge in Pedagogical Content Knowledge: A Case Study of South African Teachers Teaching the Amount of Substance and Chemical Equilibrium. *Int. J. Sci. Educ.* **2008**, *30* (10), 1365–1387. <https://doi.org/10.1080/09500690802187025>.
- (27) Malcolm, S. The Design and Validation of an Instrument to Measure Topic-Specific Pedagogical Content Knowledge in Stoichiometry of Physical Sciences Teachers. **2015**, No. November. <https://doi.org/10.13140/RG.2.1.5174.2161>.
- (28) Mavhunga, E. Improving PCK and CK in Pre-Service Chemistry Teachers. In *Exploring Mathematics and Science Teachers' Knowledge: Windows into teacher thinking.*; Venkat, H., Rollnick, M., Askew, M., Loughran, R., Eds.; Routledge: Abingdon, 2014; pp 31–48.
- (29) Rollnick, M.; Mavhunga, E. The PCK Summit and Its Effect on Work in South Africa. In *Re-examining Pedagogical Content Knowledge in Science Education*; Berry, A., Friedrichsen, P. M., Loughran, J., Eds.; Routledge: New York, 2015; pp 135–146.
- (30) Charalambous, C. Y.; Hill, H. C.; Ball, D. L. Prospective Teachers' Learning to Provide Instructional Explanations: How Does It Look and What Might It Take? *J. Math. Teach. Educ.* **2011**, *14* (6), 441–463. <https://doi.org/10.1007/s10857-011-9182-z>.
- (31) Pitjeng, P. Novice Unqualified Science Teachers' Topic-Specific Pedagogical Content Knowledge and Their Beliefs about Teaching. In *Exploring Mathematics and Science Teachers' Knowledge: Windows into teacher thinking*; Venkat, H., Rollnick, M., Askew, M., Loughran, R., Eds.; Routledge: Abingdon, 2014; pp 65–83.
- (32) Bond, Trevor G. Fox, C. *Applying the Rasch Model: Fundamental Measurement in*

the Human Sciences, 2nd ed.; New York, 2007.

- (33) Collins, D. Pretesting Survey Instruments: An Overview of Cognitive Methods. *Qual. Life Res.* **2003**, *12* (3), 229–238. <https://doi.org/10.1023/A:1023254226592>.
- (34) Dragisich, V.; Keller, V.; Zhao, M. An Intensive Training Program for Effective Teaching Assistants in Chemistry. *J. Chem. Educ.* **2016**, *93* (7), 1204–1210. <https://doi.org/10.1021/acs.jchemed.5b00577>.
- (35) Kelly, P. What Is Teacher Learning? A Socio-Cultural Perspective. *Oxford Rev. Educ.* **2006**, *32* (4), 505–519. <https://doi.org/10.1080/03054980600884227>.
- (36) Bruice, P. Y. *Organic Chemistry*, 6th ed.; Prentice Hall: Upper Saddle River, 2011.
- (37) Raker, J. R.; Reisner, B. A.; Smith, S. R.; Stewart, J. L.; Crane, J. L.; Pesterfield, L.; Sobel, S. G. Foundation Coursework in Undergraduate Inorganic Chemistry: Results from a National Survey of Inorganic Chemistry Faculty. *J. Chem. Educ.* **2015**, *92* (6), 973–979. <https://doi.org/10.1021/ed500624t>.
- (38) Alexander, P. a.; Schallert, D. L.; Hare, V. C. Coming to Terms: How Researchers in Learning and Literacy Talk about Knowledge. *Rev. Educ. Res.* **1991**, *61* (3), 315–343. <https://doi.org/10.3102/00346543061003315>.
- (39) National Institute of Advanced Industrial Science and Technology. SDBSWeb <http://sdfs.db.aist.go.jp> (accessed May 4, 2020).
- (40) PerkinElmer Informatics. ChemDraw Professional 16.0. 2017.
- (41) Rowan, B.; Schilling, S. G.; Ball, D. L.; With, R. M.; Atkins-Burnett, S.; Camburn, E.; Harrison, D.; Phelps, G. Measuring Teachers' Pedagogical Content Knowledge in Surveys: An Exploratory Study* Measuring Teachers' Pedagogical Content Knowledge in Surveys: An Exploratory Study. **2001**.
- (42) Schmelzing, S.; van Driel, J. H.; Jüttner, M.; Brandenbusch, S.; Sandmann, A.; Neuhaus, B. J. Development, Evaluation, and Validation of a Paper-and-Pencil Test for Measuring Two Components of Biology Teachers' Pedagogical Content Knowledge Concerning the "Cardiovascular System." *Int. J. Sci. Math. Educ.* **2013**, *11* (6), 1369–1390. <https://doi.org/10.1007/s10763-012-9384-6>.
- (43) Park, S.; Chen, Y. C. Mapping out the Integration of the Components of Pedagogical Content Knowledge (PCK): Examples from High School Biology Classrooms. *J. Res. Sci. Teach.* **2012**, *49* (7), 922–941. <https://doi.org/10.1002/tea.21022>.
- (44) Mavhunga, E. Transfer of the Pedagogical Transformation Competence across Chemistry Topics. *Chem. Educ. Res. Pr.* **2016**, *17* (4), 1081–1097.

<https://doi.org/10.1039/C6RP00095A>.

- (45) Hale, L. V. A.; Lutter, J. C.; Shultz, G. V. The Development of a Tool for Measuring Graduate Students' Topic Specific Pedagogical Content Knowledge of Thin Layer Chromatography. *Chem. Educ. Res. Pract.* **2016**, *17*, 700–710. <https://doi.org/10.1039/C5RP00190K>.
- (46) Chin, C. Teacher Questioning in Science Classrooms: Approaches That Stimulate Productive Thinking. *J. Res. Sci. Teach.* **2007**, *44*, 815–843.
- (47) Landis, J. R.; Koch, G. G. The Measurement of Observer Agreement for Categorical Data. *Biometrics* **1977**, *33* (1), 159–174.
- (48) Linacre, J. M. A User's Guide to Winsteps.
- (49) Haynes, S. N.; Richard, D. C. S.; Kubany, E. S. Content Validity in Psychological Assessment: A Functional Approach to Concepts and Methods. *Psychol. Assess.* **1995**, *7* (3), 238–247.
- (50) Cronbach, L. J.; Meehl, P. E. Construct Validity in Psychological Tests. *Psychol. Bull.* **1955**, *52* (4), 281–302.
- (51) Linacre, J. M. A user's guide to Facets: Rasch-model computer programs. <http://www.winsteps.com/manuals.htm>.
- (52) Linacre, J. M. Test Validity and Rasch Measurement: Construct, Content, Etc. *Rasch Meas. Trans.* **2004**, *18* (1), 970–971.
- (53) Carlson, R. E. Assessing Teachers' Pedagogical Content Knowledge: Item Development Issues. *J. Pers. Eval. Educ.* **1990**. <https://doi.org/10.1007/BF00126124>.
- (54) Hill, H. C.; Ball, D. L.; Schilling, S. G. Unpacking Pedagogical Content Knowledge: Conceptualizing and Measuring Teachers' Topic-Specific Knowledge of Students Unpacking Pedagogical Content Knowledge: Conceptualizing and Measuring Teachers' Topic-Specific Knowledge of Students. *Source J. Res. Math. Educ. J. Res. Math. Educ.* **2008**, No. 4, 372–400.

Chapter 5

Development of the NMR Lexical Representational Competence (NMR-LRC) Instrument as a Formative Assessment of Lexical Ability in ^1H NMR Spectroscopy

5.1 Initial remarks

While Chapter 4 focused on instructors' knowledge for teaching ^1H NMR spectroscopy, the investigative focus of this chapter shifts to another essential aspect of effective instruction: assessment. Specifically, this chapter describes the development and psychometric evaluation of a formative assessment to measure lexical ability in ^1H NMR spectroscopy. Formative assessments serve as powerful instructional tools for promoting learning, as they provide instructors with low-stakes opportunities for providing feedback and for ensuring that students develop relevant understanding. Further, lexical ability is a key facet of expertise in ^1H NMR spectral interpretation; organic chemists regularly use NMR spectra to support their social discourse, a practice which requires the use of words to communicate the identification, analysis, and interpretation of spectral features.¹ A formative assessment that measures lexical ability in ^1H NMR spectroscopy would thus provide instructors with a means of determining if their students are prepared to engage in such discourse and a means of providing feedback that promotes relevant competency. Further, psychometrically evaluating this formative assessment will help ensure that instructors can validly and reliably interpret assessment scores as a measure of lexical ability when using this instrument in their own classrooms.

The author first recognized the potential instructional value of such an assessment when conducting interviews with undergraduates in the study presented in Chapter 2. During these interviews, undergraduates often struggled to verbally identify spectral features, instead preferring to use the mouse cursor to point toward features they were referencing. Doctoral participants in the study presented in Chapter 3, however, were able to verbally identify spectral features with ease. This difference made evident the need for

instruction focused on cultivating lexical ability in ^1H NMR spectroscopy. Surprisingly, there are no published psychometrically-evaluated assessments on ^1H NMR spectroscopy, despite assessment being an essential component of effective instruction. Results presented in this chapter include the first of such an assessment. Specifically, results include a 10-item formative assessment of lexical ability in ^1H NMR spectroscopy, as well as a range of validity and reliability evidence to support the interpretation of data obtained from this instrument. In addition to its instructional value, the instrument will also provide researchers with a means of evaluating learning activities designed to cultivate lexical ability.

This chapter is anticipated to appear as a research article in a chemistry education journal. Benjamin Glass, the author's research mentee, assisted with instrument design, conducting cognitive interviews, quantitative data analysis, and writing of the manuscript. However, the majority of this work was completed by the author.

5.2 Abstract

Nuclear magnetic resonance (NMR) spectra are among the most common visualizations used by chemists in both academic and industrial settings, where they serve as both representations of submicroscopic entities and tools that support social discourse. The ability to use words to communicate the identification, analysis, and interpretation of features within NMR spectra is thus an essential aspect of chemists' representational competence. Collectively, these words constitute a highly specialized language specific to this technique. To adequately prepare chemists, the undergraduate curriculum must therefore cultivate students' ability to use this lexicon. We developed the NMR Lexical Representational Competence (NMR-LRC) instrument, a formative assessment to measure students' ability and perceived ability to use words to communicate the identification, analysis, and interpretation of features within ^1H NMR spectra. Following development, we administered the NMR-LRC in Spring 2020 to a total of $N=678$ second-semester organic chemistry students at a large Midwestern university. We analyzed responses using the Rasch model to collect statistical evidence of validity and reliability, and we used cluster analysis to evaluate whether the instrument could detect the

Dunning-Kruger effect (i.e., students' illusions of lexical representational competence). Supporting sources of response process, content, and associative validity evidence included cognitive interview data, instructional and disciplinary expert review, and the correlation of Rasch ability measures with an external measure, respectively. Results suggest that data obtained from the NMR-LRC can be interpreted as a measure of students' lexical representational competence and perceived lexical representational competence in ^1H NMR spectroscopy and that the NMR-LRC can detect students' illusions of such competence. These results further suggest that instructors can use the NMR-LRC to formatively assess their students' lexical ability and perceived lexical ability in ^1H NMR spectroscopy following instruction, allowing them to provide instructor feedback that promotes students' competence and perceived competence. Instructors can also use the instrument to identify students exhibiting the Dunning-Kruger effect who will require alternate forms of feedback. Lastly, instructors and researchers can use the NMR-LRC to evaluate the efficacy of learning activities designed to cultivate lexical ability in ^1H NMR spectroscopy.

5.3 Introduction

Nuclear magnetic resonance (NMR) spectroscopy is an essential tool for the structural characterization of organic compounds, and NMR spectra are among the most common visualizations used by organic chemists.^{1,2} These spectra serve dual functions in both academic and industrial settings: they act as material representations of submicroscopic entities that would otherwise be imperceptible, and they support chemists' social discourse during the synthesis of compounds.² More specifically, Kozma and Russell demonstrated in a foundational ethnographic study that chemists use NMR spectra to confirm the molecular structure of synthesis products and to convince others that synthesis products possess the intended molecular structure.² The latter action involved using particular spectral features as warrants for claims that syntheses were indeed successful. This use of NMR spectra, combined with use of other visualizations, ultimately served to integrate their participants into the scientific community of chemists.²

Kozma and Russell's study makes evident NMR spectroscopy's essential and multifaceted disciplinary role. Its importance is further underscored by multiple reports in this journal describing both NMR instructional strategies and laboratory activities.^{3–10} Within this larger body of literature are a small number of research studies that investigate how individuals interpret NMR spectra and how instructors develop knowledge for teaching this topic.^{11–14} These studies focus predominantly on the use of NMR spectra as material representations of submicroscopic entities, though these spectra also play a critical social function.² Individuals' ability to use words to communicate the identification, analysis, and interpretation of features within NMR spectra is particularly understudied, as NMR spectroscopy has its own complex lexicon (e.g., upfield, downfield, multiplicity, etc.) and yet use of this highly specialized language is essential for discourse. Sociologists also note that knowing the language surrounding a practice may contribute more to related understanding than physically engaging in the practice;¹⁵ knowing the NMR lexicon may therefore play a central, cognitive role in understanding the practice of spectral interpretation.

