

Dartmouth College

## Dartmouth Digital Commons

---

Master's Theses

Theses and Dissertations

---

9-7-2017

### A Digital Practice Tool for Chemical Resonance

Kimball Jaclyn  
*Dartmouth College*

Follow this and additional works at: [https://digitalcommons.dartmouth.edu/masters\\_theses](https://digitalcommons.dartmouth.edu/masters_theses)



Part of the [Computer Sciences Commons](#)

---

#### Recommended Citation

Jaclyn, Kimball, "A Digital Practice Tool for Chemical Resonance" (2017). *Master's Theses*. 26.  
[https://digitalcommons.dartmouth.edu/masters\\_theses/26](https://digitalcommons.dartmouth.edu/masters_theses/26)

This Thesis (Master's) is brought to you for free and open access by the Theses and Dissertations at Dartmouth Digital Commons. It has been accepted for inclusion in Master's Theses by an authorized administrator of Dartmouth Digital Commons. For more information, please contact [dartmouthdigitalcommons@groups.dartmouth.edu](mailto:dartmouthdigitalcommons@groups.dartmouth.edu).

Dartmouth Computer Science Technical Report TR2017-836  
**A DIGITAL PRACTICE TOOL FOR CHEMICAL RESONANCE**

A Thesis  
Submitted to the Faculty  
in partial fulfillment of the requirements for the  
degree of  
Master of Science  
in Computer Science  
by  
Jaclyn Kimball

DARTMOUTH COLLEGE

Hanover, NH

September 7, 2017

Examining Committee:

---

Devin Balkcom, Chair

---

Lorie Loeb

---

Xing-Dong Yang

---

F. Jon Kull, Ph.D.  
Dean of Graduate and Advanced Studies

# Abstract

Digital practice tools support online learning in math, language, computer science, and other subjects, but practice with problems whose answers are not well represented by text or quantities is underrepresented in the digital learning ecosystem beyond multiple-choice questions. This thesis project explored an alternative to multiple choice practice problems in organic chemistry that does not rely on a molecule drawing interface. This project included development and evaluation of a proof-of-concept digital practice tool for chemical resonance problems. Results of a utility study strongly suggest that the practice tool could fill a learning niche within organic chemistry practice as part of a larger integrated learning environment. The study supported the idea that the digital practice tool and others like it can meet different needs for certain learners, such as reinforcing concepts visually, allowing learners to pace themselves, encouraging learners, and providing immediate feedback. Lastly, this project identified generalizable design challenges for similar practice tools, including the need for a known deployment context, curating content for diverse learner backgrounds, and managing appropriate difficulty for diverse learner backgrounds and needs.

# Acknowledgments

With gratitude to my committee members, teachers, friends, and family, and also to those who shared their experiences learning and teaching organic chemistry with me in the early days of this thesis.

# Contents

<b>Abstract</b>	<b>ii</b>
<b>Acknowledgments</b>	<b>iii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Overview . . . . .	1
1.2 Learning and Practice . . . . .	4
1.2.1 Effective Learning . . . . .	4
1.2.2 The Need for High-Quality Practice . . . . .	6
1.2.3 Prior Work: Game-Based and Gamified Practice . . . . .	8
1.2.4 Summary . . . . .	10
1.3 Goals of the Thesis Project . . . . .	10
1.4 Chapter Summary . . . . .	10
<b>2 Selection of the Learning Topic</b>	<b>11</b>
2.1 Relevant Organic Chemistry Background . . . . .	11
2.1.1 Electron Configurations and Formal Charges . . . . .	11
2.1.2 Bond-Line Drawings . . . . .	12
2.1.3 Chemical Resonance . . . . .	14
2.2 Focusing on Organic Chemistry . . . . .	17
2.3 Selecting Chemical Resonance . . . . .	17
2.4 Chapter Summary . . . . .	19
<b>3 Design and Development</b>	<b>20</b>
3.1 Software Overview . . . . .	20
3.2 Chemical Resonance Rules . . . . .	23
3.2.1 Movement Rules . . . . .	23
3.2.2 Significance Rules . . . . .	24
3.3 Notable User-Facing Design Decisions . . . . .	26

3.3.1	User Interface Overview . . . . .	26
3.3.2	Scaffolding and Explicit Error Feedback . . . . .	31
3.3.3	Animation as Implicit Solution Feedback . . . . .	35
3.3.4	Use of Text . . . . .	36
3.3.5	Use of Sound . . . . .	37
3.3.6	Summary . . . . .	38
3.4	Formal Usability Test . . . . .	38
3.4.1	Procedure Overview . . . . .	38
3.4.2	Results . . . . .	40
3.4.3	Impact on Development . . . . .	43
3.5	Technical Implementation . . . . .	44
3.5.1	Overview of Model Classes . . . . .	45
3.5.2	Overview of Interactive Molecule Classes . . . . .	46
3.5.3	Overview of Manager Classes . . . . .	47
3.5.4	Representing Molecule Problem Data . . . . .	49
3.6	Chapter Summary . . . . .	52
<b>4</b>	<b>Utility Study</b>	<b>53</b>
4.1	Procedure Overview . . . . .	53
4.2	Results . . . . .	55
4.2.1	Results of Prior Experience Questions . . . . .	55
4.2.2	Results of Evaluation of Other Resources . . . . .	58
4.2.3	Results of Software Evaluation and Feedback . . . . .	63
4.3	Limitations . . . . .	70
<b>5</b>	<b>Discussion and Conclusions</b>	<b>72</b>
5.1	Analysis of Utility Study Results . . . . .	72
5.1.1	Discussion of Quantitative Results . . . . .	72
5.1.2	Discussion of Qualitative Results . . . . .	73
5.1.3	Generalizable Challenges and Opportunities . . . . .	78
5.2	Future Work . . . . .	79
5.3	Conclusion . . . . .	82
	<b>Appendix A Formal Usability Study</b>	<b>84</b>
A.1	Procedure . . . . .	84
A.2	Observations . . . . .	86
A.3	Questionnaire Results . . . . .	88

<b>Appendix B Utility Study</b>	<b>93</b>
B.1 Procedure . . . . .	94
B.2 Questionnaire Results . . . . .	95
B.2.1 Prior Experiences . . . . .	95
B.2.2 Evaluating Other Resources . . . . .	101
B.2.3 Software Evaluation and Feedback . . . . .	106
<b>Appendix C Software Version for Formal Usability Study</b>	<b>113</b>
C.1 Description of Software Version . . . . .	113
C.2 Deployment . . . . .	113
C.3 Screenshots . . . . .	113
<b>Appendix D Software Version for Utility Study</b>	<b>116</b>
D.1 Description of Software Version . . . . .	116
D.2 Deployment . . . . .	116
D.3 Screenshots . . . . .	116
<b>Appendix E External Assets and Resources</b>	<b>118</b>
<b>Appendix F Margin of Error Calculations</b>	<b>119</b>
<b>Bibliography</b>	<b>120</b>

# List of Tables

4.1	Utility study: academic resource usage among participants . . . . .	59
4.2	Utility study: participant evaluation of academic resources (percentages)	60
4.3	Utility study: participant evaluations of the Software . . . . .	64
A.1	Usability study multiple choice responses . . . . .	89
B.1	Utility study: participant evaluation of academic resources . . . . .	102
B.2	(Repeat of Table 4.1) . . . . .	104
B.3	(Repeat of Table 4.3) . . . . .	107



# List of Figures

1.1	Sample screenshots of the Software . . . . .	3
2.1	Annotated example of a bond-line drawing . . . . .	13
2.2	Example of two resonance structures . . . . .	15
2.3	Example of curved arrows . . . . .	16
3.1	Screenshot: example of a selected carbon atom . . . . .	21
3.2	Screenshot: example of a curved arrow . . . . .	21
3.3	Screenshot: example of feedback for insignificant solution . . . . .	22
3.4	Screenshot: skip button . . . . .	23
3.5	Screenshot: help dialog . . . . .	26
3.6	Screenshot: bond-line drawing . . . . .	27
3.7	Screenshot: error . . . . .	28
3.8	Screenshot: success dialog . . . . .	29
3.9	Screenshot: one-try success dialog . . . . .	29
3.10	Screenshot: success dialog with streak of 2 . . . . .	29
3.11	Screenshot: success dialog with streak of 11 . . . . .	30
3.12	Screenshot: solution examination . . . . .	31
3.13	Screenshot: tutorial dialog for first tier . . . . .	32
3.14	Screenshot: tutorial dialog for third tier . . . . .	32
3.15	Screenshot: sample problem from first tier . . . . .	33
3.16	Screenshot: example of error with implicit hydrogens . . . . .	34
3.17	Screenshot: error dialog example . . . . .	35
3.18	Screenshot: arrow animations . . . . .	36
3.19	Screenshot: usability study prototype, with selection . . . . .	40
3.20	Screenshot: usability study prototype . . . . .	41
3.21	Screenshot: statistics dialog . . . . .	48
3.22	Example of molecule encoding scheme . . . . .	50

5.1	Screenshot: welcome dialog . . . . .	75
5.2	Screenshot: skip dialog . . . . .	76
A.1	Usability study iPad setup . . . . .	85
C.1	Screenshot: usability prototype practice mode . . . . .	114
C.2	Screenshot: usability prototype, with selection . . . . .	114
C.3	Screenshot: usability prototype, solution review mode . . . . .	114
C.4	Screenshot: usability prototype, success dialog with streak of 2 . . . . .	115
C.5	Screenshot: usability prototype welcome dialog . . . . .	115
C.6	Screenshot: usability prototype help dialog . . . . .	115
D.1	Screenshot: WebGL build, practice mode . . . . .	117
D.2	Screenshot: WebGL build, review mode . . . . .	117

# Chapter 1

## Introduction

### 1.1 Overview

Powerful freely available digital practice tools for learning exist in math [1], foreign language [2], and programming [1, 3]. Dynamic tools can adjust to user expertise to guide beginners, boost efficiency of advanced learners, and sharpen detail-oriented skills [4, 5, 6]. Digital media can visualize practice material in different ways, and deliver real-time feedback so that the learner can try the same problem again without having seen the answer, which may help the learner better understand where and how they erred.

However, when practice problem answers are not well described by text or numbers, fewer such resources exist other than those that use multiple-choice questions. Organic chemistry is one domain where drawings and diagrams dominate how information is described. Unlike general chemistry, whose practice problems often involve quantifiable results, organic chemistry requires qualitative thinking about 3D space and logical inferences based on fundamental concepts [7]. Organic chemistry looks specifically at carbon-containing molecules and reactions that involve them [8]. What gives carbon its own domain of study is twofold: its role in living organisms and systems, and the incredible variety of reactions and structures that carbon may be involved in, which is why spatial thinking is so vital [8].

Organic chemistry practice problems in paper media often involve either delayed feedback, from waiting for a human to grade solutions, or leave it to the learner to check their answer against an answer key, with the hope that the learner will take the time to understand their mistakes. In the absence of feedback, one of the worst study habits is to simply check the solution without putting in the work to understand it deeply [9].