Findings from one study on undergraduates' interpretation of NMR spectra suggest that learning the related lexicon poses a challenge to introductory organic chemistry students.¹⁴ Kozma and Russell's findings also suggest that developing relevant lexical ability may be challenging for this population; in their study, undergraduates' laboratory discourse focused predominately on the setup of equipment and not on visualizations, unlike that of practicing chemists. Other researchers have also noted that using scientific visualizations is an inherently social practice and that learning associated terminology poses a challenge to undergraduates.¹⁶ In their study of biology undergraduates' and practicing scientists' interpretation of graphs, Bowen *et al.* focused their analysis on five essential components of a social practice, in particular the linguistic resources (i.e., nouns, verbs, specialized terminology, etc.) that individuals use to make distinctions necessary for efficiently accomplishing activities in a field.^{16,17} This study found that students were often unable to make important distinctions between terminologies, resulting in ambiguity and a failure to arrive at shared interpretations during group work.

Further, students were often unaware of these ambiguities, leaving them unable to address resulting breakdowns in their understanding and discourse.

Empirical evidence of this learning challenge among undergraduates, combined with the important social and cognitive functions of NMR spectra and associated language, indicates that organic chemistry instruction should aim to cultivate related lexical ability. One promising initial step for designing such instruction involves developing a formative assessment that measures undergraduates' ability to use words to communicate the identification, analysis, and interpretation of features within NMR spectra. Developing assessments that elicit evidence of undergraduates' engagement with scientific practices, crosscutting concepts, and disciplinary core ideas continues to play an essential role in the ongoing transformation of higher education STEM courses.¹⁸ The National Research Council's report "A Framework for K-12 Science Education" outlines these three dimensions and argues that they should be integrated into all aspects of learning, including assessment, in order for students to achieve desired learning goals.¹⁹ As noted, using NMR spectra as tools for social discourse is an essential practice in organic chemistry; developing an assessment of lexical ability in this practice would thus contribute to this larger transformation. Further, formative assessments typically provide low-stakes opportunities for students to demonstrate their performance without penalty and receive feedback that ultimately promotes learning.²⁰ This feedback can take many forms. The simplest form involves the instructor indicating that a response is correct or incorrect. More complex forms involve the instructor providing written or verbal statements that are cognitively accessible to students.²⁰ Such an assessment tool would thus allow instructors to provide feedback that promotes students' lexical ability, ensuring that undergraduates completing the chemistry curriculum are prepared to engage in necessary social discourse. It will also allow researchers and instructors to gauge the efficacy of instructional strategies designed to cultivate such ability. This tool will further serve as the first published, psychometrically-evaluated assessment focused on NMR spectroscopy; given the disciplinary value of NMR spectroscopy, such an assessment will constitute a first step toward building a repository of assessments focused on this essential technique.

5.4 Conceptual framework

This work is grounded in the theories of representational competence and methodological interactionism.^{2,15} These theories propose similar, complementary relationships between language, scientific practice, and the development of expertise. The theory of representational competence primarily guided the design of a formative assessment to measure lexical representational competence, a potential subconstruct of representational competence. Both theories then provided theoretical support for developing such an assessment and insight into how instructors could use the assessment to support learning.

5.4.1 Representational competence

First conceptualized by Kozma and Russell, representational competence is a set of skills that allows one to use representations of otherwise imperceptible physical entities and processes to reason about, communicate, and act on corresponding chemical phenomena.² This skillset includes a number of abilities identified through investigating the daily activities of practicing chemists. These abilities involve using representations to describe chemical phenomena in terms of underlying molecular entities and processes, using words to communicate about specific features of representations, making connections across representations, and describing the information that one representation affords over another, among others.² Representational competence is essential for participation in the chemical enterprise, and research demonstrates that these collective abilities evolve with increasing expertise in the domain. For instance, novices tend to engage in the rule-based application of symbols and use explicit surface features of representations to define corresponding phenomena. Conversely, experts use representations in a more rhetorical sense by relying on them to explain phenomena in terms of implicit features, identify warrants for their claims, and collectively solve problems.² This developmental trajectory aligns with the Vygotskian notion of a “zone of proximal development” and sociocultural theory, which acknowledge that one’s development is influenced by interactions with both material and social aspects of their

environment.²¹ It also aligns with situative theory, as it recognizes that the use of representational conventions integrates one into a community of practice.²²

The chemistry education community has recognized the importance of cultivating this specific skillset among undergraduate chemistry students. As a result of this recognition, several research studies have investigated how undergraduates develop representational competence in chemistry.^{23,24} Use of this framework has gained traction in studies that focus on students' learning in organic chemistry, with the construct often being used as a lens to understand students' meaning-making while they interpret and transform diagrams, structures, and mechanisms.²⁵ Findings from these studies focus primarily on students' ability to infer implicit molecular properties (e.g., electronegativity, partial charges, and reactivity) from organic structural formulae, as well as their use of the electron-pushing formalism.^{24,26} Representational competence has yet to be explicitly used as a framework in studies investigating students' learning of NMR spectroscopy, despite NMR spectra being one of the most common representations used by practicing chemists.² And while these studies provide some insight into how students make sense of submicroscopic entities using NMR spectra, students' use of NMR spectra as a tool for social discourse has as a result gone understudied.

5.4.2 Lexical representational competence

The ability to use words to communicate about features within representations is an essential skill among practicing chemists because it facilitates their participation in social discourse and integration into the scientific community.² We refer to this ability as lexical representational competence, a potential subconstruct of representational competence that merits its own investigation given its critical role. Lexical representational competence in NMR spectroscopy, or the ability to use words to communicate the identification, analysis, and interpretation of features within NMR spectra, particularly merits its own investigation given the amount of terminology inherent to this technique (e.g., shielded, de-shielded, etc.). Chemists must be familiar with this specialized language to identify and explain how features of NMR spectra support their

claims about molecular structure — the communication of data, claims, and warrants that serves as the basis of scientific argumentation.²⁷

Lexical representational competence aligns closely with the linguistic resources component of social practices,¹⁷ though we consider this subconstruct specific to representations and their associated lexicalized concepts. Lexicalized concepts are those that correspond to a word in a language; they are typically shared within a community and remain relatively stable over time.²⁸ The chemistry education community has recognized the importance of promoting students' ability to use words comprising the chemical language more generally.²⁹ Chemistry education researchers have designed and evaluated a range of literacy-based instructional strategies,³⁰ as well as identified a number of issues associated with language and learning chemistry.^{31,32} For instance, language has been identified as a contributor to students' cognitive overload, inhibiting their problem-solving ability.³¹ Language also serves as a barrier to developing understanding of complex concepts, particularly when a term used in a chemistry context has a different meaning than the same term used in a colloquial context (e.g., “strong” acids).³²

Moreover, a chemist should also be confident in their lexicon usage of if they are to employ this language for efficient communication. If one doubts their ability to use this specialized language, they will likely be less capable of participating in this social discourse than one who is confident in their ability. For instance, a novice chemist may need to communicate that a peak is shifted further left on an NMR spectrum. The chemist may believe this peak should be described as “downfield,” though they may be uncertain as they often confuse the terms upfield and downfield; this uncertainty would serve as a barrier to efficient communication. This notion is grounded in Bandura's social cognitive theory of self-efficacy, which posits that a person's belief in their ability to successfully perform a task will affect their actual performance.³³ A number of studies have demonstrated a positive relationship between self-efficacy and performance in chemistry.^{34,35} A chemist should thus likely possess perceived lexical representational competence in addition to lexical representational competence to fully engage in social discourse. Developing an inferential measure of undergraduates' lexical representational

competence and perceived lexical representational competence in NMR spectroscopy is therefore essential for gauging if students are prepared to engage in the more expert-like rhetorical usage of spectra.

5.4.3 Illusions of lexical representational competence

In addition to determining if chemistry students have both lexical representational competence and perceived competence, effective instruction will also require that instructors identify students with illusions of lexical representational competence. Illusions of competence are marked by high perceived ability in a skill but low actual ability, a common psychological phenomenon termed the Dunning-Kruger effect.³⁶ The Dunning-Kruger effect is well-documented for a range of abilities; moreover, it is a robust phenomenon in introductory chemistry courses.^{36–38} The effect is problematic in these educational contexts, as students with high perceived ability but low ability are likely unaware that they need to take corrective steps to improve.³⁷ Research further demonstrates that chemistry students' illusions of competence are not disrupted by traditional forms of feedback that would presumably prompt a recalibration of their perceived ability (e.g., low exam and homework scores).³⁷ These miscalibrations thus tend to persist over time unless they are met with alternate forms of feedback, specifically metacognitive training.^{36,37,39,40} For instance, research suggests that having students calibrate their perceived ability using guidelines or review their past performance helps with recalibration.^{39,40} Identifying students who harbor illusions of lexical representation competence will thus allow instructors to provide unique feedback that helps them to improve their lexical ability.

5.4.4 Methodological interactionalism

The theory of methodological interactionalism further underscores the importance of language to scientific practices. According to this theory, simply engaging in the physical aspects of a scientific practice is insufficient for developing related understanding and skills.¹⁵ Rather, learning a practice from someone else necessitates a shared language; in environments where one learns a practice, language is then always the learning

mechanism.¹⁵ The language of a domain plays this central role in learning because it gives meaning to specific practices; in the words of sociologist Harry Collins, “languages ‘contain’ practices.”^{15,41} Even in instances where learning a practice appears to occur primarily through immersion in related physical aspects, the role of language is still central;¹⁵ engaging in the physical aspects of a practice simply provide the condition for immersing oneself in the shared language.¹⁵ Developing expertise in a practice thus entails a collective means of learning mediated by language rather than individualist interactions with physical aspects.¹⁵ Further, knowing the language surrounding a practice allows one to understand the practice without necessarily being able to execute related physical aspects.⁴² This practical understanding is essential for making advancements in the sciences, as progress in different domains requires contributions from a range of specialists who can then coordinate their actions through a common spoken discourse.^{15,42}

In the context of interpreting NMR spectra, understanding a term thus carries with it a practical understanding of the corresponding aspect of interpreting spectra. For instance, understanding the term “coupling constant” carries with it a practical understanding that one deduces this value from signals to identify coupling partners. Likewise, understanding the terms “upfield and downfield” carries with it a practical understanding that one uses these relative peak locations to infer molecular structure. Undergraduates must first acquire this practical understanding to execute the physical aspects of interpreting spectra (e.g., physically determining coupling constants to identify coupling partners or inferring structure from relative peak position). Effective undergraduate-level instruction on interpreting NMR spectra will thus involve cultivating students’ knowledge of the lexicon already shared among organic chemists, instructors, and curricular materials. Instruction that fosters knowledge of this lexicon and, in turn, practical understanding will more effectively support students’ ability to execute physical aspects of interpreting NMR spectra and to engage in organic chemists’ common spoken discourse. This latter ability will be essential for students pursuing careers in chemistry not involving spectral interpretation, as these individuals will still need to engage in common spoken discourse to coordinate activities in the larger domain. Instructors can

thus use a formative assessment of lexical representational competence in NMR spectroscopy to help determine if students have this practical understanding.

5.5 Research goals

The primary aim of this study was to develop an instrument that provides a measure of students' ability and perceived ability to use words to communicate the identification, analysis, and interpretation of features within ^1H NMR spectra. ^1H NMR spectroscopy was selected as the focus of the instrument because this technique is the most widely used among chemists.³⁸ More specifically, this investigation was guided by the following goals:

1. To develop an instrument that provides an inferential measure of second-semester organic chemistry students' lexical representational competence and perceived lexical representational competence in ^1H NMR spectroscopy
2. To determine if this instrument can be used to detect illusions of lexical representational competence among second-semester organic chemistry students

The first goal will provide instructors with a means of formatively assessing if their students possess lexical representational competence and perceived lexical representational competence in ^1H NMR spectroscopy and are thus prepared to engage in relevant social discourse and learn the physical aspects of interpreting spectra. Using data obtained from this instrument, instructors could then provide feedback that promotes students' competence or perceived competence. For instance, instructors could provide students exhibiting low ability and low perceived ability with written statements regarding misused terms and suggested areas to review. Similarly, instructors could provide students exhibiting high ability but low perceived ability with encouragement to promote their self-efficacy.⁴⁴ The second goal will provide instructors with a means of identifying students with low lexical ability but high perceived lexical ability, or those who may be resistant to the above forms of feedback. Instructors can then provide students who exhibit the Dunning-Kruger effect with alternate forms of feedback like metacognitive training to improve their lexical ability.

5.6 Methods

5.6.1 Participants and instructional context

Participants included two student cohorts enrolled in a second-semester organic chemistry laboratory course at a large, public Midwestern university during Fall 2019 (Sample 1) and Spring 2020 (Sample 2). Sample 1 was part of the instrument development phase of the study, and Sample 2 was part of the instrument evaluation phase of the study. Sample 1 included five participants, with all participants enrolled in a single section of the course. Of these participants, four were female and all identified as white. Participants in Sample 2 were enrolled across three sections of the course, with two sections taught by one instructor and one section taught by another instructor. Each section included a combination of majors and nonmajors. The three sections contained a total of 814 students, and 678 students consented to participate and completed the study. Of the participants in Sample 2, 64% were female and 47% identified as nonwhite. Two students from Sample 2 participated in interviews; of these participants, one was female and both identified as nonwhite. Instruction for Sample 2 occurred in-person for half of the Spring 2020 semester; following the onset of the COVID-19 pandemic, instruction then occurred online. All individuals voluntarily consented to participate in the study, and Institutional Review Board approval was obtained.

^1H NMR spectroscopy is taught in detail in the laboratory course in which the study took place. Instruction on this topic occurred during weekly, one-hour lectures and covered concepts necessary to (1) interpret features of a ^1H NMR spectrum (e.g., number of peaks, peak position, integration, splitting, some second-order splitting, coupling constants, etc.); (2) match compounds to appropriate NMR spectra; (3) use an NMR spectrum to predict molecular structure; and (4) compare ^1H NMR spectra obtained in lab to spectra from literature. Instructors shared lecture notes with one another, though each instructor independently selected which ^1H NMR concepts and corresponding terminology they incorporated into their respective lecture. Sections being taught by different instructors added a degree of variability to the instructional context; this variability provides some support for interpreting data obtained from an instrument similarly across

instructional contexts. In addition to lecture-based instruction, students also interpreted ^1H NMR spectra in several laboratory sessions with teaching assistants' help. These sessions occurred virtually following the onset of the COVID-19 pandemic. Lastly, students were provided with a coursepack containing optional practice problems involving ^1H NMR spectral interpretation. Laboratory protocols and coursepacks did not differ across sections or semesters.

5.6.2 Instrument development

To develop the NMR Lexical Representational Competence (NMR-LRC) instrument, we first used two organic chemistry textbooks to identify terms specific or closely related to ^1H NMR spectra.^{45,46} We then designed ten items to assess individuals' ability and perceived ability to use words to communicate the identification, analysis, and interpretation of features within ^1H NMR spectra. Each item included a multiple-choice question to assess lexical representational competence and an associated confidence tier to assess perceived competence (Figure 5.1). Each multiple-choice question incorporated either one term (e.g., chemical shift) or one set of closely related terms (e.g.,

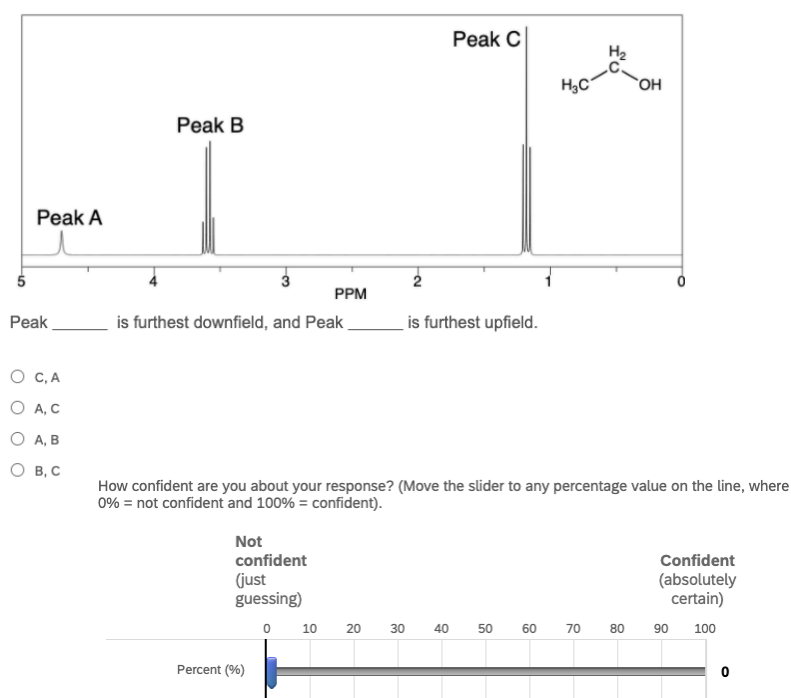


Figure 5.1. NMR-LRC multiple-choice question and associated confidence tier.

Table 5.1. Ten NMR-LRC items, including terminology in each multiple-choice question (in italics) and a question summary. Each question summary reflects the final question format following changes to ensure response process validity.

Item	Terminology and question summary
Upfield and downfield	Select ^1H NMR peaks that are furthest <i>upfield</i> and furthest <i>downfield</i> on a spectrum
Chemical shift	Identify the <i>chemical shift</i> of a peak on a spectrum
Proton exchange	Identify ^1H NMR peak(s) corresponding to hydrogen atom(s) undergoing <i>proton exchange</i>
Shielded and de-shielded	Identify ^1H NMR peaks corresponding to protons that are the most <i>shielded</i> and <i>de-shielded</i>
Peak area	Identify ratio of <i>peak area</i> using the number of protons corresponding to ^1H NMR peaks
The N+1 rule	Identify number of protons adjacent to protons corresponding to an ^1H NMR peak using the <i>N+1 rule</i>
Multiplicity	Assign <i>multiplicity</i> to peaks on a spectrum
Topicity	Identify how <i>topicity</i> of hydrogen atoms will affect appearance of corresponding peaks on a spectrum
Spin-spin coupling	Identify groups of protons involved in <i>spin-spin coupling</i>
Coupling constant	Determine the <i>coupling constant</i> from an ^1H NMR peak given the coupling constant of the coupling partner

upfield and downfield) (Table 5.1). Questions required individuals to either select terms that describe spectral features, to identify spectral features corresponding to given terms, or to identify molecular features corresponding to given terms (Table 5.1). Questions included simple ^1H NMR spectra and molecules to assess general lexical ability rather than advanced understanding.

Research demonstrates that pairing a confidence tier with multiple-choice questions on concept inventories serves as an effective means of measuring chemistry students' understanding and perceived understanding, as well as identifying students exhibiting the Dunning-Kruger effect.⁴⁷ To obtain a measure of perceived competence, an interval confidence tier originally published by McClary and Bretz (2012) was paired with each multiple-choice question and asked students to report their confidence from 0% (Just Guessing) to 100% (Absolutely Certain) using a sliding scale (Figure 5.1).⁴⁸ The sliding scale could be adjusted in increments of 1%. Higher confidence scores could be interpreted as higher perceived lexical representational competence. An interval confidence tier was selected rather than a Likert scale to circumvent potential issues associated with conducting statistical analyses on ordinal data.⁴⁷ The alignment of terms with widely published instructional materials contributed to the content validity of the instrument, or the degree to which a given measure represents the intended domain.⁴⁹

To collect additional evidence of the NMR-LRC's content validity, a disciplinary and instructional expert in organic chemistry was interviewed to provide feedback on the extent to which items assessed lexical ability in ^1H NMR spectroscopy. This expert had over ten years of research and teaching experience relating to ^1H NMR spectroscopy. A small number of changes to question wording were made following this interview to be more in line with verbiage used by practicing organic chemists.

Following initial development, the NMR-LRC was administered via Qualtrics, an online survey tool, to Sample 1 as part of a pilot study. The instrument was administered following instruction on NMR spectroscopy; this instruction is described in detail above. The pilot study served to collect evidence of response process validity, or the extent to which individuals' interpretation of items in a survey match the interpretation intended by instrument developers.⁴⁹ During piloting, individuals completed the NMR-LRC and then participated in one-on-one, semi-structured cognitive interviews.⁵⁰ Cognitive interviews serve as a common means of collecting evidence of response process validity. During each interview, participants verbalized all of their thinking involved in responding to each item. Participant recruitment and data collection continued until participants expressed no new thinking, indicating data saturation had been achieved.⁵¹ These verbalizations allowed for the identification of multiple-choice questions with wording that either resulted in participants responding correctly without knowing given terms or responding incorrectly while knowing given terms. The wording was then altered as necessary to help ensure that questions would be interpreted as intended. Alterations are described in the Results and Discussion section.

5.6.3 Instrument evaluation

Following development, the NMR-LRC was administered to Sample 2. Two students from this study population also participated in cognitive interviews to further ensure that questions were interpreted as intended. The instrument was administered after students had completed instruction on NMR spectroscopy. Students received bonus points for completing the survey, and responses were exported from Qualtrics for analysis. Students who completed the survey but did not consent to participate still received bonus

points. Data obtained from the NMR-LRC were first analyzed using a variety of descriptive statistics. Data were then psychometrically analyzed using the Rasch model to collect statistical evidence of validity and reliability; this evidence was collected to support the interpretation of data obtained from the NMR-LRC and its use as a formative assessment of introductory organic chemistry students' lexical ability and perceived lexical ability in ^1H NMR spectroscopy. Supporting evidence of associative validity was collected by evaluating the correlation of participant's Rasch ability measure and their cumulative quiz score from the laboratory course in which they were enrolled. The cumulative quiz score was the sum of five individual quiz scores, where quiz questions related to NMR spectroscopy, reaction mechanisms, percent yield, and safety procedures. Following analysis using the Rasch model, the instrument's ability to detect the Dunning-Kruger effect was evaluated using cluster analysis. A variety of statistical packages were used for analyses, including the R Stats Package in RStudio, Winsteps 4.0.0, and IBM SPSS Statistics 25.0.^{46–48}

Rasch analysis. The Rasch model is a one-parameter item response theory (IRT) model that is commonly used to psychometrically evaluate assessment data.⁴⁹ This model is unidimensional and provides measures of item difficulty and individuals' ability for a single latent trait.⁴⁹ The probability that an individual will correctly respond to an item is then given as function of the individual's ability and item difficulty measures, where individuals with higher ability measures will have a greater probability of correctly responding to items of a given difficulty.⁴⁹ The model is particularly useful because it transforms individuals' raw test scores on the ordinal scale to ability measures in units of "log odds" (i.e., logits), which are on the interval scale with uniform spacing.⁴⁹ Raw test scores cannot be assumed as interval-scale measures of ability because differences in scores may not reflect uniform differences in ability. For instance, an instrument may contain one very difficult question and nine very easy questions. Students correctly answering eight or nine easy questions likely have small differences in ability. However, students correctly answering nine or ten questions likely have larger differences in ability, as the tenth question was very difficult. Transformation to the interval scale allows for a meaningful comparison of differences in individuals' ability. The Rasch model also places

measures of individuals' ability and item difficulty on the same scale, allowing instrument developers to ensure that questions exhibit a range of difficulty and in turn reliably measure a range of ability.⁴⁹

Unlike other IRT models which involving fitting models to data, the Rasch model requires that data fit the model. Ensuring fit to the model requires evaluating the data's unidimensionality, the local independence of items, and a number of fit statistics, all of which serve as evidence of validity and reliability (or lack thereof) to support the interpretation of data obtained from an instrument.⁵⁰ Given its utility, the Rasch model is often used in chemistry education research for evaluating concept inventories, i.e., instruments designed to inferentially measure understanding in a given domain.^{50,51} While the NMR-LRC is not a concept inventory, it is an assessment designed to inferentially measure ability in a given domain. The Rasch model was thus well suited to collect evidence of validity and reliability that would support the interpretation of data obtained from the NMR-LRC, as well as provide insight into which terms may be most accessible or challenging for students. For our analysis, multiple-choice responses on the ten NMR-LRC items were analyzed using the Rasch model. This analysis provided evidence that would support the interpretation of multiple-choice responses and confidence ratings as measures of lexical representational competence and perceived lexical representational competence, respectively. Evidence from this analysis supports the interpretation of data obtained from both item components, as confidence ratings on questions measuring a latent trait would necessarily measure one's confidence for that latent trait. Correct responses to multiple-choice questions on the NMR-LRC were given a score of one and incorrect responses were given a score of zero. This raw data was then imported into Winsteps 4.0.0 for subsequent analysis using the dichotomous Rasch model.

Cluster analysis. Cluster analysis is a form of data mining that involves partitioning a dataset into subsets, i.e., clusters.⁵² Clusters contain data objects that are similar to one another and yet dissimilar to data objects in other clusters;⁵² this class of statistical techniques thus provides a means of identifying groups of study participants exhibiting similar response patterns within a dataset. The set of techniques employs a variety of algorithms to identify clusters through unsupervised learning. Given that

clusters are not known *a priori*, the meaning and significance of clusters must be interpreted by the researcher. When using cluster analysis to identify groups of study participants with similar response patterns, the researcher must therefore interpret the meaning and significance of response patterns for their given context. To investigate the NMR-LRC's potential to detect illusions of lexical representational competence, Rasch ability measures and average confidence ratings were used to cluster participants with similar response patterns. Participants exhibiting the Dunning-Krueger effect would be expected to exhibit low Rasch ability measures and high average confidence ratings; a cluster of participants with this response pattern would thus provide evidence that the NMR-LRC can be used to detect illusions of competence. Rasch ability measures were used for this analysis rather than total scores, as these interval-scale measures have uniform spacing that more accurately reflect differences in ability.⁴⁹

A TwoStep clustering procedure was used to identify clusters of participants with similar response patterns.⁵³ The TwoStep algorithm generates clusters quickly by relying on only one iteration through raw data, and it provides solutions without predefining an assumed number of clusters.⁵³ The clustering algorithm is based on a distance measure that can produce solutions based on continuous variables, and these solutions are best when variables are independent and normally distributed. However, the algorithm is quite robust and functions well when these assumptions are not met.⁵³ The log-likelihood criterion was used as a distance measure for this analysis, as it functions with either continuous or categorical data.⁵³ The first step of the procedure involves generating preclusters; during this step, the algorithm uses the distance measure to evaluate data objects (e.g., participants' responses) one-by-one to determine if they should be merged with an existing precluster or used to start a new precluster. Once all data objects are evaluated, the second step involves using a standard hierarchical clustering algorithm that iteratively merges preclusters until larger clusters remain.⁵³ The Schwarz Bayesian Criterion is then used to select the optimal number of clusters.⁵³ The TwoStep cluster analysis was conducted using IBM SPSS Statistics 25.0.⁴⁸

5.7 Results and discussion

5.7.1 Descriptive statistics

Descriptive statistics of participants' total raw scores and average confidence ratings are provided in Table 5.2. Possible total raw scores on the NMR-LRC ranged from 0 to 10, and possible average confidence ratings ranged from 0% to 100%. The median total raw score on the NMR-LRC equaled 6.0, suggesting that the NMR-LRC was not overly easy or difficult for the study population. Similar to this median value of 6.0, the mean average confidence rating equaled 62.5%. These comparable values, combined with the similar distribution of total scores and average confidence ratings (Figure 5.2A-B), initially suggest that the study population appropriately gauged their perceived ability.