With digital media, on the other hand, feedback is immediate and may be personalized to the learner’s solution. Textbook publishers, who are the major content providers for organic chemistry digital practice tools, tend to rely on multiple choice as the dominant problem format when answers cannot be neatly expressed by text or quantities [8, 10, 11, 12, 13, 14]. However, multiple choice is not an effective practice strategy for long-lasting retention or deep learning [15, 16, 17].

To counteract this deficiency, and to build relevant organic chemistry skills, several digital platforms from textbook publishers do include problems that have the user draw molecules with a special drawing interface [11, 12, 13, 18, 19]. These digital practice tools that involve drawing are good for homework due to immediate feedback and labor-saving grading [20], but not necessarily good for self-paced practice. Homework is assigned and graded; practice is self-paced, time-efficient, and targeted, prioritizing some skills over others. The general-purpose molecule drawing interfaces can have dozens of buttons—both Pearson’s Mastering Chemistry and McGraw-Hill Connect have drawing interfaces that display more than three dozen buttons [11, 12, 13, 19]. With these drawing interfaces, a lot of time may be spent copy-pasting, dragging, and fussing with the software, such as making sure that no part of a molecule is drawn outside the specified box [19]. The added complexity and effort of using molecule drawing software may be why textbook publishers showcase multiple-choice problems when advertising quick, targeted organic chemistry practice [8, 10, 11, 12, 13, 14].

With this thesis project, I sought to find an alternative to both multiple choice and cumbersome digital drawing interfaces. I designed and developed a proof-of-concept organic chemistry practice tool that avoids multiple choice without making the user spend time drawing molecules. I narrowed the scope down to one subtopic, chemical resonance, to demonstrate this experimental approach. Relevant organic chemistry background will be described in Chapter 2.

The digital practice tool will hereafter be referred to as the Software for clarity. The format of the practice problems gives one resonance structure and ask the user to find another significant resonance structure. The user taps the molecule to draw arrows from one part of it to another, and receives feedback upon submitting their solution. The feedback includes explicit feedback, such as alerting the user to an error, and implicit feedback, such as visualizing how the user’s arrows would move electrons to change the structure.

With a utility study, I collected evaluations of the Software and of other types of resources from prior students of organic chemistry, in order to gather data about

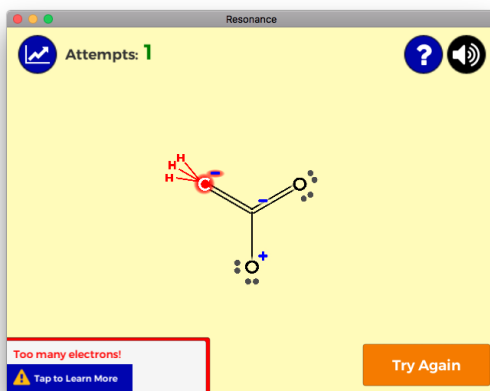
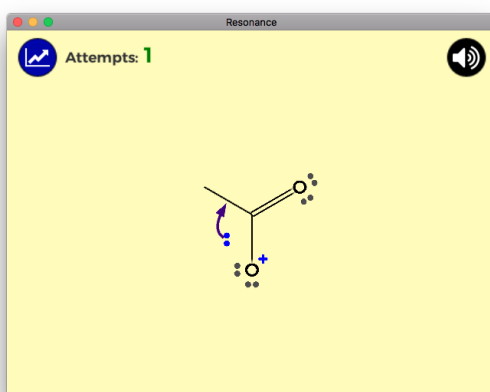
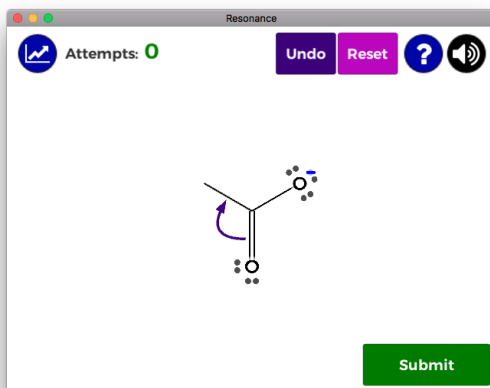


Figure 1.1: Sample screenshots of the Software: arrow drawn by the user in practice mode, animation-in-progress (after the user submits their solution) showing the effect of the user's arrows, and error feedback about the resulting structure.

whether the proof-of-concept may be viable as a practice tool. The results strongly suggest that the Software could fill a learning niche within organic chemistry practice, but its appeal and usefulness require situating the Software and similar tools in an integrated learning environment.

This thesis will describe key concepts related to learning that motivated creating the Software and shaped its design in Section 1.2 of this chapter. The goals of the thesis project are explicitly stated in Section 1.3 of this chapter. Chapter 2 will provide a high-level overview of chemical resonance and related foundational organic chemistry knowledge, and describe how I came to select organic chemistry and specifically chemical resonance as the learning topic. Chapter 3 describes the Software features, as well as their design rationale and high-level technical implementation. That chapter will also discuss a formal usability test during development. Chapter 4 covers the utility study, the results of which Chapter 5 analyzes and discusses. Chapter 5 also discusses future work and concludes this thesis.

## 1.2 Learning and Practice

This section introduces the reader to key terms, fundamental concepts, and pertinent research about learning, so that the reader can understand the basis for this project and its design decisions.

### 1.2.1 Effective Learning

**Learning** is the process of encoding information into usable knowledge in memory, by interpreting the information and linking it to one's own existing knowledge structures [15, 21]. How an individual processes information, rather than how the individual encounters the information, is the most significant factor for turning that information into useful knowledge, though the context and medium may influence processing [15]. We cannot directly observe learning, however, and so we must rely on external indirect measures, such as task performance, to infer whether learning has occurred [21].

People access knowledge structures to perform tasks. As with computational data structures, these knowledge structures have trade-offs. For example, processing information linearly facilitates recalling that information as a sequence, but it impairs other types of access to the information [15]. For example, if a person memorizes a poem by practicing from start to finish repeatedly, and then tries to start from somewhere else in the poem, they will encounter difficulty and may have to recite

from the beginning in their head before reaching the new start position [15].

Thus, knowledge is more than mere data; it also encompasses the structures encoding that information. In a computer science context, depending on how one intends to access, modify, or restructure information in the future, some structures achieve better performance than others. The same is true for individual human knowledge. A data structure in a person's mind is called a **schema** (plural: schemata) [15]. As with computational data structures, a schema can be a complex structure that organizes and links smaller schemata. A schema may structure knowledge visually, verbally, procedurally (scripts), sequentially (lists), declaratively, and in other forms, including mixed forms and links to related schemata [15, 21]. Extending the computer science analogy, declarative knowledge is to procedural knowledge as a constant is to an algorithm. Procedural knowledge informs how to use knowledge found in related schemata and discerns whether using that knowledge is appropriate [15, 21, 22, 23]. If a person lacks procedural knowledge for how and why to use a piece of declarative knowledge, such declarative knowledge is considered **inert knowledge** [15].

At its heart, expertise in some domain of knowledge is the cumulating effects of complex, interconnected schemata which allow experts to form correct inferences when confronting new problems in that domain [15, 21]. How one processes material determines learning outcomes far more than time spent on the material [15]. Moving towards expertise requires a virtuous cycle of learning, because how learners organize knowledge determines not only how that knowledge can be used, but also how it can promote future learning by creating a meaningful framework in which to fit new knowledge [15, 21, 22]. The converse is also true: defects in prior knowledge impair how people learn related material in the future [21].

Unfortunately, too many beginner students of organic chemistry expect that the course is all about memorization, or feel that sheer memorization is the only way to keep pace with the course [7, 9, 24]. Though repetition-based memorization is relatively quick and has some value when combined with other learning strategies, by itself repetition-based memorization is a particularly ineffective strategy for building expertise and establishing schemata [15]. The resulting schemata organize information around superficial features and thus offer poor foundations for integrating new information meaningfully [15, 21]. In fact, knowledge acquired purely through repeated exposure to material may not even stay in memory until the end of an academic term [15, 17]. As Mastascusa et. al. note in their book *Effective Instruction for STEM Disciplines*, the same strategies that allow undergraduate students to cram knowledge for a midterm prevent them from recalling that knowledge by

the time of the final exam unless they reprocess the knowledge more actively, such as practicing challenging problems and testing their knowledge [15, 16]. Actively reprocessing material through practice reinforces and extends knowledge, which not only improves task performance, but also establishes robust schemata that will foster future learning.

### 1.2.2 The Need for High-Quality Practice

Explicit instruction can only take a learner so far. The best instruction cannot substitute for putting the onus on the learner to build their schemata through firsthand practice and self-teaching [15, 23]. High-quality practice builds better schemata and automatizes low-level processes, both of which promote learning related high-level material more deeply. The characteristics of high-quality practice include a specific, explicit goal; material that challenges but does not overwhelm the learner; and appropriate volume, variety, and frequency of task performance to develop relevant skills [16, 21].

Often, facilitating high-quality practice requires controlling the difficulty level to avoid overwhelming or confusing the learner. One such approach is **scaffolding**, which are supportive techniques for introducing material at an appropriate level of challenge and, as learners gain mastery, gradually lessening the cognitive support [21]. Scaffolding can allow learners to practice a subset of skills first before broadening practice to include more complex skills [21]. It also helps learners approach higher-level concepts and tasks that would otherwise be too challenging for their current abilities and knowledge [21, 25].

Similarly, some initial guidance before practice can prepare the learner to integrate new information. For understanding complex systems, best outcomes involve providing some initial explicit guidance before immersing the learner into varied experiences with the system, in which the learner can try and practice procedures [23]. Ultimately, learners teach themselves, but good instruction at the right time can direct learners' attention to salient features and relationships if the system is complex [23]. The initial guidance should clue the learner in to the underlying structure of the system, with the intent to guide the learner in crafting their own schema in a logical way [23].

People have a small and finite amount of **working memory** for use at any given moment, which can be filled with external information from the environment or with information retrieved from **long-term memory** which can store vast amounts of



knowledge [15, 26]. High-quality practice develops rich and robust schemata through high-quality **retrieval events**, in which the learner pulls knowledge out of long-term memory so that it can be used in the current situation [15, 16, 17]. Reprocessing knowledge during a retrieval event strengthens the resulting schema and is essential for deep learning [15, 16].

Two techniques increase retrieval difficulty: spacing out the retrieval events over time and reducing cues [15, 16, 17]. Spacing out the retrieval events ensures that no trace of the knowledge hides in working memory, thereby forcing a complete retrieval [15, 16, 17]. Reducing cues, on the other hand, demands that the learner recognizes the opportunity or need to apply that particular knowledge; it is why recalling knowledge to solve a multiple-choice question is easier than facing that question without a list of potential answers [15, 16, 17]. Difficulty creates high-quality retrieval events whose learning benefits greatly surpass those of low-quality retrieval events [15, 16, 17]. In other words, the hard work of encoding information deeply, with quality retrievals over time, eventually pays off: in the long term, such knowledge can be retained, applied, and built upon to forge strong and versatile schemata [15, 16, 17].