Table 5.2. Descriptive statistics of participants' total scores and average confidence ratings.

N=678	Total raw scores (0 – 10)	Average confidence ratings (0-100%)
Mean	n/a for ordinal data	62.5%
Std. dev.	n/a for ordinal data	15.6%
Median	6.0	64.1%
Minimum	1.0	0.0%
Maximum	10.0	100.0%

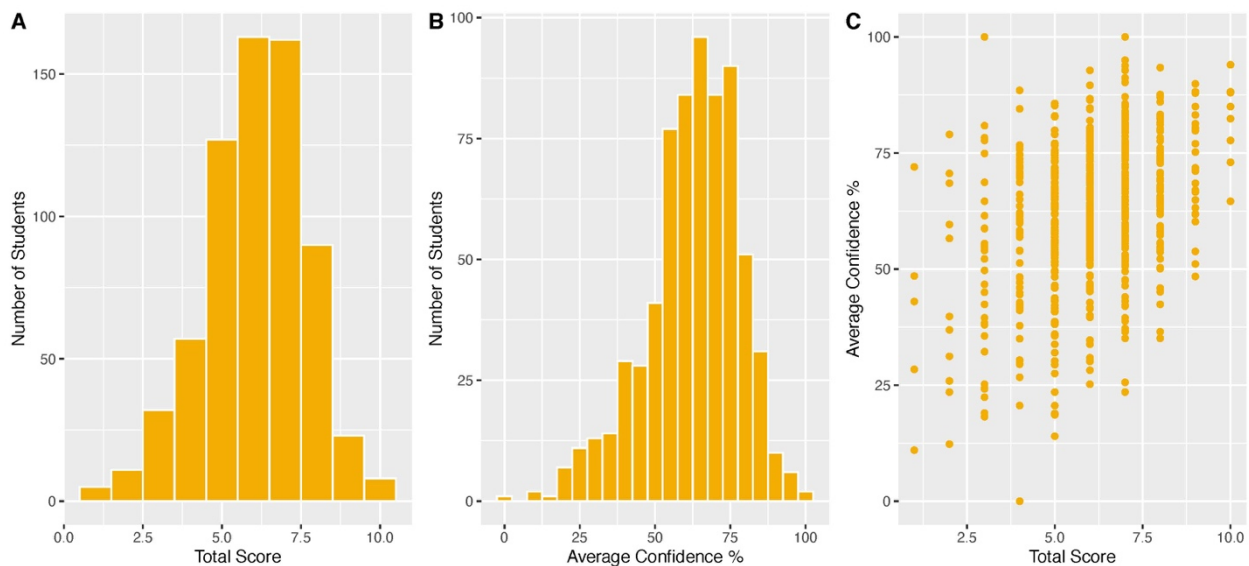


Figure 5.2. Descriptive statistics of NMR-LRC data with A) distribution of participants' total raw scores on the NMR-LRC, B) distribution of participants' average confidence ratings, and (C) a scatterplot of participants' total raw scores and average confidence ratings.

However, the scatter plot of participants' total scores and average confidence ratings reveals a wide range of perceived abilities at low total raw scores (Figure 5.2C), suggesting that the study population may contain individuals holding illusions of competence and that the NMR-LRC can be used to identify such individuals. This possibility was further investigated via cluster analysis following psychometric evaluation of the NMR-LRC using the Rasch model.

5.7.2 Statistical evidence of validity and reliability from Rasch analysis

Analysis of NMR-LRC multiple-choice responses using the Rasch model provided a range of validity and reliability evidence, including statistical evidence of structural validity, content validity, response process validity, and item response reliability; this evidence provides support for interpreting data obtained from the NMR-LRC as a measure of lexical representational competence and perceived lexical representational competence in ^1H NMR spectroscopy, as well as support for using the NMR-LRC as a formative assessment of introductory organic chemistry students' lexical ability and perceived lexical ability in ^1H NMR spectroscopy.

Unidimensionality of NMR-LRC data. The Rasch model is a unidimensional measurement model, meaning that data analyzed via Rasch analysis must also be unidimensional, i.e., measure a single latent trait. Evaluating the unidimensionality of data obtained from the NMR-LRC thus serves as a means to evaluate whether the Rasch model is an appropriate measurement model for psychometric analysis. Further, evaluating unidimensionality also serves as means of collecting evidence of the NMR-LRC's structural validity. Structural validity concerns the internal structure of an instrument, i.e., the relationship among instrument items. Evidence of structural validity establishes the degree to which the internal structure of an instrument matches its intended structure, where the intended structure is designed based on the construct(s) being measured.⁴⁹ For instance, if an instrument is designed to measure a single construct, its internal structure should ideally reflect that instrument items measure only

one dimension. If instrument items instead measure multiple dimensions, data obtained from the instrument cannot be interpreted as measuring the construct of interest.⁴⁹ The NMR-LRC was designed to measure a single construct, lexical representational competence (and associated perceived competence) in ¹H NMR spectroscopy. Establishing unidimensionality of data obtained from the NMR-LRC would thus provide evidence that the instrument measures the construct of interest and that data can be interpreted as an inferential measure of lexical representational competence and perceived competence in ¹H NMR spectroscopy.

Multiple-choice responses on the 10 NMR-LRC items were evaluated for unidimensionality using principal component analysis to determine the correlated variance of items' standardized residuals not explained by the Rasch model. Items with loadings less than ± 0.4 on the second contrast or with associated eigenvalues less than 2.00 would provide evidence for unidimensionality.⁵³ Eight of the ten items exhibited loadings less than ± 0.4 on the second contrast. Two items exhibited loadings of 0.76 and -0.48; however, the eigenvalue of the second contrast was 1.25 and so below the 2.00 criterion, meaning these items were not a threat to unidimensionality. Data obtained from the NMR-LRC therefore met the assumption of unidimensionality for Rasch analysis, and evidence of its unidimensional structure supports its interpretation as a measure of lexical representational competence and perceived lexical representational competence in ¹H NMR spectroscopy. All factor loadings are provided in the Supporting Information.

Local independence of NMR-LRC items. In addition to its assumption of unidimensionality, the Rasch model also assumes the probability of correctly responding to one item is independent of the probability of correctly responding to another item.⁵³ To evaluate whether NMR-LRC items met this assumption of local independence, inter-item correlations were evaluated again using principal component analysis of multiple-choice responses. For questions to be considered locally independent, correlations between items' standardized residuals must be below 0.7.⁵³ All inter-item correlations were below 0.20, providing evidence that items met the Rasch model's assumption of local independence. Data obtained from the NMR-LRC were therefore appropriate for Rasch analysis.

Fit of NMR-LRC items to the Rasch model. The Rasch model calculates individuals' ability and item difficulty measures by assuming that individuals with higher ability will have a greater probability of correctly responding to items than individuals with lower ability. During Rasch analysis, the fit of item responses to the model is evaluated through fit statistics.⁵³ Good item fit to the model indicates that a question functions as intended, where participants with higher ability respond correctly more often than participants with lower ability.⁵⁵ Conversely, poor item fit indicates that a question does not function as intended and that the probability of responding correctly is independent of ability level. Unexpected response patterns are indicators of poor fit, and two residual analyses are used to identify different types of such patterns and thus evaluate fit: outfit and infit. Outfit is an unweighted index sensitive to unexpected response patterns involving large differences in individuals' ability and item difficulty, such as when participants guess the correct response to items with difficulty measures far above their ability measures.⁵³ Conceptually, outfit is used to identify outlier response patterns, and the index can be easily skewed by such observations.⁶⁰ Conversely, infit is a weighted index sensitive to unexpected response patterns involving small differences in individuals' ability and item difficulty measures, where the Rasch model should be able to accurately predict the probability of a correct response. Infit is used to identify inlier response patterns, which often point toward larger issues that threaten response process validity.⁵³ Reliability concerns the reproducibility of assessment measures and addresses the hypothetical question of whether students with identical ability in a latent trait would respond to items similarly, both in the same assessment administration and across timepoints.⁴⁹ Outfit and infit provide statistical evidence of reliability, as acceptable indices indicate that participants' responses are consistent with their ability level and not random.⁵⁶ Outfit and infit indices are provided in the form of a mean-square statistic (MNSQ), both of which are chi-squared statistics divided by degrees of freedom. MNSQ values should fall between 0.5 and 1.5 for questions to be productive for measurement.⁵³ MNSQ values have an associated Z-statistic (ZSTD) that is used to assess statistical significance, though this value can be ignored when MNSQ values are acceptable or when the sample size exceeds 300.⁵³

Table 5.3. Psychometric estimates of NMR-LRC items from Rasch analysis, including infit and outfit statistics and item difficulty measures. Items are ordered from easiest to most difficulty.

Item	Infit	Outfit	Item Difficulty Measure
Multiplicity	0.89	0.73	-2.51
Chemical shift	0.96	0.87	-2.00
The N+1 rule	0.88	0.74	-1.74
Upfield and downfield	1.01	1.10	-0.97
Peak area	0.93	0.87	-0.15
Shielded and de-shielded	0.93	0.93	-0.09
Spin-spin coupling	1.10	1.24	1.01
Topicity	1.02	1.05	1.25
Proton exchange	1.06	1.24	1.52
Coupling constant	1.00	3.29	3.69

All NMR-LRC items had infit statistics within the acceptable range of 0.5 to 1.5, providing statistical evidence of response process validity for all items (Table 5.3). Further, nine of ten NMR-LRC items had outfit statistics within the acceptable range (Table 5.3). The most difficult question, “coupling constant,” exhibited an outfit value above 1.5, suggesting that a subset of participants guessed when responding to this question. This statistic is not particularly problematic, as outfit values are easily skewed by such observations. Overall, nine out of ten NMR-LRC items had acceptable infit and outfit values and thus displayed a good fit to the Rasch model. Fit to the model provides statistical evidence of item response reliability, as acceptable infit and outfit values indicate that participants did not respond in a random, inconsistent fashion.⁵⁶ The unacceptable outfit value for the most difficult question, “coupling constant,” indicates that data obtained from this question should be interpreted with caution. While its infit value suggests no associated response process validity issues, correct responses may be the result of guessing and not a reliable indication of lexical representational competence. The acceptable infit value suggests that while there may have been some guessing, the question was interpreted as intended and thus still provides a measure of perceived lexical representation competence.