Practice not only strengthens existing knowledge, it also builds new knowledge from prior knowledge. All types of knowledge benefit from practice, but procedural knowledge requires it [22]. It is the difference between knowing about the rules, and knowing how the rules work, sometimes without being able to articulate the rules or even consciously recognize them [22, 23]. Without experience, a person's knowledge of the formulas and rules governing a system are only inert, shallow, declarative knowledge [23].

To build new knowledge from prior knowledge through practice requires appropriate feedback. Feedback should guide learners in constructing their schemata properly by explicitly relating some aspect of their performance to a known goal, without replacing self-teaching by appearing too often or giving away too much [15, 21]. With repeated practice, learners test hypotheses and, with useful feedback, correct inaccurate prior knowledge and acquire new knowledge [21]. Good feedback intercepts information while it still sits in the learner's working memory, so that the learner can self-correct misconceptions and fill in foundational knowledge before depositing information into long-term memory [15, 21].

One of the most powerful performance benefits of richer schemata is called **chunking**, in which links from the schemata in working memory enable simultaneous access to a larger quantity of relevant knowledge than working memory could otherwise hold [15, 27]. To use a computing analogy, chunking allows working memory to

store pointers to knowledge rather than storing all of the information itself, thereby greatly expanding the capabilities of working memory [15]. For effective chunking, the learner’s schemata must emphasize meaningful relationships between information; for example, chess experts remember chessboard configurations better than novices only if the configuration could have appeared during a game, since the experts remember the configuration as meaningful relational chunks rather than as individual pieces [15]. Working memory is fragile and information contained in it can degrade and disappear rapidly, as often occurs unintentionally when a person gets distracted by information from the environment [22, 28], so not only can chunking help a learner access more information at once, the information itself may be more reliable and better quality.

In addition to improving schemata and forming new knowledge, practice automatizes low-level processes, which frees up cognitive resources to tackle more complex higher-level problems, speeds up task performance, and improves performance under stress [22, 23, 27, 29, 30]. These improvements arise because acquiring greater skill reduces the required conscious thought and effort to perform the task under most conditions, and automatic tasks have less performance impact on concurrent tasks [22, 28]. For students stressed during an exam, prior high-quality practice may lessen the negative impact of stress on their performance, as unconscious processes suffer less from the effects of stress than conscious ones [23, 26, 28, 29, 30].

### **1.2.3 Prior Work: Game-Based and Gamified Practice**

Long before the rise of digital learning tools, spelling bees were a widely known example of gamified learning, even though they exemplify how an educational game can fail to facilitate learning: the mechanics do not help struggling learners figure out how to improve their spelling [31]. That defect has not disappeared from the gamified learning landscape with time; memory drill activities disguised as games may be fun for the player who is already “good” at the subject, but provide little information about how to improve performance and fail to highlight what is innately enjoyable about the subject [32, 33].

What games and high-quality practice have in common is that they create and strengthen procedural knowledge through making mistakes and learning from those mistakes [34, 35, 36, 37]. Some “chocolate-covered broccoli”—a term for gamified practice tools, some of which try to pass as games—may fall into the same trap as the spelling bee, but others can sweeten the hard work of practice while providing

useful feedback [38]. Frequent exposure to encouraging, clear, and constructive feedback can bolster intrinsic motivation [36]. Andy Matuschak, co-director of long-term research at Khan Academy, writes that education research and game design often investigate the same questions about how people learn and how much explicit guidance they need [39].

For general chemistry, digital games and game-like learning tools abound, though simulations and other exploratory tools appear to be much more common than narrowly practice-focused tools [40, 41, 42, 43].

A board game called React! recently received Kickstarter funding [44], and when released to the public, the game aims to offer a social and competitive way to practice laboratory-related synthesis skills [45]. Aimed at a general audience, React! may not be the most effective study strategy, though it can reinforce existing knowledge [44, 45]. However, it is far from the first tabletop game to tackle organic chemistry [40]. Similarly, several organic chemistry card games have been developed for practicing synthesis-related logic skills, some of which are intended for an audience with some prior knowledge [40, 46, 47]. These games allow learners to practice synthesis logic without having fully memorized the reactions, since they are printed on the cards; the cards act as scaffolding for developing synthesis logic skills earlier than would otherwise be possible.

Though these games may let the player practice useful skills, many of these games are in effect offering a kind of multiple choice—with choices as cards in a person’s hand, for example. Still, they allow players to practice high-level skills before completely mastering low-level details, which may help learners keep the big picture in mind. Digital games for organic chemistry often also rely on multiple choice-like mechanics, some even going as far to be simply gamified multiple-choice practice problems [40, 48].

One digital practice tool that takes a different approach is Chairs!, a mobile app that takes a game-based approach to practicing one very specific organic chemistry spatial skill under time pressure: drawing the bonds of a cyclohexane isomer, given the other isomer [49]. Though very small in scope, the app effectively offers an alternative to multiple choice for quickly practicing a spatial analytical skill in organic chemistry. The unique approach of Chairs! partially inspired this thesis project, even though Chairs! and the Software use very different practice mechanics.

### 1.2.4 Summary

This section explained how effective learning works: through connections to prior knowledge and repeated active processing of material through high-quality practice. This section also briefly discussed prior work in gamified organic chemistry learning.

## 1.3 Goals of the Thesis Project

This thesis project sought to achieve the following goals:

1. Develop a proof-of-concept digital practice tool for a subtopic of organic chemistry, without relying on multiple choice or user drawings of the molecule, using research on learning to guide design decisions
2. Evaluate whether such a tool has potential to benefit learners of organic chemistry
3. Identify some of the generalizable design challenges to developing similar practice tools for advanced subjects

## 1.4 Chapter Summary

This chapter presented an overview of this thesis and introduced interdisciplinary research, key terms, and prior work to prepare the reader for understanding content in subsequent chapters. This chapter concluded by summarizing the goals of this thesis project.

# Chapter 2

## Selection of the Learning Topic

This chapter explains key organic chemistry terms and concepts related to chemical resonance, which is the learning topic of the Software, and then discusses the impetus for selecting the learning topic.

### 2.1 Relevant Organic Chemistry Background

This section briefly introduces the reader to an essential subset of organic chemistry terms and conventions, emphasizing those that are most crucial to understanding the design and implementation of the Software. I have taken the liberty of simplifying my explanations to an appropriate level for these purposes.

#### 2.1.1 Electron Configurations and Formal Charges

A **molecule** is defined by a set of atoms joined by covalent bonds [50].

The electrons that may be involved in chemical reactions are the outermost electrons surrounding an atom, known as **valence electrons** [9]. The periodic table organizes elements by atomic number, which is the number of positively charged protons any atom of that element will have, and thus the location of an element in the periodic table indicates the expected number of negatively charged valence electrons for a neutral atom of that element [8, 9].

If an atom has more or fewer valence electrons than what its atomic number predicts, the atom has a **formal charge** [8, 9]. Excess valence electrons produce a negative formal charge; missing valence electrons produce a positive formal charge [8, 9]. A negative formal charge is typically represented by a minus sign placed near the atom on a drawing, and a positive formal charge by a plus sign [8, 9]. Notably,

formal charges are an imaginary device to track the discrepancy between expected and actual valence electron count, and they do not expressly indicate the presence of a real charge [8].

If an atom has valence electrons, each electron either forms a bond between the atom and another atom, or the electron does not form a bond. In the latter case, the electron is most often part of a pair of nonbonding electrons which are called a **lone pair** [8, 9]. A lone pair adds two electrons to an atom's valence electron count.

If an electron forms a bond between two atoms, it does so as a pair of electrons. Covalent bonds involve two, four, or six electrons shared between two atoms, as a single, double, or triple bond respectively. However, when counting charges on an individual atom, a bond is considered to contribute half of the electrons involved to each atom's valence electron count, under the simplifying assumption that the atoms are sharing the electrons equally between them [9].

Formal charges are always drawn in any conventional 2D representation of a molecule; since lone pairs can be inferred from an atom's formal charge and bond configuration, lone pairs can be and often are omitted from drawings [8, 9].

### 2.1.2 Bond-Line Drawings

The ability to quickly read and interpret a drawing of a molecule is one of the most fundamental skills in organic chemistry [9]. Within organic chemistry, drawings are how textbooks and lectures explain concepts, facts, and relationships; verbal and mathematical descriptions are insufficient to convey the nuances of organic chemistry's spatial and relational aspects to a human audience.

Among the ways of drawing molecules, the most prevalent representation is **bond-line drawings**, also called bond-line diagrams, bond-line structures, line-bond structures, or Kekulé structures [9, 8, 51]. What makes bond-line drawings particularly quick to draw and read is a minimalist approach: conventions de-clutter the molecule by omitting information which can be inferred. These conventions reward fluency with efficiency and clarity, but for novices, omitted information steepens the learning curve.

Briefly, I will discuss the salient features of the bond-line drawing representation system.

Bond-line drawings depict a molecule's carbon skeleton, defined by the carbon atoms and their bonds, where each straight line represents a covalent bond and each corner or endpoint represents a carbon atom, unless such a corner or endpoint is

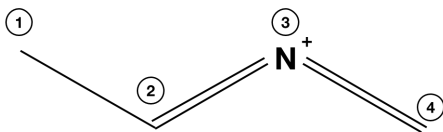


Figure 2.1: Example of a bond-line drawing, with circled number labels added for illustrative purposes. There are four atoms depicted in the molecule, at each numbered spot. All except #3 are carbons. However, the atom has six implicit hydrogens attached to carbons, so the actual number of atoms in this molecule is ten. The "+" at #3 tells us that the nitrogen atom ("N") has a positive formal charge, which for nitrogen whose expected valence electron count is five, means that it has only four valence electrons. This implies that the nitrogen at #3 has no lone pairs since it has four bonds (two double bonds, with #2 and #4 respectively) contributing four electrons to its valence electron count.

otherwise marked with another element's symbol [8, 9]. When hydrogen bonds to carbon, both the hydrogen atom and the bond are omitted, as the presence of hydrogen may be inferred from the formal charge of the carbon in an organic chemistry context; such a hydrogen atom is called **implicit hydrogen** [9]. Hydrogen atoms are only shown if they bond to a non-carbon atom, frequently by grouping the hydrogen with the non-hydrogen atom and omitting a visual depiction of the bond [9]. See Figure 2.1 for an example of a bond-line drawing.

Recognizing an unmarked carbon atom is the first hurdle for beginners, but the more difficult feature to see is how electrons are distributed in the molecule due to omitted implicit hydrogens and lone pairs [9]. Without automatic pattern recognition of the different bond and formal charge configurations on carbon, a beginner must do some quick algebra to understand a carbon atom's electron configuration. If a beginner cannot read the lone pair count and implicit hydrogen count on each carbon, they will either have to take the time to externally record it, or by keeping it all in their head risk losing track of the counts across multiple atoms. Moreover, while a fluent reader of bond-line drawings may diagnose an error at a glance, a novice may easily forget about an implicit hydrogen and thus have a defective mental model of the molecule and any reactions it may be involved in.

### 2.1.3 Chemical Resonance

As discussed in 2.1.1, the simplified way to think about valence electrons is that they either form a bond between two atoms, or do not form a bond. Drawings of molecules also describe electrons in this way. However, this model can be inaccurate for certain molecules.

Rather than thinking about electrons as stationary particles, chemists think about electron density clouds [8, 9]. In chemistry, the location of an electron is typically specified by **orbitals** (zones around atoms that may contain electrons, described by a wave function) rather than a single fixed location [8]. The shape and placement of electron-containing orbitals with respect to other atoms indicates whether a covalent bond is formed or not. Each covalent bond on an atom completely occupies one orbital; lone pairs can also completely occupy an orbital.

However, when a molecule exhibits **resonance**, the electron configuration described by a single bond-line drawing fails to convey the electron density of the molecule [8]. (Outside of a chemistry context, resonance may be referred to as "chemical resonance" to distinguish it from the unrelated phenomenon of resonance in physics and other usage of the term.) Resonance denotes distributions of electrons across orbitals that conventional representation systems cannot describe accurately with one drawing alone [8, 52, 53].

Multiple drawings represent a molecule exhibiting resonance, where each drawing describes a **resonance structure**, also called a resonance form or resonance contributor [8, 9, 52, 53]. These structures are imaginary: useful for representation but not descriptive of a physical state [8, 53].

Chemists put a straight double-headed arrow between drawings that represent the resonance structures of the same molecule, and frequently (but not always) enclose the set of drawings in square brackets (see Figure 2.2 for an example) [8, 9, 52, 53]. The only difference between the structures is the configuration of nonbinding electrons (lone pairs) and pi (double or triple) bonds; all pairs of bonded atoms share at least one covalent sigma (single) bond that does not change across structures [8, 53].

A common misconception is that a real-world molecule exhibiting resonance switches its electron configurations between two or more resonance structures, or that multiple forms of the real-world molecule exist, but in fact, all instances of the real-world molecule share one stable electron configuration [8, 9, 53]. Even though two or more different structures represent the molecule in bond-line drawings, a single weighted composite of these structures best represents the real-world **resonance hybrid** [8,



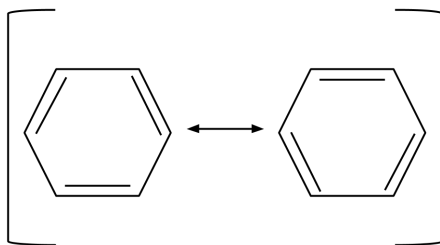


Figure 2.2: Example of two resonance structures linked by a double-headed arrow and surrounded by square brackets. The brackets are optional; the double-headed arrow is not.

9, 52].

For example, one resonance structure may be:

1. Atoms A and B have a double bond
2. Atoms B and C have a single bond

The other resonance structure may be:

1. Atoms A and B have a single bond
2. Atoms B and C have a double bond

The resonance hybrid composites the resonance structures such that:

1. Atoms A and B have a bond stronger than a single bond but weaker than a double bond
2. Atoms B and C have a bond stronger than a single bond but weaker than a double bond

The resonance hybrid most accurately models the real-world molecule.

Resonance occurs because the imaginary resonance structures of the molecule would be unstable, and rather than oscillate between unstable electron configurations, the real-world molecule settles at the most stable configuration. Thus, given one resonance structure of a molecule, chemists deduce other valid resonance structures by identifying unstable regions of the given resonance structure. To communicate how

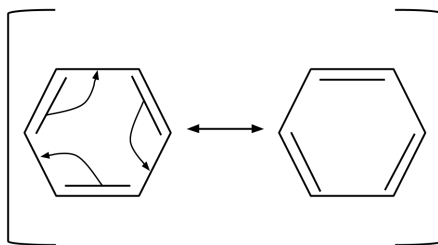


Figure 2.3: Example of curved arrows illustrating how the resonance structure on the left may be transformed into the resonance structure on the right. Each curved arrow in this example has a pi (double) bond as its electron source at the arrow tail, with the arrow head leading to an adjacent bond.

knowing one resonance structure implies the existence of another, **curved arrows** represent *imaginary* movement of electron density from one configuration to another [8, 9, 53]. Though this movement is imaginary (since the molecule exists as a constant hybrid of the structures), when students encounter reaction mechanisms later on in an organic chemistry course, curved arrows reappear to signify actual changes in electron density [9, 53].

A curved arrow runs from the source of the electrons to where the electrons are going (an imaginary process for resonance, an actual process for reaction mechanisms) [8, 9]. Moving electrons with curved arrows is called **pushing electrons**, or, alternately, electron pushing or arrow pushing [9, 52].

The rules regarding imaginary electron density movement will be discussed later in 3.2.1.

Not all resonance structures contribute equally to represent the real-world hybrid [8, 9]; those that make significant contributions are called **major contributors** or **significant** structures, and those that do not are called **minor contributors** or **insignificant** structures [9]. In general, more stable electron configurations contribute more to the resonance hybrid than less stable configurations, and the resonance hybrid itself is more stable than any of its contributors [8, 52]. Even though insignificant resonance structures are valid, the distinction between significant and insignificant structures is useful when considering the molecule's real-world characteristics, as significant structures may overshadow an insignificant structure's influence on the resonance hybrid [9]. The rules about how to determine whether a structure is insignificant will be discussed later in 3.2.2.

## 2.2 Focusing on Organic Chemistry

Organic chemistry has a widely acknowledged reputation among undergraduates as a notoriously difficult, memorization-heavy course that will destroy many students' grades, and as a result students often begin an organic chemistry course with fear and dread [7, 9, 24]. The reputation is not entirely unmerited, as many educators agree that students often struggle with organic chemistry, but the characterization of the course's memorization demands is perilously misleading [7, 9].

With a strong general chemistry background, organic chemistry is not conceptually difficult, as there are only a few fundamental concepts underlying the vast majority of course content [7]. However, not all students come in with a strong general chemistry background, and furthermore, even those that do have the background may fail to see the connection between their coursework and the deeper fundamentals because they rely too much on shallow memorization [7].

The way to succeed in the course is to learn the fundamentals and then work on applying them repeatedly with practice problems [8, 9]. The vast amount of practice involved in learning organic chemistry and its formidable reputation piqued my interest as a place for a learning intervention. Through dozens of informal conversations with former organic chemistry students, current professors, teaching assistants, and tutors, I began to hone in on a few early course topics where a digital tool might help struggling students solidify the fundamentals through scaffolded practice.

## 2.3 Selecting Chemical Resonance

After I narrowed the topic to organic chemistry, I selected resonance as the subtopic for three primary reasons.

First, I had to consider my own limits as a lone developer with no formal education in organic chemistry. Therefore, I gravitated towards early topics as the most promising candidates.

Second, despite being an early topic in organic chemistry courses, resonance resurfaces in later coursework. As one educator warns, "Resonance is the one topic that permeates the entire subject matter from start to finish. It finds its way into every chapter, into every reaction, and into your nightmares if you do not master the rules of resonance" [9]. Most reactions that occur between organic molecules result from regional differences in electron density across two molecules, and thus, predicting whether and how a molecule will react will occasionally require a firm understanding

of resonance [9].

Third, resonance is a major checkpoint for learners in many organic chemistry curricula. Learners must bring together foundational skills and knowledge areas that they have recently acquired in order to succeed at resonance. Later coursework also relies on this foundational knowledge. Additionally, later coursework depends on learners grasping electron pushing, which students first encounter in resonance problems. Weak or incorrect prior knowledge may interfere with acquiring and integrating new knowledge, and the more fundamental that prior knowledge is, the more severe the impact may be [21]—and educators agree that resonance and its underlying concepts are very fundamental [9, 53].

Given that some learners may not yet have mastered the foundational knowledge and skills, a digital tool can offer greater scaffolding than paper technology to help struggling learners develop their foundational knowledge while practicing resonance problems. Perhaps the most important foundational skill that may still be in progress is fluently reading bond-line diagrams. Building bond-line drawing fluency is important, along with related skills such as recognizing the electron density patterns in a molecule.

Synthesis is one topic that students learn later in the course. A study on successful synthesis problem-solving strategies among organic chemistry students found that the largest gap in skills between successful and unsuccessful answers involved correctly identifying the differences in bond formation between two molecules [54]; in other words, the skills to correctly read and discern bonding differences when looking at two drawings.

Lack of fluency greatly hurts learning, just as having to sound out every letter to read a text impairs understanding that text due to the limits of working memory and cognitive capacity [28]. Like language, the rules of how to read the diagram are a representational convention, but they are essential to understanding the science. To a fluent reader of bond-line drawings, two molecules whose only visible difference is a double bond will seem as different as "snake" and "shake" seem to an experienced reader of English, despite strong visual similarities, and catching an error in one's drawing becomes as automatic as catching a typo in one's writing.

Thus, I settled on chemical resonance as the topic for practice, with the intent to use scaffolding to help learners cement their bond-line diagram knowledge to better understand the fundamentals behind resonance.

## 2.4 Chapter Summary

This chapter presented a high-level overview of the basic organic chemistry topics required to understand the Software's design, including topics such as formal charges, bond-line diagrams, and chemical resonance. This chapter also discussed the challenges that students face in undergraduate organic chemistry and how chemical resonance practice presents an opportunity to reinforce foundational knowledge with a digital practice tool.

# Chapter 3

## Design and Development

This chapter presents major design and technical decisions during the Software's development. First, this chapter provides an overview of the Software and its key features; second, this chapter describes the rules of chemical resonance that the Software enforces; third, this chapter discusses notable design decisions; fourth, this chapter explains the impact of a formal usability study; and lastly, this chapter gives a high-level overview of the Software's technical implementation.

### 3.1 Software Overview

The Software is a proof-of-concept digital practice tool for organic chemistry that does not involve multiple choice answers and does not require the user to draw molecules. The subtopic in organic chemistry is chemical resonance, and the practice problems give one resonance structure and ask the user to find another significant resonance structure. The core skills that the Software intends to build in its users are fluency with reading bond-line drawings (including formal charges, implicit hydrogens, hidden lone pairs, and curved arrows) and pattern recognition for deriving significant resonance structures from an initial structure. The Software assumes that the user has some basic prior knowledge of bond-line diagrams, resonance, and fundamental chemistry concepts.

Users interact with the given molecule by tapping to draw curved arrows from either an atom or a bond, to another atom or bond. When a user is satisfied with their solution, they tap the "Submit" button. Upon submission, the Software verifies first that all curved arrows have a valid electron source at their tail; if not, then the Software alerts the user to an error and gives the option to try again. If the arrows have valid electron sources, then the Software animates the imaginary movement of

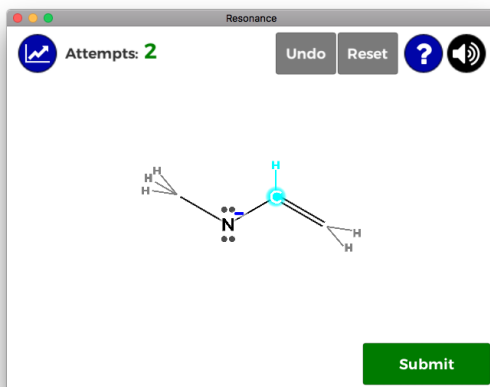


Figure 3.1: Example of a carbon atom selected by the user.