Additional statistical evidence of reliability. The Rasch model places student ability measures and item difficulty measures on the same scale in units of logits, facilitating a direct comparison between student abilities and item difficulties during assessment evaluation. This placement is depicted using a Wright map, a plot including student ability measures and item difficulty measures on the same axis (Figure 5.3). When a student's ability measure and an item difficulty measure are equal on a Wright map, the student has a 50% probability of correctly responding to the item.⁵⁵ The student then has greater than a 50% probability of correctly responding to items below their ability level, and they have less than a 50% probability of correctly responding to items above their ability level. The map provides a convenient means of evaluating item targeting, or the degree to which item difficulties match student abilities. Evaluating the match between average participant ability and average item difficulty provides an initial indication of item targeting, where similar values indicate the assessment appropriately targeted the ability of the student population. When assessment items exhibit a spread of difficulties that matches the distribution of abilities in a student population, items are able to provide a more reliable indication of student ability level. Evidence of item targeting thus serves as

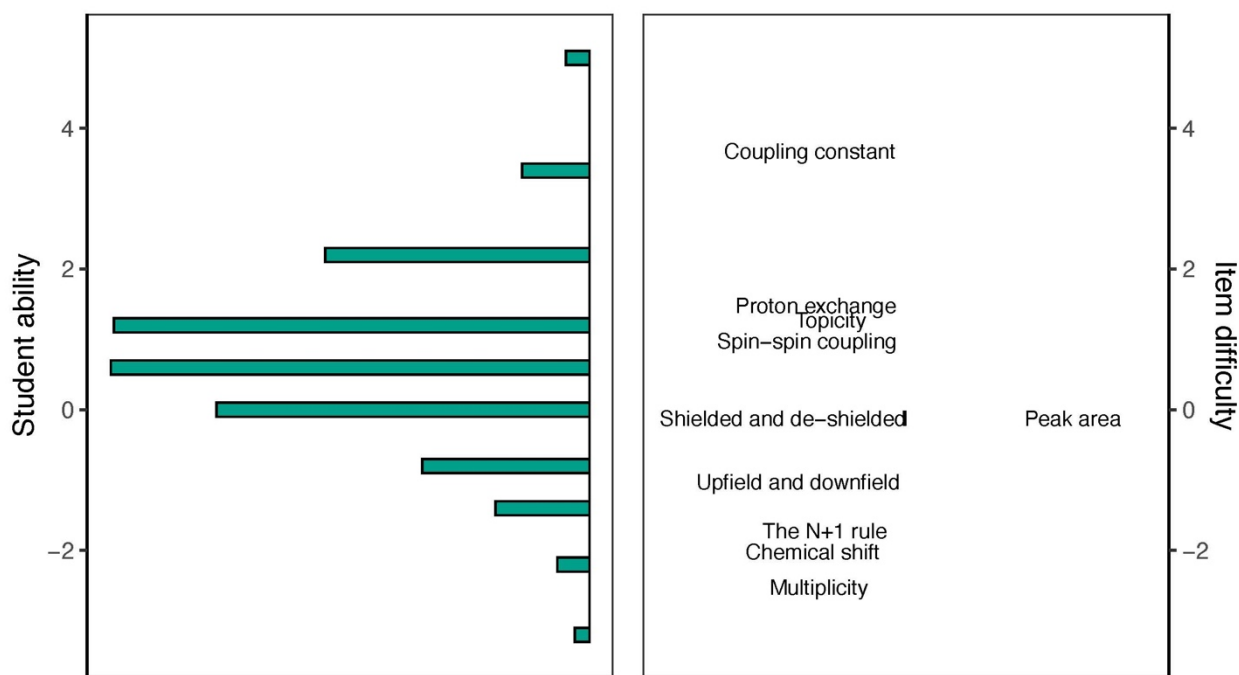


Figure 5.3. Wright map of NMR-LRC participant ability and item difficulty measures on a logit scale. Student ability measures are on the left, where the size of each bar corresponds to the number of participants with a given ability measure. Item difficulty measures are on the right.

statistical evidence of reliability, and it supports the interpretation of data obtained from an instrument as a reliable measure of students' ability.

Item targeting on the NMR-LRC was evaluated by generating a Wright map of participant ability measures and item difficulty measures (Figure 5.3). Average item difficulty was set to zero during Rasch analysis and average participant ability was found to equal 0.67, indicating that the assessment was somewhat easy for the study population though not overly easy or difficult. Further, the spread of item difficulties covers the distribution of student abilities, with all items exhibiting difficulty measures between the highest and lowest student ability measures. Items are also centered around the average item difficulty measure, with four items above zero logits, four items below zero logits, and two items near-zero logits. This spread provides additional evidence of item targeting and thus reliability. A small subset of participants had ability measures above all item difficulty measures, making estimating these ability measures somewhat less reliable. However, this mismatch is only problematic when designing a high-stakes summative assessment to determine precise ability level. The NMR-LRC was designed to be a low-stakes formative assessment, so it is not intended for evaluating students' precise ability level.

The Rasch model provides additional statistical evidence of reliability in the form of the person reliability coefficient. High values of the person reliability coefficient indicate a high probability that individuals with high ability measures actually have higher ability measures than individuals with low ability measures.⁵³ High values thus indicate good "reproducibility of relative measure location."⁵³ This measure provides insight into whether an assessment reliably discriminates a student sample into enough levels for the instrument's intended purpose.⁵³ Coefficients near 0.9 are required to reliably discriminate samples into three or four levels,⁵³ which would be necessary for more high-stakes summative assessments. However, coefficients near 0.5 indicate the sample can still be reliably discriminated into one or two levels,⁵³ where two levels (e.g., low and high) would be sufficient for a low-stakes formative assessment. The person reliability coefficient for the NMR-LRC equaled 0.48, suggesting that the instrument may be suitable for use as a formative assessment to classify each student into either a high or low ability

level. While the coefficient falls below the general reliability coefficient criterion of 0.7, this value is similar to the reliability coefficients of multiple concept inventories in chemistry.^{48,56} Reporting a reliability coefficient is standard in evaluating assessments, though researchers argue that this reliability measure may not be appropriate for concept inventories as students' knowledge tends to be fragmented and not highly consistent.⁶¹ While the NMR-LRC is not a concept inventory, we hypothesize that participants' lexical ability was similarly fragmented and thus resulted in a low reliability coefficient.

The number of person strata was evaluated to provide further insight into whether the instrument discriminates into one or two levels. This value equals the estimated number of statistically distinguishable performance levels when very high and very low scores are likely due to very high and very low ability, respectively.⁵³ The number of person strata is the counterpart of the separation index, which assumes extreme scores are accidental. The number of person strata was estimated as 1.96, indicating that the instrument discriminated the sample into approximately two levels. Collectively, the person reliability coefficient and the number of strata suggest the NMR-LRC is suitable for use as a formative assessment to classify each student into either a high or low ability level. However, the instrument may not be reliably used as a high-stakes summative assessment such as a graded quiz.

Statistical evidence of content validity. In addition to statistical evidence of reliability, the Wright map also provides evidence of content validity in the form of the NMR-LRC's item difficulty hierarchy. The Wright map depicts this item difficulty hierarchy, where the most challenging term for this study population is at the top of the map and the most accessible term is at the bottom (Figure 5.3). Items near the bottom of the map (e.g., "multiplicity," "N+1 rule," and "chemical shift") correspond to terms commonly used during introductory-level instruction on interpreting ¹H NMR spectra, and items near the top of the map (e.g., "coupling constant," "proton exchange," and "topicity") correspond to terms less commonly used during instruction.⁴⁶ The hierarchy serves as evidence of content validity for each item, as items designed to measure one's ability to use common terms should be more accessible than items designed to measure one's ability to use less-common terms. Further, the distribution of terms from accessible to challenging suggests

that the content domain was adequately sampled, providing additional content validity evidence.

5.7.3 Supporting sources of validity evidence

Rasch analysis afforded a range of statistical evidence of validity and reliability, providing support for interpreting the NMR-LRC as a measure of lexical representational competence and perceived competence in ^1H NMR spectroscopy. Additional qualitative sources of content validity and response process validity, as well as a statistical source of associative validity, provide further support for this interpretation.

Qualitative evidence of content validity. A final source of content validity includes a disciplinary and instructional expert's review of the degree to which items assessed lexical ability in ^1H NMR spectroscopy. This qualitative evidence supports the statistical evidence of content validity obtained through Rasch analysis. The expert provided insight into question wording and the extent to which it accurately reflects the content domain. Several small changes to question wording were made to be more in line with verbiage used by practicing organic chemists. However, these changes did not impact the terms selected from widely published instructional materials.

Qualitative evidence of response process validity. Data obtained from cognitive interviews conducted during the instrument development and evaluation phases serve as qualitative evidence of response process validity. This qualitative evidence supports the statistical evidence of response process validity obtained through Rasch analysis. During the instrument development phase, cognitive interviews were used to identify questions with irregular response patterns and question wording resulting in such patterns. From these interviews, a subset of participants correctly responded to the item "shielded and de-shielded" using incorrect reasoning involving steric hindrance. This question was revised to obtain responses more indicative of students' competence. Participants responded to all other questions either correctly by using correct reasoning or incorrectly by using incorrect reasoning, so no additional changes were made. Following the revision, cognitive interviews were conducted again in the instrument

evaluation phase to determine if this change addressed response process issues. No irregular response patterns were identified in the evaluation phase, suggesting that this change was successful.

In addition to identifying response process issues, cognitive interviews were also used to determine if incorrect response options were appealing to students and thus capable of providing insight into the nature of students' limited competence. It should be noted that the NMR-LRC is not a concept inventory, so incorrect response options do not reflect different levels of understanding identified through a body of research. Rather, incorrect response options simply reflect limited competence regarding a set of two related terms (e.g., confusing the terms shielded and de-shielded) or limited competence regarding a single term (e.g., thinking that the chemical shift of a peak refers to its width in ppm rather than its position on a spectrum). Participants often selected incorrect response options using erroneous thinking, suggesting that these options provide such insight and are thus informative for instructors. For instance, a subset of participants incorrectly responded to the item "upfield and downfield" while expressing that they often confuse the direction that each term references. Further, several participants incorrectly responded to the item "proton exchange" by incorrectly reasoning that this term referred to spin-spin coupling, leading them to identify split peaks as those corresponding to hydrogen atoms undergoing proton exchange. Some participants also incorrectly responded to the item "coupling constant" by incorrectly reasoning that this value is measured using the height of peaks and not spacing between peaks, leading them to rationalize that coupling partners would have unequal coupling constants. Therefore, incorrect responses to the NMR-LRC were not necessarily random and instead have the potential to provide instructors with insight into their students' understanding.

Statistical evidence of associative validity. Evidence of associative validity was collected by evaluating the correlation between participants' Rasch ability measures and their cumulative quiz score from the laboratory course in which they were enrolled. Evaluating the correlation between measured variables and similar external variables affords evidence that an instrument functions as intended and measures ability in the domain of interest, as students who performed well on one assessment would be

expected to perform well on other assessments containing similar, related content.⁴⁹ The Spearman's rank-order correlation coefficient (ρ), the nonparametric version of the Pearson-product moment correlation coefficient, was evaluated as the correlation measure given that quiz scores were ordinal. This correlation was both moderate ($\rho=0.30$) and significant ($p<0.0001$),⁶² providing additional support for interpreting data obtained from the NMR-LRC as a measure of lexical ability and perceived lexical ability in ¹H NMR spectroscopy.

5.7.4 Cluster analysis

The cluster analysis of participants' Rasch ability measures and average confidence ratings suggests that the NMR-LRC can be used to identify students who hold illusions of lexical representational competence, i.e., high perceived lexical ability but low ability. This analysis further suggests that the NMR-LRC can be used to identify a range of student groups with characteristic ability and perceived ability level combinations, allowing instructors to provide a range of feedback to promote competence or perceived competence. The cluster solution is presented herein, along with a discussion of the meaning of different clusters in the context of chemistry education.

Cluster solution. The TwoStep cluster analysis of participants' Rasch ability measures and average confidence ratings generated a five cluster solution (Figure 5.4). A description of clusters is provided in Table 5.4. A range of evidence supports the validity and stability of the cluster solution. The solution's interpretability and reasonable number of homogeneous clusters serve as primary evidence of validity.⁵⁹ As will be further discussed below, the solution is interpretable in the context of chemistry education. Clusters can be interpreted as moderate ability/high perceived ability (Cluster 1), low ability/high perceived ability (Cluster 2), high ability/high perceived ability (Cluster 3), moderate ability/moderate perceived ability (Cluster 4), and low ability/low perceived ability (Cluster 5). Cluster 2 indicates the well-documented Dunning-Kruger effect, and Cluster 3 indicates students who are likely prepared to engage in relevant social discourse and learn physical aspects of interpreting spectra. Given that clusters can be ascribed meaning, the number of clusters is not unreasonable. The clusters are also sized

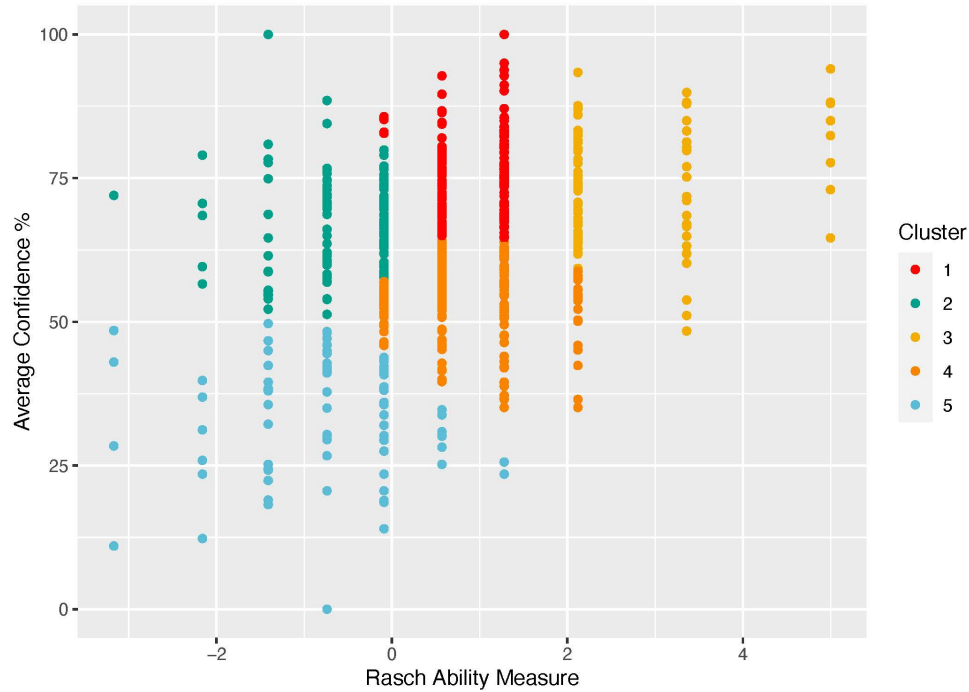


Figure 5.4. Cluster solution of Rasch ability measures and average confidence ratings.

relatively uniformly, with the number of participants in each cluster ranging from 80 to 191. Further, the silhouette coefficient provides evidence that the cluster solution is valid. This measure describes a solution’s cohesion and separation, or the degree to which in-cluster distances are small and between-cluster distances are large, respectively.⁵⁹ The measure provides insight into the degree to which clusters are internally uniform but also distinct from one another. The silhouette coefficient equaled the cutoff criterion of 0.5, suggesting a reasonable degree of cohesion and separation.⁵⁹

Table 5.4. Description of clusters, including cluster interpretation, mean Rasch ability measure, median total score, mean average confidence rating, and cluster size. All values are from one of the seven five-cluster solutions. This solution was selected randomly, and all values across solutions were highly similar.

Cluster	Cluster interpretation	Mean Rasch ability measure	Median Total Score	Mean average confidence rating	Cluster size (N)
1	Moderate ability/high perceived ability	0.92	7	75.2%	179 (26.4%)
2	Low ability/high perceived ability (i.e., Dunning-Kruger effect)	-0.54	5	66.6%	131 (19.3%)
3	High ability/high perceived ability	2.65	8	73.6%	97 (14.3%)
4	Moderate ability/moderate perceived ability	0.92	6	54.3%	191 (28.2%)
5	Low ability/low perceived ability	-0.73	4	33.7%	80 (11.8%)

The TwoStep clustering algorithm is sensitive to the ordering of data, so to evaluate the stability of the solution the cluster analysis was completed ten times with participants' responses being randomized with each iteration.⁵⁹ The five-cluster solution was obtained seven out of ten times, where mean Rasch ability measures and average confidence ratings remained highly similar across solutions. A four-cluster solution was obtained two out of ten times, and a six-cluster solution was obtained one out of ten times. The five-cluster solution was therefore considered stable.⁴⁷ This procedure for establishing the stability of a cluster solution is similar to that used in another chemistry education study measuring the Dunning-Kruger effect.⁴⁷

Cluster corresponding to illusions of competence. Cluster 2 includes responses reflecting low lexical representational competence in ¹H NMR spectroscopy but high perceived competence. This cluster thus corresponds to students with illusions of such competence, and it serves as evidence that the NMR-LRC can be used to identify students exhibiting the Dunning-Kruger effect. Moreover, this cluster also provides evidence that the Dunning-Kruger effect extends to lexical ability in ¹H NMR spectroscopy. The cluster's moderate size further suggests that this phenomenon is prevalent, with approximately 19% of the study population exhibiting the effect (Table 5.4). Evidence of this effect in ¹H NMR spectroscopy, combined with its apparent prevalence, suggests that instructors should expect a proportion of their students to exhibit illusions of lexical representational competence and prepare instruction that effectively promotes learning among these individuals. These results also highlight the importance of using a formative assessment that can identify students holding illusions of lexical representational competence, providing support for using the NMR-LRC as a formative assessment in introductory organic chemistry.

Interpretation of other clusters. While cluster analysis was conducted to determine if the NMR-LRC could detect illusions of competence, the other identified clusters also provide useful insight into the instrument's functioning. These results suggest that the NMR-LRC can be used to identify a range of student groups with characteristic ability and perceived ability level combinations. Cluster 3 included responses reflecting high ability and high perceived ability, suggesting that the NMR-LRC

can be used to identify students who are prepared to engage in social discourse and learn physical aspects of interpreting spectra. Also, Cluster 5 includes responses reflecting low ability and low perceived ability; these students demonstrate limited ability, but they would likely respond to traditional forms of feedback (e.g., low assessment scores or written feedback about terms they may be unfamiliar with) given that their perceived ability is well-calibrated. Clusters 1 and 4 include responses reflecting moderate lexical ability with high and moderate perceived ability, respectively. While these clusters suggest that a subset of students are in an intermediate, transitional phase toward the expert-like rhetorical usage of spectra, Rasch analysis demonstrated that the NMR-LRC can only reliably discriminate students into approximately two levels (i.e., low versus high). These clusters should therefore be interpreted with caution. Lastly, a cluster of responses reflecting high ability and low perceived ability was not identified, though it is still possible that the NMR-LRC could be used to identify such responses among other student populations. Students with such responses would likely benefit from feedback in the form of encouragement or high assessment scores to promote their self-efficacy.⁴⁴

5.8 Limitations

Noteworthy limitations of this investigation relate to the study population. Like all scale development studies, evidence of validity and reliability supporting the interpretation of assessment data is specific to the study population. Participants in this study were recruited from a single institution, so the degree to which validity and reliability evidence supports the interpretation of NMR-LRC data in other instructional contexts is uncertain. However, several factors suggest that NMR-LRC data can be interpreted as a measure of lexical representational competence and perceived competence in other instructional contexts. Participants in the evaluation phase of the study were taught by one of two instructors who independently selected the terminology they included in their instruction, so the study population did contain a degree of variability regarding instructional context. Also, instruction in organic chemistry is remarkably consistent and uniform across institutions at the national level, unlike instruction in fields such as inorganic chemistry where content coverage varies considerably.⁶³ Terms included in the NMR-LRC were also selected from widely-published instructional materials, and questions were designed

to assess general ability rather than advanced understanding. The uniformity of organic chemistry instruction, combined with these item design elements, suggests that the NMR-LRC's content would align with the content presented in most introductory organic chemistry courses covering ^1H NMR spectral interpretation. Ultimately, however, additional studies are needed to determine the degree to which validity and reliability evidence supports NMR-LRC data interpretation in other instructional contexts. Further, the majority of participants in this study identified as white. The degree to which validity and reliability evidence would support the interpretation of NMR-LRC data collected from more diverse student populations is therefore also uncertain. Cognitive interviews with participants who identified as nonwhite provide some evidence for response process validity in more diverse contexts, though this evidence is minimal. Additional studies are therefore also needed to evaluate if the NMR-LRC functions as intended in diverse contexts where more students identify as nonwhite or speak English as an additional language. We welcome additional studies that provide evidence of validity and reliability with different student samples; if you would like a copy of the NMR-LRC to use in your classroom, please contact the corresponding author. Instructors seeking to use the NMR-LRC prior to additional evaluation studies are also welcomed to contact the corresponding author for a copy; however, these instructors should evaluate the degree to which their students' responses may be informative before administering the assessment, as well as use the obtained data in a low-stakes manner.

5.9 Conclusions and implications

This article describes the development and evaluation of the NMR-LRC, a 10-item formative assessment of organic chemistry students' lexical ability and perceived lexical ability in ^1H NMR spectroscopy. Development of the NMR-LRC involved the use of widely published instructional materials and expert review of assessment content to ensure that the instrument adequately represented the domain of interest (i.e., content validity). Development also involved the use of cognitive interviews to help ensure questions were interpreted as intended (i.e., response process validity). Evaluation of the NMR-LRC involved psychometric analysis using the Rasch model, during which we collected statistical evidence that the instrument measured lexical ability and perceived lexical

ability in a valid and reliable manner. Rasch analysis afforded statistical evidence that the NMR-LRC measures a single dimension (i.e., structural validity), that items measure lexical ability for a range of distinct terms (i.e., content validity), that response patterns were consistent with participants' ability level and that items were targeted to measure low and high ability levels (i.e., reliability), and that there were no inlier response patterns indicating issues with question interpretation (i.e., response process validity). To support this evidence of validity, we also established a correlation between NMR-LRC ability measures and course quiz scores (i.e., associative validity). This collective evidence supports the interpretation of data obtained from the NMR-LRC as a measure of students' lexical ability and perceived lexical ability in ^1H NMR spectroscopy, and it supports its use as a formative assessment in introductory organic chemistry courses.

Results from cluster analysis of Rasch ability measures and confidence ratings suggest that the NMR-LRC can be used to identify students with both competence and perceived competence who are likely prepared to engage in social discourse and learn the physical aspects of interpreting spectra. In turn, it may also be used to identify students who are likely unprepared and require additional instructor feedback. Cluster analysis also demonstrated that the NMR-LRC can be used to identify students holding illusions of lexical representational competence, or high perceived competence but low competence. These are students who likely require alternate forms of feedback, which are discussed in the implications section. These results further support the use of the NMR-LRC as a formative assessment to help instructors cultivate students' practical understanding and expert-like, rhetorical usage of spectra. Instructors seeking to use the NMR-LRC in their classroom are welcomed to contact the corresponding author for a copy.

5.9.1 Implications for research

The NMR-LRC is the first published psychometrically-evaluated instrument focusing on NMR spectroscopy, particularly the use of ^1H NMR spectra as tools for social discourse. However, NMR spectra play an essential role not only in supporting chemists' social discourse and but also representing submicroscopic entities.² Psychometrically-

evaluated assessments focused on NMR spectra as submicroscopic representations are thus also needed. Existing studies provide a theoretical foundation for designing assessments to measure conceptual understanding of underlying physical principles and conceptual understanding and reasoning involved in relating molecular structures to spectra.^{14,64} The design of such instruments is an important next step toward building a repository of assessments focused on this essential technique. Such a repository will help instructors cultivate students' ability to interpret NMR spectra and use these representations for discourse. Further, the NMR-LRC was designed to be a formative assessment. Summative assessments on NMR spectroscopy are needed in addition to low-stakes formative assessments to assist these instructors further. Lastly, evidence of validity and reliability for the NMR-LRC was gathered at a single institution. While the organic chemistry curriculum is remarkably consistent across institutions and widely-published curricular materials were used to develop the NMR-LRC, additional evaluation studies are needed to ensure that the instrument functions as intended in other contexts.

Beyond the development and evaluation of assessments, this investigation provides important implications for the design of future chemistry education studies focusing on representational competence. The unidimensional nature of the NMR-LRC provides empirical support for the notion that lexical representational competence is a subconstruct of representational competence; the ability to use words to communicate the identification, analysis, and interpretation of features within representations is thus a distinct component of representational competence that merits attention. Future studies should thus intentionally consider this subconstruct as they seek to understand individuals' representational competence overall.

5.9.2 Implications for teaching

Results from this study provide several implications for ^1H NMR spectroscopy instruction. Most notably, psychometric evidence supports that lexical representational competence in ^1H NMR spectroscopy is a distinct component of one's overall representational competence. Therefore, instruction should focus in part on cultivating students' understanding of NMR-specific terminology and their ability to use these words to

communicate. Researchers have identified a range of literacy-based strategies for chemistry instruction which may help cultivate such familiarity.³⁰ Learning also occurs when individuals participate in a social practice with their peers and more competent individuals,¹⁶ suggesting that undergraduates may benefit from communicating their identification, analysis, and interpretation of NMR spectral features in small groups containing lab mates and an instructor or teaching assistant. Results from cluster analysis further suggest that students can hold illusions of lexical representational competence in ¹H NMR spectroscopy, meaning instructors should anticipate that a subset of their students will exhibit the Dunning-Krueger effect and not respond to traditional forms of feedback. Effective instruction in this topic will therefore require that students with illusions of competence receive alternate forms of feedback to calibrate their perceived lexical ability. The importance of lexical representational competence in ¹H NMR spectroscopy, combined with the possibility for illusions of competence, further indicates that effective instruction will require formatively assessing students' lexical ability and perceived lexical ability. The NMR-LRC serves as a psychometrically-evaluated tool for such an assessment.

Alternate forms of feedback for students holding illusions of competence may take on many forms. Benchmarking, or evaluating one's performance by comparing it to that of others, is one promising approach for calibrating perceived ability.^{40,65} For instance, research demonstrates that students' proofreading abilities improve after reviewing other students' proofreading edits.⁶⁵ Students holding illusions of lexical representational competence in NMR spectroscopy would thus likely benefit from group work in which students verbally describe spectra. An instructor could first use the NMR-LRC to identify three to four students with varying levels of ability and perceived ability, followed by grouping these individuals for a verbal task. Instructors could also incorporate writing prompts that require students to describe spectra. Students could then engage in peer review, providing an opportunity for students with illusions of competence to calibrate their perceived ability. A number of studies have investigated writing and peer review in chemistry classrooms, providing insight into the effective design of such an activity.⁶⁶⁻⁶⁹ A subset of these studies are also situated in organic chemistry classrooms.^{68,69} Further,

one study demonstrates how writing and peer review was implemented in an introductory biology course to identify and address students' misconceptions, i.e., deficits in knowledge marked by low understanding but high perceived understanding.^{47,70} Writing and peer review may be particularly suited to address students' illusions of lexical representational competence in NMR spectroscopy.

Lastly, evidence from Rasch and cluster analysis evidence supports the interpretation of data obtained from the NMR-LRC and its use as a formative assessment. However, instructors seeking to use the NMR-LRC do not need to conduct Rasch or cluster analysis to infer meaning from their obtained data. Ideally, students should demonstrate an ability to use all terms on the NMR-LRC (i.e., receive a score of 10), as these terms appeared in multiple widely-published instructional materials. For students scoring below 10, the item difficulty hierarchy presented herein provides instructors with a means of determining students' general ability level, and in turn the amount of additional instruction they will require to receive a score of 10 (Figure 3). Psychometric evidence suggests that the NMR-LRC can discriminate students into two ability levels (i.e., low and high). Students who mainly provide correct responses to the four items well below the average item difficulty of zero could thus be considered lower ability and requiring more instruction, and students correctly responding to these items in addition to items above zero could be considered higher ability and requiring less instruction (Figure 3, Table 3). In addition to using terms, students should also ideally report confidence ratings closer to 100% than 0% (i.e., above 50%). Using these general guidelines, instructors can classify students into four general groups having high or low ability and perceived ability. Students with responses near an intermediate level (i.e., correct responses to the four items below zero and two items near zero, with confidence ratings near 50%) could then be grouped at the instructor's discretion.