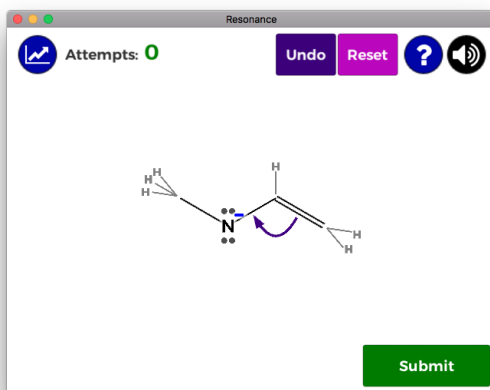


Figure 3.2: Example of a curved arrow drawn by the user.

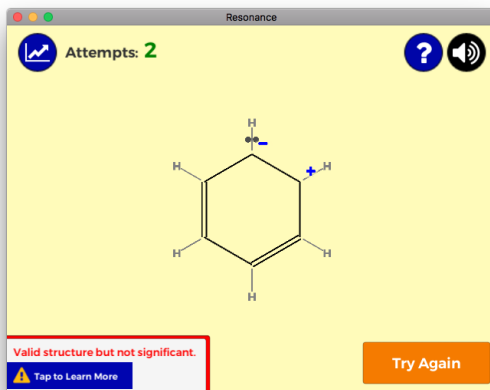


Figure 3.3: Example of feedback for an insignificant solution.

electrons along the curved arrows and shows the resulting structure. The Software verifies that the resulting structure is a valid resonance structure and that it is significant. If the resulting structure is not a resonance structure or is insignificant, the Software warns the user and gives the option to try again.

After the user has worked through the initial set of problems marked in code as beginner problems, whenever the user submits three or more unsuccessful solutions to a problem, the user will have the opportunity to skip the problem in addition to trying again. Skipping a problem moves on to a new problem, but increases the odds that the skipped problem will reappear soon. The skip button does not appear during problems marked by the Software as "beginner problems" since there are no easier problems available; if the Software were embedded in a learning environment with other resources, it would be ideal to redirect a user to instructional material if they cannot solve the beginner problems in fewer than three tries, because that would suggest that their prior knowledge is inadequate for this type of practice.

When the Software alerts the user to an error, the alert's level of detail depends on the scaffolding tier that the user is on. As the user shows mastery, scaffolding is removed. The four scaffolding tiers in the Software progress from having all scaffolds toggled on, to hiding the implicit hydrogen unless there is an error, to also hiding the lone pairs unless there is an error (so the non-error state of the drawing looks like a normal bond-line drawing), to never showing the implicit hydrogens or lone pairs (so the user always sees normal bond-line drawings), to generically alerting the user that there is an error. This progression helps ease learners towards a more exam-like experience, where they will be responsible for diagnosing their own errors. The Software tracks the number of attempts per problem and the user's streak of solving



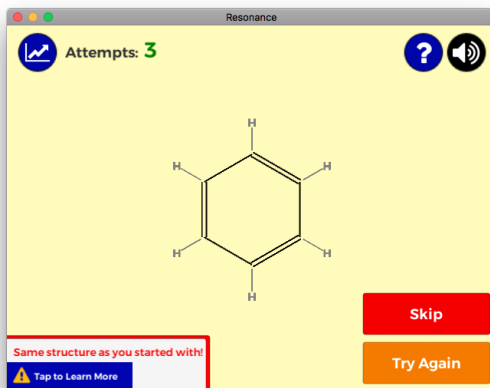


Figure 3.4: The skip button appears after three unsuccessful attempts to solve the problem.

problems on the first try to encourage the user to try to catch their own errors before submitting.

The Software was built in Unity and thus can be deployed on multiple platforms (e.g. web, mobile, desktop).

## 3.2 Chemical Resonance Rules

This section outlines some of the rules of chemical resonance problems that the Software enforced.

### 3.2.1 Movement Rules

For resonance, the rules about which electrons may participate in imaginary movement and where they may go are as follows:

1. Single (sigma) bonds cannot be broken [8, 9, 53]
2. **The Octet Rule:** Valence electrons cannot overflow an atom's available orbitals [8, 9, 53]
3. Two electrons from a pi (double or triple) bond can move to become lone pairs on one of the atoms in the bond, but not on any other atoms [53]
4. Two electrons from a pi (double or triple) bond can move to another bond only if the other bond includes one of the atoms in the initial bond [53]

5. Two electrons from a lone pair on an atom can move to a bond on that atom, but not to any bond that does not include that atom [53]

Though extraordinarily rare exceptions to these rules exist, first-semester organic chemistry students should assume that these rules are always in effect unless an instructor specifically presents one of the exceptional cases [9]. The rule about respecting an atom's orbitals is typically called **the octet rule** since most common organic molecules (except hydrogen) have four orbitals that can hold no more than two valence electrons each.

It is important to note that drawings with multiple arrows treat the electron movements as happening simultaneously, not sequentially. Thus, while a curved arrow might violate the octet rule on its own, another curved arrow could pull electrons away from the atom simultaneously and thus prevent the rule from being broken [9]. In addition, an arrow whose tail is at a single (sigma) bond or at an atom without any lone pairs is always incorrect, even if another arrow's head is pointed at that same location—each arrow represents moving *one* unique pair of electrons, whether from a bond or lone pair.

The Software enforces the last three rules, about not moving electrons too far, by undoing the arrows right after they are drawn and displaying an error message, even though the user is not in the solution review state; the idea is to condition the user to automatically prefer arrows that do not take electrons too far from their atoms.

The other rules, which involve a more intellectual understanding of electron positions and stable orbitals, are enforced by error feedback in solution review mode. The Software first checks the tails of all arrows: that the electron source is not a single bond (which would violate the first rule) and that the source is not an atom without any lone pairs (which has no electrons that can move). If an arrow does not pass the test, it turns red and an error message indicates that electrons cannot be moved from that spot (for all scaffolding tiers except the last, which only shows a generic error message). If all arrows pass the test, then the arrows animate the movement of electrons. The Software then evaluates the resulting structure for any electrons that do not fit into an available orbital (violations of the octet rule).

### 3.2.2 Significance Rules

After the Software confirms that a user's submitted solution is not the same as the initial resonance structure and has no errors, the Software checks whether the user's solution is significant, meaning that it has a significant influence on the resonance

hybrid that represents the real-world molecule.

The Software enforces the following rules for determining significance:

1. Permit no more than two formal charges [9]
2. Electronegative atoms (greater expected valence electron count than carbon) may only have a positive formal charge if they have an octet (meaning that all of their valence orbitals are full) [9]
  - (a) A brief example: a nitrogen atom has four orbitals and an expected valence electron count of five, so if it has four bonds which each fill an orbital, its orbitals are full (an octet) and it has a positive formal charge since it counts one electron from each bond against its expected valence electron count ( $5 - 4 = +1$ )—this nitrogen does *not* disqualify a structure from being significant
3. If any carbon atom has a formal charge, there cannot be another carbon atom with the opposite formal charge, i.e., both a C+ and C- in the same structure are not allowed [9]

Breaking any of these rules results in an unstable but valid resonance structure, which is usually not significant.

Though the Software only uses a binary significant/insignificant distinction for all user-facing feedback, the evaluative algorithm that the Software uses distinguishes between three cases, breaking up significant into highly stable significant structures (no formal charges) and less significant (but still not insignificant) structures (one or two formal charges) [9]. The algorithm was developed in an early iteration of the Software, but the non-binary distinction seemed potentially too complex for a tool focused on scaffolding the learning of those who are still trying to master bond-line diagrams.

A more advanced resonance practice problem tool might include more nuanced significance rules [53], and include special exceptions (for example, certain nitrogen-containing substructures are an exception to the formal charge limit [9])—the challenge would be effectively communicating nuance and exceptions to the user without making them seem arbitrary. A more advanced problem format might also permit finding insignificant solutions (which still sometimes matter [53]) in addition to the significant ones, for example by giving each structure a relative stability score, or having the user assign such a score.

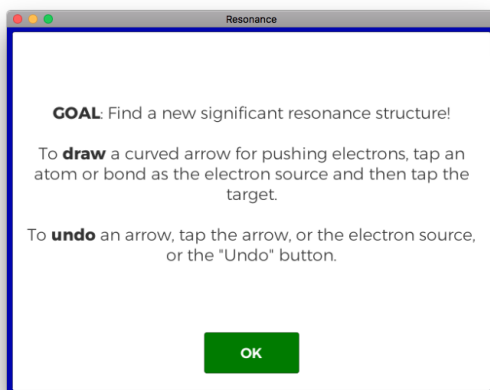


Figure 3.5: The help dialog that can be accessed by tapping the help button during practice mode. It also appears when the user first starts using the tool, but the user may not read it carefully.

### 3.3 Notable User-Facing Design Decisions

This section outlines some of the major user-facing design decisions while developing the Software: user interface, scaffolding, animations, and the small details that go into communicating with the user. Small details subtly influence user attitudes, whether those details are word choices or showing a simple gold star.

#### 3.3.1 User Interface Overview

Users frequently do not read instructions, and even if they do, the instructions may quickly be forgotten [55]. Therefore, in addition to the presence of a help button to show the instructions again (shown in Figure 3.5), the Software supports trial-and-error approaches to learning how to use it. However, the Software's biggest weakness for a trial-and-error approach is that the molecule may not look like an interactive element to the user at first.

Conventions help users know what to expect and how to use an interface without explicit instruction or significant time spent on trial and error [55]. A worthwhile goal is to make anything that the user can interact with look obviously interactive by relying on conventions [55]. While meeting this goal is straightforward with buttons and other common user interface elements, the bond-line drawing representing the problem suffers from conflicting conventions. The main priority of the bond-line drawing is to look like other bond-line drawings in other media, including static media, so by convention it does not look interactive.

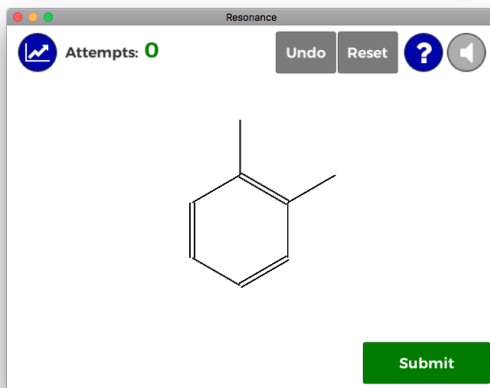


Figure 3.6: A conventional-looking bond-line drawing does not seem interactive at first sight.

For the purpose of this thesis project, the Software was developed in Unity to be similar across all platforms, so a mouse moving over any user interface element does not change the element's appearance because such capabilities are not common to mobile platforms. However, were the Software or a digital tool like it to be developed for non-mobile platforms, a color change when the mouse moves over the molecule would help with making the molecule appear more interactive for a first-time user. For mobile devices with a touch interface, based on the results of the usability study discussed in Section 3.4, future work to redundantly support both tapping and dragging to draw arrows should cover a majority of users' instincts on such platforms.