5.10 Acknowledgements

The authors would like to acknowledge students and the instructional and disciplinary expert who participated in this study, as well as Eleni Zotos for her help with coordinating data collection.

5.11 Supporting Information

Table 5.5. Principal component analysis factor loadings of NMR-LRC items on the second contrast.

Item	Loading
Multiplicity	-0.20
Chemical shift	0.31
The N+1 rule	-0.39
Upfield and downfield	0.15
Peak area	-0.48
Shielded and de-shielded	0.16
Spin-spin coupling	-0.14
Topicity	0.76
Proton exchange	-0.32
Coupling constant	0.00

5.12 References

- (1) Schummer, J. The Impact of Instrumentation on Chemical Species Identity: From Chemical Substances to Molecular Species. In *From Classical to Modern Chemistry: The Instrumental Revolution*; Morris, P., Ed.; The Royal Society of Chemistry: Cambridge, 2002; pp 188–211.
- (2) Kozma, R.; Russell, J. Modelling Students Becoming Chemists: Developing Representational Competence. In *Visualization in Science Education*; Gilbert, J. K., Ed.; Springer: Dordrecht, 2005; pp 121–146. https://doi.org/10.1007/1-4020-3613-2_4.
- (3) Parmentier, L. E.; Lisensky, G. C.; Spencer, B. A Guided Inquiry Approach to NMR Spectroscopy. *J. Chem. Educ.* **1998**, *75* (4), 470–471. <https://doi.org/10.1021/ed075p470>.
- (4) Debska, B.; Guzowska-Swider, B. Molecular Structure from ¹H NMR Spectra: Education Aided by Internet Programs. *J. Chem. Educ.* **2007**, *84* (3), 556–560. <https://doi.org/10.1021/ed084p556>.
- (5) Veeraraghavan, S. NMR Spectroscopy and Its Value: A Primer. *J. Chem. Educ.* **2008**, *85* (4), 537–540. <https://doi.org/10.1021/ed085p537>.
- (6) Livengood, K.; Lewallen, D. W.; Leatherman, J.; Maxwell, J. L. The Use and Evaluation of Scaffolding, Student Centered-Learning, Behaviorism, and Constructivism to Teach Nuclear Magnetic Resonance and IR Spectroscopy in a Two-Semester Organic Chemistry Course. *J. Chem. Educ.* **2012**, *89* (8), 1001–1006. <https://doi.org/10.1021/ed200638g>.

- (7) Azman, A. M.; Esteb, J. J. A Coin-Flipping Analogy and Web App for Teaching Spin-Spin Splitting in ^1H NMR Spectroscopy. *J. Chem. Educ.* **2016**, *93* (8), 1478–1482. <https://doi.org/10.1021/acs.jchemed.6b00133>.
- (8) Erhart, S. E.; McCarrick, R. M.; Lorigan, G. A.; Yeziarski, E. J. Citrus Quality Control: An NMR/MRI Problem-Based Experiment. *J. Chem. Educ.* **2016**, *93* (2), 335–339. <https://doi.org/10.1021/acs.jchemed.5b00251>.
- (9) Graham, K. J.; McIntee, E. J.; Schaller, C. P. Web-Based 2D NMR Spectroscopy Practice Problems. *J. Chem. Educ.* **2016**, *93* (8), 1483–1485. <https://doi.org/10.1021/acs.jchemed.6b00007>.
- (10) Vosegaard, T. ISpec: A Web-Based Activity for Spectroscopy Teaching. *J. Chem. Educ.* **2018**, *95* (1), 97–103. <https://doi.org/10.1021/acs.jchemed.7b00482>.
- (11) Cartrette, D. P.; Bodner, G. M. Non-Mathematical Problem Solving in Organic Chemistry. *J. Res. Sci. Teach.* **2010**, *47* (6), 643–660. <https://doi.org/10.1002/tea.20306>.
- (12) Topczewski, J. J.; Topczewski, A. M.; Tang, H.; Kendhammer, L. K.; Pienta, N. J. NMR Spectra through the Eyes of a Student: Eye Tracking Applied to NMR Items. *J. Chem. Educ.* **2017**, *94* (1), 29–37. <https://doi.org/10.1021/acs.jchemed.6b00528>.
- (13) Connor, M. C.; Shultz, G. V. Teaching Assistants' Topic-Specific Pedagogical Content Knowledge in ^1H NMR Spectroscopy. *Chem. Educ. Res. Pract.* **2018**, *19* (3), 653–669. <https://doi.org/10.1039/c7rp00204a>.
- (14) 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>.
- (15) Collins, H. Language and Practice. **2011**, *41* (2), 271–300. <https://doi.org/10.1177/03063127111399665>.
- (16) Bowen, G. M.; Roth, W. M.; McGinn, M. K. Interpretations of Graphs by University Biology Students and Practicing Scientists: Toward a Social Practice View of Scientific Representation Practices. *J. Res. Sci. Teach.* **1999**, *36* (9), 1020–1043. [https://doi.org/10.1002/\(SICI\)1098-2736\(199911\)36:9<1020::AID-TEA4>3.0.CO;2-#](https://doi.org/10.1002/(SICI)1098-2736(199911)36:9<1020::AID-TEA4>3.0.CO;2-#).
- (17) Denning, P.; Dargan, P. Action-Centered Design. In *Bringing design to software*; Winograd, T., Ed.; 1996; pp 105–120.
- (18) Lavery, J. T.; Underwood, S. M.; Matz, R. L.; Posey, L. A.; Carmel, J. H.; Caballero, M. D.; Fata-Hartley, C. L.; Ebert-May, D.; Jardeleza, S. E.; Cooper, M. M. Characterizing College Science Assessments: The Three-Dimensional Learning Assessment Protocol. *PLoS One* **2016**, *11* (9), 1–21. <https://doi.org/10.1371/journal.pone.0162333>.

- (19) National Research Council. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*; 2012.
- (20) Sadler, D. R. Formative Assessment: Revisiting the Territory. *Assess. Educ.* **1998**, 5 (1), 77–84.
- (21) Vygotsky, L. *Thought and Language*; MIT Press: Boston, 1986.
- (22) Lave, J.; Wenger, E. *Situated Learning*; Cambridge University Press: New York, 1991.
- (23) 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>.
- (24) 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>.
- (25) Graulich, N. The Tip of the Iceberg in Organic Chemistry Classes: How Do Students Deal with the Invisible? *Chemistry Education Research and Practice*. Royal Society of Chemistry January 1, 2015, pp 9–21. <https://doi.org/10.1039/c4rp00165f>.
- (26) Flynn, A. B.; Featherstone, R. B. Language of Mechanisms: Exam Analysis Reveals Students' Strengths, Strategies, and Errors When Using the Electron-Pushing Formalism (Curved Arrows) in New Reactions. *Chem. Educ. Res. Pract.* **2017**, 18 (1), 64–77. <https://doi.org/10.1039/c6rp00126b>.
- (27) Erduran, S. *Methodological Foundations in the Study of Argumentation in Science Classrooms*; Springer, Dordrecht, 2007; pp 47–69. https://doi.org/10.1007/978-1-4020-6670-2_3.
- (28) Gelman, S. A. Learning from Others: Children's Construction of Concepts. *Annu. Rev. Psychol.* **2009**, 60, 115–140. <https://doi.org/10.1146/annurev.psych.59.103006.093659>.
- (29) Pyburn, D. T.; Pazicni, S.; Benassi, V. A.; Tappin, E. E. Assessing the Relation between Language Comprehension and Performance in General Chemistry. *Chem. Educ. Res. Pract.* **2013**, 14 (4), 524–541. <https://doi.org/10.1039/c3rp00014a>.
- (30) Rees, S. W.; Kind, V.; Newton, D. Can Language Focussed Activities Improve Understanding of Chemical Language in Non-Traditional Students? *Chem. Educ. Res. Pract.* **2018**, 19 (3), 755–766. <https://doi.org/10.1039/c8rp00070k>.
- (31) Cassels, J. R. T.; Johnstone, A. H. *Words That Matter in Science*; Royal Society of Chemistry: London, 1985.

- (32) Jasien, P. G. What Do You Mean That “Strong” Doesn’t Mean “Powerful”? *J. Chem. Educ.* **2011**, *88* (9), 1247–1249. <https://doi.org/10.1021/ed100744x>.
- (33) Bandura, A. *Self-Efficacy: The Exercise of Control*; W. H. Freeman and Company: New York, 1997.
- (34) Kan, A.; Akbaş, A. *Affective Factors That Influence Chemistry Achievement (Attitude and Self Efficacy) and The Power Of These Factors To Predict Chemistry Achievement-I*; 2006; Vol. 3.
- (35) Zusho, A.; Pintrich, P. R.; Coppola, B. Skill and Will: The Role of Motivation and Cognition in the Learning of College Chemistry. *Int. J. Sci. Educ.* **2003**, *25* (9), 1081–1094. <https://doi.org/10.1080/0950069032000052207>.
- (36) Kruger, J.; Dunning, D. Unskilled and Unaware of It: How Difficulties in Recognizing One’s Own Incompetence Lead to Inflated Self-Assessments. *J. Pers. Soc. Psychol.* **1999**, *77* (6), 1121–1134. <https://doi.org/10.1037/0022-3514.77.6.1121>.
- (37) Pazicni, S.; Bauer, C. F. Characterizing Illusions of Competence in Introductory Chemistry Students. *Chem. Educ. Res. Pract.* **2014**, *15* (1), 24–34. <https://doi.org/10.1039/c3rp00106g>.
- (38) Ehrlinger, J.; Johnson, K.; Banner, M.; Dunning, D.; Kruger, J. Why the Unskilled Are Unaware: Further Explorations of (Absent) Self-Insight among the Incompetent. *Organ. Behav. Hum. Decis. Process.* **2008**, *105* (1), 98–121. <https://doi.org/10.1016/j.obhdp.2007.05.002>.
- (39) Bol, L.; Hacker, D. J.; Walck, C. C.; Nunnery, J. A. The Effects of Individual or Group Guidelines on the Calibration Accuracy and Achievement of High School Biology Students. *Contemp. Educ. Psychol.* **2012**, *37* (4), 280–287. <https://doi.org/10.1016/j.cedpsych.2012.02.004>.
- (40) Dunning, D.; Heath, C.; Suls, J. M. *Flawed Self-Assessment Implications for Health, Education, and the Workplace*; 2004; Vol. 5.
- (41) Collins, H. Three Dimensions of Expertise. *Phenom. Cogn. Sci.* **2013**, *12*, 253–273. <https://doi.org/10.1007/s11097-011-9203-5>.
- (42) Collins, H. Studies of Expertise and Experience. *Topoi* **2018**, *37* (1), 67–77. <https://doi.org/10.1007/s11245-016-9412-1>.
- (43) Gerothanassis, I. P.; Troganis, A.; Exarchou, V.; Barbarossou, K. Theme Issue 2 on Structural Concepts Contribution from Science (Invited Review Paper) Nuclear Magnetic Resonance (NMR) Spectroscopy: Basic Principles and Phenomena, and Their Applications to Chemistry, Biology, and Medicine. *Chem. Educ. Res. Pract. Eur. Educ. Res. Pr. Eur* **2002**, *3* (3), 229–252.

- (44) Villafañe, S. M.; Xu, X.; Raker, J. R. Self-Efficacy and Academic Performance in First-Semester Organic Chemistry: Testing a Model of Reciprocal Causation. *Chem. Educ. Res. Pract.* **2016**, *17* (4), 973–984. <https://doi.org/10.1039/c6rp00119j>.
- (45) Ege, S. *Organic Chemistry: Structure and Reactivity*, 5th ed.; Staudt, M., Stranz, M., Eds.; Cengage Learning: Mason, 2009.
- (46) Bruice, P. Y. *Organic Chemistry*, 6th ed.; Prentice Hall: Upper Saddle River, 2011.
- (47) Brandriet, A. R.; Bretz, S. L. Measuring Meta-Ignorance through the Lens of Confidence: Examining Students' Redox Misconceptions about Oxidation Numbers, Charge, and Electron Transfer. *Chem. Educ. Res. Pract.* **2014**, *15* (4), 729–746. <https://doi.org/10.1039/c4rp00129j>.
- (48) McClary, L. M.; Bretz, S. L. Development and Assessment of A Diagnostic Tool to Identify Organic Chemistry Students' Alternative Conceptions Related to Acid Strength. *Int. J. Sci. Educ.* **2012**, *34* (15), 2317–2341. <https://doi.org/10.1080/09500693.2012.684433>.
- (49) Arjoon, J. A.; Xu, X.; Lewis, J. E. Understanding the State of the Art for Measurement in Chemistry Education Research: Examining the Psychometric Evidence. *J. Chem. Educ.* **2013**, *90* (5), 536–545. <https://doi.org/10.1021/ed3002013>.
- (50) Collins, D. Pretesting Survey Instruments: An Overview of Cognitive Methods. *Qual. Life Res.* **2003**, *12* (3), 229–238. <https://doi.org/10.1023/A:1023254226592>.
- (51) Cohen, L.; Manion, L.; Morrison, K. *Research Methods in Education*, 7th ed.; Routledge: London, 2011.
- (52) R Core Team. R: A Language and Environment for Statistical Computing. Vienna 2019.
- (53) Linacre, J. M. A User's Guide to Winsteps.
- (54) IBM Corp. IBM SPSS Statistics. IBM Corp: Armonk, NY 2017.
- (55) Bond, Trevor G. Fox, C. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*, 2nd ed.; New York, 2007.
- (56) Wren, D.; Barbera, J. Psychometric Analysis of the Thermochemistry Concept Inventory. *Chem. Educ. Res. Pract.* **2014**, *15* (3), 380–390. <https://doi.org/10.1039/c3rp00170a>.
- (57) Barbera, J. A Psychometric Analysis of the Chemical Concepts Inventory. *J. Chem. Educ.* **2013**, *90* (5), 546–553. <https://doi.org/10.1021/ed3004353>.

- (58) Han, J.; Kamber, M.; Pei, J. *Data Mining: Concepts and Techniques*, 3rd ed.; Elsevier: Amsterdam, 2012. <https://doi.org/10.1109/ICMIRA.2013.45>.
- (59) Norusis, M. J. Cluster Analysis. In *IBM SPSS Statistics 19 Statistical Procedures Companion*; Prentice Hall, 2011; pp 375–404.
- (60) Wren, D.; Barbera, J. Psychometric Analysis of the Thermochemistry Concept Inventory. *Chem. Educ. Res. Pr.* **2014**, *15* (3), 380–390. <https://doi.org/10.1039/C3RP00170A>.
- (61) Adams, W. K.; Wieman, C. E. Development and Validation of Instruments to Measure Learning of Expert-Like Thinking. *Int. J. Sci. Educ.* **2011**, *33* (9), 1289–1312. <https://doi.org/10.1080/09500693.2010.512369>.
- (62) Ravid, R. *Practical Statistics for Educators*, 4th ed.; Rowman & Littlefield Publishers: Lanham, 2011.
- (63) Raker, J. R.; Reisner, B. A.; Smith, S. R.; Stewart, J. L.; Crane, J. L.; Pesterfield, L.; Sobel, S. G. Foundation Coursework in Undergraduate Inorganic Chemistry: Results from a National Survey of Inorganic Chemistry Faculty. *J. Chem. Educ.* **2015**, *92* (6), 973–979. <https://doi.org/10.1021/ed500624t>.
- (64) Domin, D.; Bodner, G. Using Students' Representations Constructed during Problem Solving to Infer Conceptual Understanding. *J. Chem. Educ.* **2012**, *89* (7), 837–843. <https://doi.org/10.1021/ed1006037>.
- (65) Farh, J. L.; Dobbins, G. H. Effects of Comparative Performance Information on the Accuracy of Self-Ratings and Agreement Between Self- and Supervisor Ratings. *J. Appl. Psychol.* **1989**, *74* (4), 606–610. <https://doi.org/10.1037/0021-9010.74.4.606>.
- (66) 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>.
- (67) 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* (3), 922–939. <https://doi.org/10.1039/c9rp00292h>.
- (68) 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>.

(69) 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>.

(70) 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), ar28. <https://doi.org/10.1187/cbe.17-10-0212>.

Chapter 6

Closing Remarks

This body of work constitutes a significant portion of the initial chemistry education research on teaching and learning ^1H NMR spectroscopy. From this research, we see that undergraduates initially learning to interpret ^1H NMR spectra face significant cognitive barriers to cultivating ability in this practice. These barriers most often arise when students first over-extend learned principles (e.g., the N+1 rule) to contexts where experimental and implicit chemical variables must be evaluated to accurately interpret spectral data. Implicit chemical variables refer to information which is not directly observable but instead must be inferred from a molecular structure (e.g., the exchangeable nature of hydroxyl protons). Undergraduates then subsequently rely on these overgeneralizations to inaccurately make decisions using just one spectral feature conflicting with their limited understanding. This constrained thinking must be met with evidence-based instruction that will shift novices to analytical thinking, several examples of which are provided in Chapter 2.

Further, development of this ability continues to prove challenging for undergraduates even once they surpass these barriers. When interpreting spectra, undergraduates tend to rely on less sophisticated conceptual understanding regarding experimental and implicit chemical variables affecting splitting and chemical shift. They also lack familiarity with characteristic chemical shift values. Their limited understanding and familiarity then appear to inhibit the selective and efficient processing of spectral data. When compared to doctoral chemistry students, undergraduates allocate significantly more visual attention to spectral data requiring the evaluation of underlying variables affecting splitting and chemical shift, potentially trying to make sense of spectral data conflicting with their limited understanding. They also allocate significantly more visual attention to reference material, as well as transition more between reference material and

spectra to refresh their working memory. Undergraduates then appear to search within spectra for particular signals, potentially after identifying an expected spectral feature following this extensive processing of unfamiliar reference material. Conversely, doctoral students appear to rely on their more sophisticated understanding and familiarity with reference values to efficiently process this information and move on to other task-relevant features. Doctoral students not only allocate less visual attention to features requiring the evaluation of underlying variables and transition less between spectra and reference material, but they also transition more directly from molecular structures to spectra and from the ^1H NMR spectrum to the complementary IR spectrum. These later transitions suggest an informed interpretation approach, in which doctoral students first look to molecular structures to predict spectral features and then look to spectra to confirm their presence, as well as look across spectra for complementary pieces of data. Doctoral students thus likely appear to have knowledge of characteristic chemical shift values or the relative location that peaks will appear on a spectrum (e.g., further downfield versus further upfield), or both. Interview data suggests that this knowledge is a combination of both; doctoral participants would sometimes first observe the molecule and then express the chemical shift regions in which they would expect signals (e.g., expressing that signals corresponding to aromatic protons should appear around 7 ppm), and at other times they would simply evaluate relative peak location. In their occasional use of reference material, doctoral participants would then either confirm predicted characteristic regions or determine unfamiliar values.

From these findings, we primarily see that providing undergraduates who are initially learning this practice with entire spectra, general guidelines for interpretation (e.g., the N+1 rule), and chemical shift reference material does not adequately support learning. These students are simply working to process all information in spectra and reference material, let alone process this information in an informed and efficient manner. They will likely develop relevant ability through such practice, but as previous research suggests, this development may take years of graduate-level study. Further, the tasks that undergraduate participants completed were relatively accessible compared to many problems presented in introductory organic chemistry textbooks, as those often require

complete structural elucidation using ^1H NMR spectra. Curricular materials may thus be ill-equipped to support development. In addition, providing initial learners with such tasks may not only fail to support learning but also propagate constrained thinking; when completing a task for which they lack relevant understanding and familiarity, it is unsurprising that students would base their decisions on single pieces of data.

Findings on developing the ability to interpret ^1H NMR spectra have considerable implications for the design of instruction. Collectively, these findings primarily suggest that a transformation of undergraduate-level instruction on ^1H NMR spectroscopy is needed. In this transformed instruction, initial learners should first develop conceptual understanding regarding experimental and implicit chemical variables across the five identified areas before they are presented with full spectra. Individual spectral features could be used to illustrate the effect of an underlying variable on spectral appearance, but tasking these students with interpreting entire spectra will not effectively support learning. Further, familiarity with characteristic chemical shift regions seems to be an essential component of expertise; however, providing initial learners with a chemical shift reference table is not a productive supplement for this lack of familiarity. Providing this reference material does not seem to support informed interpretation strategies, but rather results in extensive processing and subsequent searching behavior. Introductory instruction may thus need to cultivate conceptual understanding of relative peak location, with reference material being introduced in more advanced courses. Once initial learners develop understanding across the five areas, they should then develop the ability to evaluate and weigh a small number of variables (e.g., two to three) as they make decisions using spectral data. Full spectra could then be introduced after students develop this ability, with instruction encouraging the informed interpretation approaches identified in Chapter 3. Graduate-level instruction could then further cultivate students' ability to use informed interpretation approaches, as well as cultivate familiarity with characteristic chemical shift values relevant to compounds these students regularly synthesize in their research laboratories. In order to accomplish this transformation, chemistry educators may likely need to redefine their desired learning outcomes for introductory-level instruction on this technique. For instance, it may be possible that students complete introductory organic

chemistry courses without interpreting full spectra. However, their strong conceptual understanding and ability to weigh relevant variables will ultimately expedite their development of expertise as they progress in the curriculum. Then when eventually entering the organic chemistry community, newest members will be capable of using this essential technique.

Accomplishing this transformation will unfortunately require a comprehensive redesign of undergraduate curricular materials. However, findings from this research point toward several immediate changes that may have a considerable impact on student learning. One of the most potentially contentious but impactful of these changes will involve renaming the N+1 rule. Undergraduates in our investigations regularly overextended this principle without considering underlying variables and made decisions using single features conflicting with this principle. This rule-based thinking may in part be so prevalent because this guideline for determining splitting is termed a “rule.” This principle could easily be termed the “N+1 guideline” to account for and convey its flexible nature. From our study on developing expertise in interpreting ^1H NMR spectra, we see that intermediate levels of conceptual understanding focus on context and not underlying variables; instruction that shifts students away from simply evaluating context will thus also be essential for supporting learning. This renaming would ideally encourage students to evaluate implicit variables affecting splitting, helping them not only stop overextending this principle but also move beyond simply considering context and viewing certain contexts as “exceptions” to a rule.

To further encourage the evaluation of implicit variables and shift students from focusing simply on context, curricular materials should organize content around underlying variables rather than explicit molecular features. For instance, widely-published textbooks often have subsections titled “protons attached to nitrogen and oxygen.” Such titles may encourage students to construct knowledge that is organized around explicit features (e.g., the presence of certain atoms). These sections should be retitled to cultivate knowledge structures centered around implicit variables (e.g., instead, using the title “effects of proton exchange and quadrupole moments on splitting”). If instructors then deem underlying variables like quadrupole moments as content too

advanced for introductory instruction, then perhaps instruction should not include protons like those attached to nitrogen until requisite conceptual understanding *is* accessible. As one final, immediate potential change to instruction, curricular materials should avoid providing students with small ranges of characteristic chemical shift values. By providing very wide characteristic ranges or even relative ranges (e.g., stating that aromatic protons appear moderately downfield and carboxylic acid protons appear far downfield) initial learners will be encouraged to evaluate spectra holistically rather than extensively process reference material. More narrow characteristic ranges can then be introduced in advanced undergraduate courses where students synthesize particular compounds or in graduate-level courses where students synthesize particular classes of compounds in research laboratories. Again, introductory-level instructors must question if they would like their students to efficiently cultivate the requisite skillset that will later allow them to interpret a range of spectra, or if they would like students to cultivate this ability incrementally and through extensive practice with a range of spectra over years of graduate-level study.

In addition to redesigning undergraduate curricular materials, transforming instruction will also require adequately preparing teaching assistants to promote multivariate, analytical thinking among their undergraduate students. These instructors regularly interact with undergraduates as they learn to interpret spectra in laboratory courses and office hours. As a result, they have considerable potential to positively impact students' learning of ^1H NMR spectroscopy. However, teaching assistants in our investigation did not report knowledge of instructional strategies that would shift students to analytical thinking, despite being aware of constraints on undergraduates' thinking. Further, some teaching assistants were unaware of these constraints and reported instructional strategies that would promote constrained thinking. Teaching assistants should thus be trained in using the evidence-based instructional strategies identified in Chapter 2. Lastly, additional evaluated formative assessments in NMR spectroscopy are needed to adequately support learning. While this work contributes one such assessment, there are multiple aspects of expertise in interpreting ^1H NMR spectra that instructors must be able to formatively assess and in turn support. For instance, instructors need

assessments that help them determine if their students are considering and weighing underlying variables as they evaluate spectra or if they are instead using explicit features and rule-based reasoning. Such assessments will help further ensure positive learning outcomes.

Transformation of undergraduate-level instruction on ^1H NMR spectroscopy will be a challenge that requires a concerted effort among chemistry education researchers, instructors, teaching assistant coordinators, and curricular material authors. For instance, chemistry education researchers need to investigate how individuals develop conceptual understanding across the five identified areas, to design instructional innovations that effectively cultivate this understanding, and to design formative assessments that will help instructors support this development. As part of designing instructional innovations, chemistry education researchers also need to investigate how NMR spectra are used within the organic chemistry community so that this instruction can incorporate relevant tasks and thus cultivate relevant understanding. Instructors should also work to design evidence-based instructional innovations, where researchers can then assist with evaluating the efficacy of such innovations. In addition, instructors and curricular material authors must work to redesign the order and manner in which they present content using findings from this research; addressing the immediate potential changes identified above will be an essential first step in this redesign. Lastly, teaching assistant coordinators must also work to train teaching assistants in the use of evidence-based instructional strategies, where the design and evaluation of such training programs will require the assistance of chemistry education researchers. Through this concerted effort informed by our research, the organic chemistry community will then hopefully begin to see its newest members enter the field being fully capable of using this essential tool.