Other user interface elements exist at the four corners of the screen, with each of the four groupings implying a logical relation [55]. See Figure 3.2 for an example. The upper left corner shows the current number of attempts and a button that summons a dialog which shows more of the user's performance statistics. The upper right corner has the "Undo" and "Reset" button grouped together for their related functions (both functions can also be achieved by direct interaction with the molecule), as well as the help and sound buttons. The lower right corner has the "Submit" button in practice mode while the "Try Again" button replaces it in solution review mode if the solution is not correct or significant; these buttons toggle between modes. When the "Skip" button appears (after three or more unsuccessful submissions on the same problem for non-beginner problems), it appears directly above the "Try Again" button as an alternative next step (see Figure 3.4). When an error appears, the message pops into the lower left corner unobtrusively, so that the user's line of sight on the molecule is not interrupted, which gives the user the chance to catch their own error before

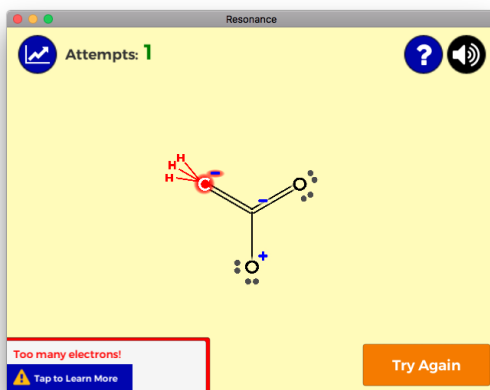


Figure 3.7: An example of an error in solution review mode.

reading the explicit feedback (see Figure 3.7).

When the user submits their solution, they exit practice mode and enter solution review mode. Beyond button changes, the modes are further visually distinguished by a change in background color, from white in practice mode to a light yellow in solution review mode (compare Figures 3.2 and 3.7 for an example). This visual change heightens the visibility of the current mode; with the more paper-like white background, the user can interact with the molecule, whereas when the color is no longer white, the user cannot interact directly with the molecule.

In solution review mode, an animation shows the effects of the user's arrows, and if the solution contains an error or is insignificant, the resulting structure remains on the screen with a notification in the lower left corner about the issue. If the solution is valid and significant, after a brief pause a dialog appears, informing the user of their success (see Figure 3.8). If the success occurs on the first attempt, then the dialog includes a gold star and a congratulatory message for succeeding in only one try (see Figure 3.9). If the user continues to solve multiple problems in a row in one try, then additional stars and text track the user's one-try streak in the success dialog (see Figures 3.10 and 3.11).

Though the stars do give the user some performance feedback about the quantity of their streak, the primary purpose of the stars are as rewarding stimuli—part of simple operant conditioning as found in many games and gamified software that motivates the user to repeat the behavior [30, 56]. In other words, the stars add extrinsic reinforcement to the intrinsic rewards of successfully solving a problem, and specifically encourage the user to do their best work on their solution before submitting it in order to continue their streak, rather than letting the Software catch

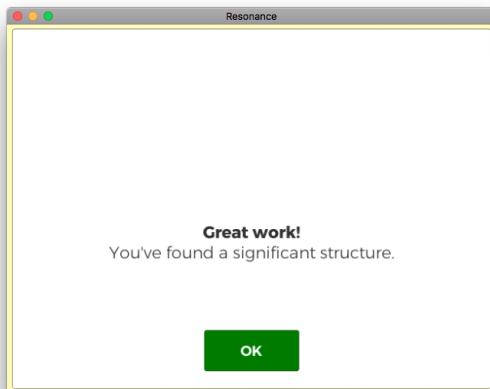


Figure 3.8: The success dialog, after more than one attempt.

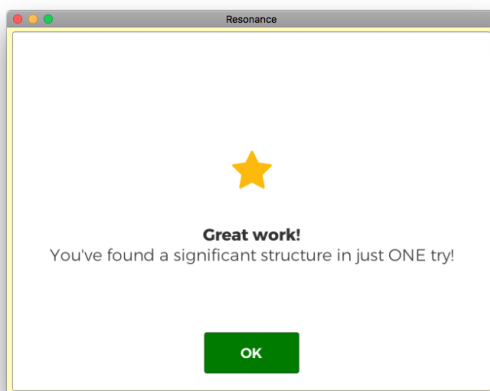


Figure 3.9: The success dialog, after just one attempt.

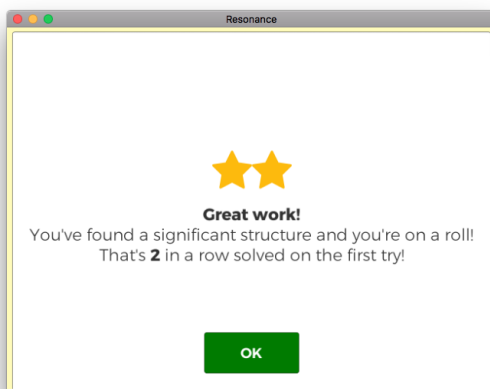


Figure 3.10: The success dialog, after two problems solved in a row in just one attempt each.

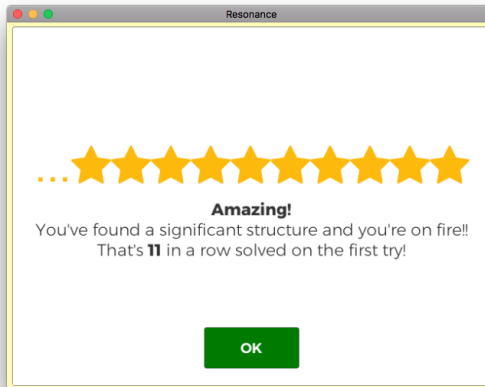


Figure 3.11: The success dialog, after eleven problems solved in a row in just one attempt each. When the one-try streak is ten or more problems, the text changes, and the number of stars stays at ten stars with a leading ellipsis ("...") implying more stars.

their errors. Khan Academy's World of Math gamified practice tools, which use a gold star to reinforce success, inspired the decision to celebrate the user's success with stars [57].

Measuring user performance by number of attempted solutions has three primary benefits. First, it does not discourage experimentation or trial-and-error—the user can draw and undo as many arrows as they want to visualize their thinking, as long as they catch their mistakes themselves before submitting. Thus, the emphasis is on self-checking rather than perfect performance, since the Software penalizes incorrect submissions rather than incorrect intermediary work. Second, users can work at their own pace, and incidentally become faster with efficiency gains from high-quality practice over time as relevant processes become automatic, rather than use sloppy shortcuts to try to beat a timer. Third, one-try streaks feed a sense of accomplishment and let the player set their own long-term performance goals, similar to the idea of competing against one's own high score.

When users dismiss the success dialog, they are still in solution review mode, though they now see the initial resonance structure and their solution linked by a double-headed arrow (see Figure 3.12). The user can examine their successful solution more closely and decide when they are ready to move on with the "Next Problem" button.



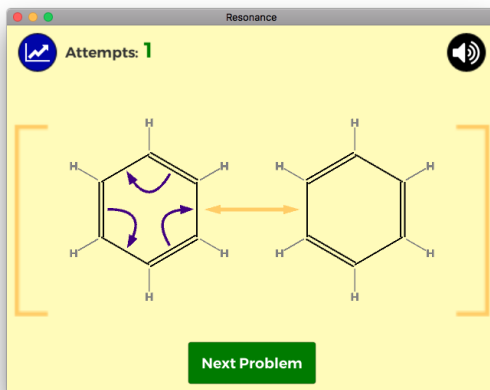


Figure 3.12: After dismissing the success dialog (see Figure 3.8), the user sees the initial resonance structure with their arrows and the result of their solution side by side.

### 3.3.2 Scaffolding and Explicit Error Feedback

Scaffolding prevents learners from feeling overwhelmed when confronting complex systems and procedures [21, 25]. Given that chemical resonance is often when organic chemistry students start putting all of their foundational knowledge together to solve problems, using digital media to scaffold the experience seemed like it may benefit some learners with less firm foundational knowledge.

The scaffolds that can be toggled are:

1. Always show the implicit hydrogens in the bond-line drawing
2. Always show the lone pairs in the bond-line drawing
3. Show implicit hydrogens when an error occurs on an atom
4. Show lone pairs when an error occurs on an atom
5. Show error type (e.g. too many electrons) and specify location of the error (e.g. highlighting the atom)

All but the last of these scaffolds specifically help learners who struggle with reading bond-line diagrams fluently, but have learned the basics of how to read them.

The Software has four tiers of decreasing scaffolding, each of which are announced by tutorial dialogs when the user moves to a more advanced tier (see Figures 3.13 and 3.14 for examples). The Software promotes users to a more challenging tier after the user solves five problems in a row from the current tier on the first try, not including

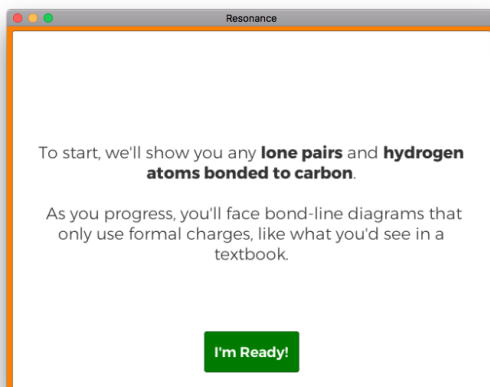


Figure 3.13: The tutorial dialog at the start of the first scaffolding tier.

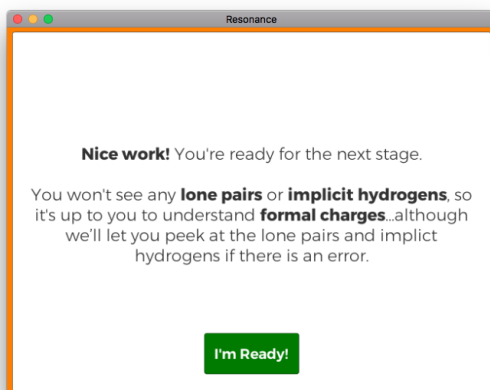


Figure 3.14: The tutorial dialog at the start of the third scaffolding tier.

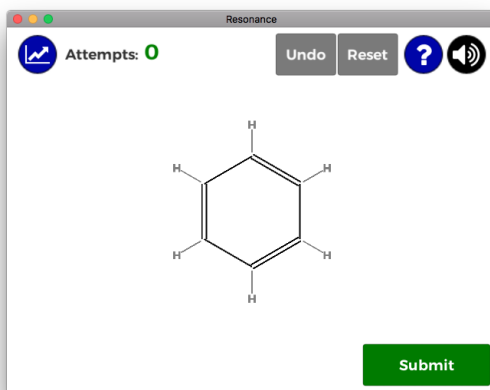


Figure 3.15: Example of a problem at the first scaffolding tier. Implicit hydrogens are shown, and if there were any lone pairs, they would also be shown.

the initial beginner problems. The designation of five-in-a-row in one try as signifying mastery is a somewhat arbitrary example of how this promotion criteria might work; with further user tests it may be that a better criteria accounts for other statistics or varies the criteria for promotion depending on the scaffolding tier.

The four scaffolding tiers in the Software, from most scaffolding to least, are:

1. All scaffolds toggled on — user sees implicit hydrogens and lone pairs always, and receives complete information about an error
2. Like previous tier except implicit hydrogens only appear for errors — user has to recognize where implicit hydrogens are hiding when drawing arrows
3. Like previous tier except lone pairs also only appear for errors — the non-error state of the drawing looks like a normal bond-line drawing
4. Like previous tier except implicit hydrogens and lone pairs are never shown — the user always works with normal bond-line drawings
5. Like previous tier except errors are generic — if user's solution is not a valid resonance structure, user must find the error themselves

The last tier represents a more exam-like experience in that the learner must diagnose their own error. Though reduced feedback may impede initial understanding, learners eventually need the opportunity to build their own schemata [15]. In general, less feedback forces the learner to think more deliberately about their own performance and knowledge [15].

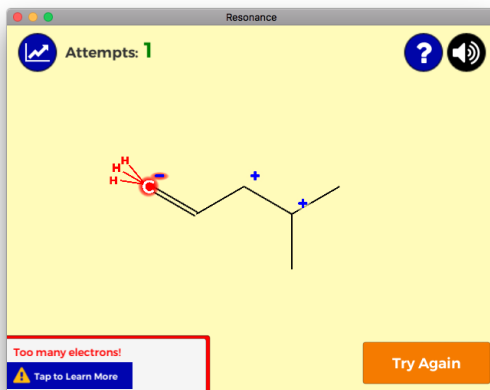


Figure 3.16: Example of a problem in a scaffolding tier where the implicit hydrogens are only shown when the Software identifies an error; other carbon atoms in this molecule do not have their implicit hydrogens visible.

However, the Software still provides feedback about whether or not a submission is significant even at the most difficult tier, to prevent learners from confusing error-containing invalid solutions with insignificant but valid solutions. Eventually, another practice tool for ranking the relative significance of resonance structures may be beneficial for more expert learners.

When an error message appears in the lower left corner, it gives the user a cue of brief text as to what the error is and, unless the user is in the most difficult scaffolding tier, the atom in question turns red (see Figure 3.7). For a user who recognizes what the error is, the brief text should be enough for them to figure out how to correct their solution; for a user who does not see how the error applies, tapping the message brings up an error dialog with more specific information about what the error means (see Figure 3.17 for an example of the dialog).

A more detailed error message requires a tap, rather than appearing right away, for two primary reasons.

First, longer textual explanations of the error require more space; some users will not read longer explanations carefully, unless they deliberately seek them out [55]. More text taking up more space may require putting the error message in a more prominent place, which could distract the user from looking at the molecule as the primary way to investigate an error in the solution. If the error text demands user attention, users who might recognize the error once they see it in the resulting structure might be denied the opportunity to see the error for themselves before reading the text.

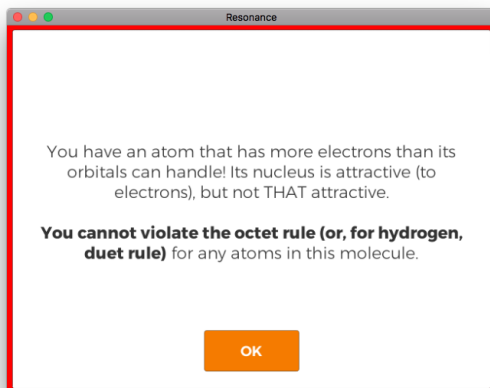


Figure 3.17: The detailed error dialog for an atom with too many electrons.

Second, the tiny amount of extra friction nudges users to process the brief cue about the error and look at the molecule first before deciding that they need more information. Access to a more detailed error message is not restricted, but it is not the default, which encourages the user to more actively investigate the error [58].

### 3.3.3 Animation as Implicit Solution Feedback

The Software animates curved arrows: drawing them, undoing them, or showing their effect on the molecule. For drawing and undoing, the animation emphasizes that the arrow appears or disappears in response to the user's input. It also provides a light metaphorical allusion to the paper method of drawing and erasing arrows. The animation for drawing goes from tail to head in the same order as the user's taps (and in the same order that the user would typically draw an arrow). For undoing an arrow, the animation reverses and goes from head to tail.

When the user submits their solution, in addition to explicit feedback the Software also gives a kind of implicit feedback about the solution by visualizing the results of the user's arrows through animation. This visualization does not explicitly judge the solution, but it gives neutral information about the impact of the arrows on the system which the learner can compare to their intended impact.

During the animation sequence after a user has submitted their solution, each arrow animates to show the effect of moving electrons from the tail to the head (see Figure 3.18). An electron pair appears during the animation to emphasize that this movement is what the arrow signifies. The correct conceptual way to imagine the imaginary movement of electrons between resonance structures is to imagine that

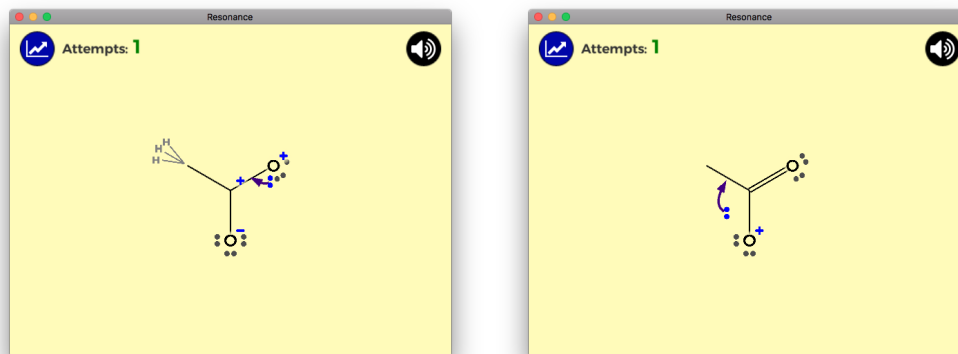


Figure 3.18: Two examples that screenshot animations-in-progress, which are illustrating the effects of the user’s curved arrows.

the movements all occur simultaneously [9]; however, this visualization shows each movement sequentially in the order that the user drew the arrows. Though less conceptually correct, sequential movement allows the user to follow along and understand the impact of each arrow on the resulting structure. Simultaneous animations may be less beneficial for a user doing practice problems, as they cannot split their visual attention effectively and would miss seeing the consequences of each arrow [23, 28].

### 3.3.4 Use of Text

In general, for longer sections of text, the Software bolds crucial actionable information to call attention to it, as in Figure 3.17 for example. Bolding text is a visual cue that calls the user’s attention to the text and signifies its importance [55]. Longer text is necessary to explain errors in greater detail, but in many places, text is brief.

For text, often less is more, because users typically do not read any interface’s text attentively, if they read at all [55]. For many interfaces, therefore, best practices include eliminating text that does not deliver immediately actionable information so that users can hone in on the information they need [55]. However, text can do more than deliver information in learning tools. Practice-based learning involves emotion more than many other tasks, with frequent failures. Text can play a role in shaping that emotional experience. For example, text can intervene in a learner’s self-reflection, by disrupting self-defeating narratives with an alternative that changes the learner’s emotional response [30, 59, 60]. Text can frame a failure experience as a learning opportunity, not defeat.

In the Software, text indicating an incorrect solution emphasizes the existence of

an error rather than a blanket statement of wrongness, which subtly implies that the solution is fixable rather than a failure. Users are given the option to "Try Again" when their solution is incorrect or insignificant. Successful solutions earn praise, and one-try successes earn additional praise. Skipping a difficult problem comes with the reminder that the user will get a chance to try the problem again in the future.

### 3.3.5 Use of Sound

The Software gives the user the option to turn the sound effects off, using conventional symbols and color changes to signify the current audio mode. For users who opt to keep sound turned on, the Software takes a conservative approach to adding sound effects, applying them to reinforce visual communication with the user.

While the user draws arrows, no sound is needed, because the user is attuned to their actions and the Software's visual response occurs in the immediate vicinity of their taps. However, undoing an arrow involves a subtle swoosh sound, added specifically to help alert a user if they accidentally remove an arrow by tapping on it so that they can recover from that unintended action. The swoosh also provides auditory feedback for tapping either the "Undo" or "Reset" button to reinforce that the Software is responding to the user's action, even though the visual response is not in the immediate vicinity of the buttons.

Similarly, a subtle sound occurs when the user clicks a button that triggers a dialog, such the help, statistics, or error detail dialogs. This sound reinforces that the dialog is the result of the user's action.

After the user submits their solution, sounds during the animation sequence call attention to each electron movement and emphasize that the Software rather than the user is manipulating the molecule for the duration of the animation.

A unique sound also calls attention to dialogs that notify the user that they have progressed to a new scaffolding tier. This helps alert the user to the upcoming changes in their experience.

The sound that does the least for communicating practical information is the celebratory sound that the Software plays when notifying the user that their solution is valid and significant. Though it does reinforce that the success dialog is different than other dialogs, the primary purpose is as a rewarding stimulus for operant conditioning and for fostering self-efficacy. Similar to the gold stars that reward one-try streaks discussed previously in 3.3.1, this congratulatory use of sound also owes some inspiration to Khan Academy's gamified The World of Math, which uses only two

sounds purely for reinforcing success [57].

### **3.3.6 Summary**

This section discussed how user-facing design decisions support desirable high-quality practice habits and attempt to build the user's self-efficacy through use of scaffolding, visualizations, feedback, phrasing, and sound.

## **3.4 Formal Usability Test**

A usability test, in which users individually interact with a prototype and attempt specified tasks, helps identify what works in the current prototype, helps uncover what does not work, and provides some insight into how to improve future iterations [55].

This section describes how the findings of a formal usability test during the Software's development informed subsequent changes, with a brief description of the study procedure. For a more in-depth description of the usability test, see Appendix A.

### **3.4.1 Procedure Overview**

The widely used number of users for a usability study—five—may be too generalized for particular cases, but the main idea behind such a small number is to test new prototypes frequently and improve each iteration rather than test one prototype with many users [22, 55, 61].

Various iterations of the Software underwent informal user tests, but one formal usability study specifically sought out five users who were more like the intended audience: the proxy users recruited from the Hanover, NH area had taken organic chemistry within the last four years and were not friends of the Software's creator. The purpose of this usability study was to identify issues with the prototype of the Software at the time of the testing, with users from a population similar to that of the future utility study, in order to address usability issues in the prototype prior to the utility study.

The user testing involved the user "thinking aloud" while interacting with the prototype on an iPad for five minutes, under observation; then demonstrating features of the prototype in response to verbal prompts; then verbally answering questions about features of the prototype after the investigator performs an action; and then filling out a questionnaire on Qualtrics. Observations during iterative user testing



typically yield more useful insight than questionnaires [62], though the questionnaire provided supplementary information about what was observed.

The verbal prompts were:

1. Can you show me how to push electrons from an atom to a bond?
2. Can you show me how to undo that arrow?
3. Can you show me how to check a solution?

The demonstration sessions verified whether the participant had a basic understanding of how to use the prototype: recognizing atoms and bonds, drawing an arrow, undoing an arrow, and checking a solution. Note that at the time of testing, the "Submit" button described in Section 3.3 was labeled "Check Solution" instead.

The verbal questions in response to investigator actions were:

1. Investigator makes two arrows, and points to the one made first. Asks, "Can you tell me how to undo this arrow without undoing the other one?"
2. Investigator makes an arrow whose source is a single bond and taps "Check Solution" button. Asks, "Can you tell me why this arrow is red?" and follows up with: "How did you know?" then "What could you do if you didn't understand why it was an error?"
3. Investigator makes an arrow whose source and target are too far apart. Asks, "Can you tell me why my arrow disappeared?" and follows up with: "How did you know?" then "What could you do if you didn't understand why it disappeared?"
4. Investigator makes many arrows. Asks, "Can you tell me how to start over without undoing each arrow individually?"
5. Investigator points to solution checks count. Asks, "Can you tell me what this green number represents?" and follows up with: "Can you tell me what the number beneath it represents?" (the user's solution checks per level average). (These user interface elements were changed in later versions of the Software; see Figure 3.19 to see what practice mode looked like in this prototype.)
6. Investigator solves the problem to move to the next problem, and then solves the new problem in one go and taps "Check Solution" button. Asks, "Can you tell me what the star represents?"

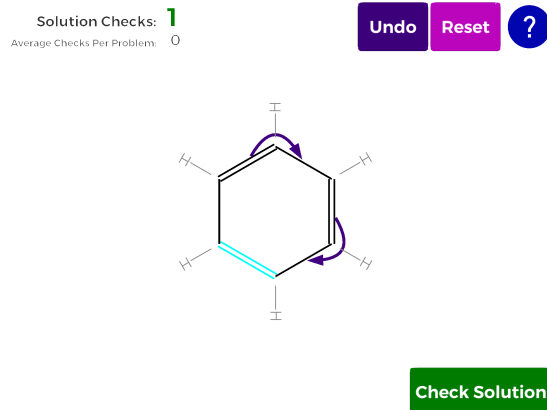


Figure 3.19: Example of practice mode in the usability study prototype, with a bond selected.

These questions tested whether the user understood more advanced features, including getting more information about an error. It also provided an opportunity to introduce the user to the one-try star, if the user did not previously solve a problem in one try.

The questionnaire first asked users to rate aspects of the prototype in a multiple-choice series of questions, simply to get a sense of how users compared to each other when reviewing the observations and questionnaire results from each session.

After the page of multiple-choice answers, the questionnaire moved into open-ended questions about specific features and perceived problems.

### 3.4.2 Results

Issues observed during initial unguided use:

1. A majority of participants first attempted to drag to draw arrows. Two participants did not immediately then try tapping.
2. A participant mistook "H" (hydrogen atoms) for "I" due to their orientation and got confused, briefly. (See Figure 3.20 for an example of the rotated "H" symbols.)
3. A participant tried to use the "Undo" button to deselect their most recent selection, rather than undo an arrow.
4. Two participants verbally noted that the lone pair electrons looked somewhat messy. (See Figure 3.20 for an example of the unappealing representation of lone pairs.)

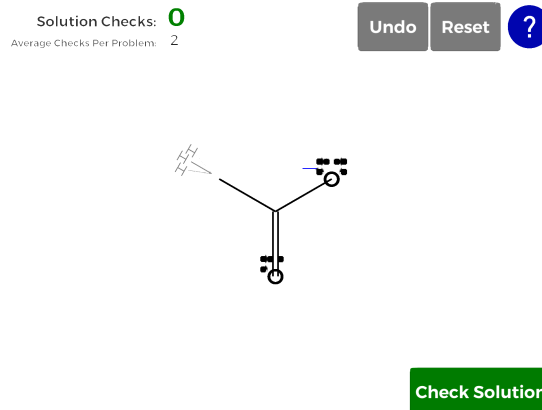


Figure 3.20: Example of practice mode in the usability study prototype.

- Two participants verbally indicated that they prefer "Submit" to "Check Solution" as the latter implied to them the idea of giving up and seeing someone else's solution. One of these participants kept changing their arrows because they thought that the tool would automatically inform them that they had arrived at a correct solution. (See Figure 3.20 to see the "Check Solution" button, whose label is "Submit" in later versions.)
- A majority of participants verbalized that the error messages and explanations were helpful.
- A participant became frustrated with a problem and verbally expressed a desire to be able to skip it.
- Some participants tapped non-tappable parts of the screen (whitespace, formal charges, hydrogens bonded to carbon) at first; two were particularly persistent about tapping the hydrogens.
- One participant verbally indicated that being able to draw an arrow from a single bond (a forbidden move) seemed odd.

Results of demonstration and explanation sessions:

- All five participants demonstrated basic proficiency when asked to show how to push electrons from an atom to bond, undo an arrow, and check a solution.
- To undo the arrow when asked, two participants tapped the arrow, one participant tapped the arrow's source, and two participants used the "Undo" button. All approaches were correct and successful.

3. All participants successfully explained how to undo an arrow without undoing a subsequent arrow, although one participant noted that it was a guess (this participant had only used the "Undo" button previously).
4. All participants successfully explained the cause of red arrows and disappearing arrows, referencing the error message and the "Learn More" detailed message as resources for someone who did not recognize the cause of the error.
5. All participants successfully chose the "Reset" button to clear a problem of multiple arrows at once.
6. All participants correctly explained the significance of the solution checks count; one participant incorrectly interpreted the average solution checks as referring to the average checks on the problem at hand across all users, rather than referring to their own average number of checks across all prior problems.
7. All participants who had solved a problem in one try explained the significance of the gold star as representing solving a problem without any incorrect solution checks; one participant, who had not solved a problem in one try on their own, agreed that it signified that the solver had done a "good job" on the problem, but did not explicitly connect it to solving the problem in one try.

Issues raised by the open-ended questions:

1. One participant noted that it was frustrating to not be able to skip a difficult problem, and that they expected to be able to do so and to view someone else's solution (as one can do with a textbook's problem). Another participant also briefly described a wish to have the option to see a solution after trying an incorrect solution.
2. The participant who had confused the hydrogen "H" symbols with "I" noted that issue. Another participant noted that showing the hydrogens bonded to carbon made the molecule look more complicated and confusing (indeed, that is why organic chemistry favors implicit hydrogens in bond-line diagrams, though it requires greater fluency).
3. Three participants noted that they initially sought to drag rather than tap to draw arrows.
4. Two participants noted that they expected or preferred a "Submit" button to a "Check Solution" button.

5. Two participants noted that they were dissatisfied with how lone pair electrons looked.
6. The participant who had expected automatic confirmation of a correct solution noted that they had been confused by the "Check Solution" button.

Other insights from the open-ended responses:

1. A majority of participants noted that the current solution checks count could tell a user how well they were doing on a particular problem, as a way to self-check understanding or flag problems for follow-up. One participant noted that this was especially important for preparing for an exam, when the only way to check your answer was to self-check.
2. A majority of participants noted that the average solution checks could assess overall success or mastery.
3. A majority of participants wrote that the gold stars brought some satisfaction or confidence when they solved a problem in one try. Two participants noted that when they realized they could get more stars for streaks of one-try successes, they challenged themselves to catch their own errors.
4. One participant expressed a wish to be able to pick problems.
5. One participant expressed a wish to make the user identify all contributors rather than just one contributor to the resonance hybrid.
6. One participant noted that this tool seemed like it would save time as an alternative to having to draw the molecule and then the solution.
7. One participant briefly expressed a wish to be able to take notes, without much detail about the type of note.
8. One participant expressed a wish to save a difficult problem for later and come back to it.

### 3.4.3 Impact on Development

The usability study influenced the change from "Check Solution" to "Submit" as the button text for entering review solution mode, though all functionality remained the same. "Submit" made it more clear that the user would not have their work

automatically interrupted when they solved the problem, and "Check Solution" could have misleadingly suggested "Check The Answer Key Solution" instead.

One user wanted to be able to skip a problem, which influenced the decision to include that feature in a later iteration of the Software. However, the request to show an answer key solution raised questions about dealing with multiple solutions, as did one user's request to require all solutions; depending on the context in which the Software exists, how to best handle multiple solutions may vary, and an answer key solution was not added to subsequent versions of the Software. Similarly, abilities to pick problems and save problems were not included in later version of the Software, since the best way to handle those features would also depend on the environment in which the Software exists.

The issue users had figuring out what is interactive and what is not influenced the decision to make everything that is part of the molecule interactive: formal charges and hydrogens bonded to carbon previously were not interactive. Later versions of the Software had it so that selecting a formal charge or a hydrogen bonded to carbon would select the selectable atom that those components were affiliated with, as had already been the case for lone pairs on an atom.

For future work on mobile platforms, it seems that covering both tapping and dragging as ways to draw arrows may be worthwhile. For the purposes of developing the Software for the utility study, however, which was to be web-based on desktop browsers, such development did not occur.

The issue of rotated hydrogen "H" symbols was an artifact of rapid development and an already known issue before user testing; it was fixed prior to the utility study. Similarly, a graphics overhaul for lone pair electrons was also already planned.

## 3.5 Technical Implementation

This section outlines the technical implementation of the Software. As the Software is a proof-of-concept prototype, low-level details and unimportant implementation decisions are omitted.

The Software was developed with Unity, using C# for custom scripts. For a reader unfamiliar with Unity, note that most scripts are classes whose instances act as components on Unity `GameObject` instances, and `GameObject` instances have hierarchical relationships with other `GameObject` instances. Though the Software contains 28 custom classes, in part due to working with Unity `GameObject` prefabs (preconfigured `GameObjects`, of which the Software has 35), technical discussion will

focus only on notable classes that have unique relevance to the Software.

### 3.5.1 Overview of Model Classes

To represent the molecule in the practice problem, multiple `GameObject` prefabs represented different parts. Scripts attached to these prefabs include model classes `Molecule`, `Atom`, `Bond`, and `Electrons`, which represent the actual data of the molecule in the problem.

**Molecule** — This class has a list of all `Atom` instances included in the molecule, a set of all `Bond` instances, and a list of substructures as a 2D list of `Atom` instances. `Molecule` instances have a state, which is a string representation of the configuration of electrons in the molecule that includes bond strengths, implicit hydrogens, and lone pair electron locations as data. The state string is used for comparing user solutions to the initially given resonance structure to ensure they are not the same structure.

**Atom** — Instances of this class that are the same element share the same static element symbols, expected valence electron counts, and valence orbital counts, since real-world atoms of the same element share those characteristics. `Atom` instances also have a dictionary where other `Atom` instances are keys and `Bond` instances are values, such that an `Atom` knows if it is bonded to another `Atom`. `Atom` instances may have an implicit hydrogen count if their element is carbon, in which case the implicit hydrogen count is also counted as part of the bonding electrons. An `Atom` instance can calculate all of the bonding electrons it has by iterating through values to get each `Bond` instance's bond order, which represents numerically whether the represented bond is a single, double, etc. bond, and summing up the number of implicit hydrogens. `Atom` instances also keep a list of `Electrons` instances that represent nonbonding electrons on the `Atom`. Therefore, `Atom` instances can calculate their formal charge by summing the number of bonding and nonbonding electrons, and comparing it to their expected valence electron count. `Atom` instances can also calculate whether or not their orbitals are full.

**Bond** — Instances of this class keep an array of two `Atom` instances, which represent the bonded atoms. `Bond` instances also track their bond order, for which 1 represents a single bond, 2 represents a double bond, and so forth. (Users can propose solutions that have unrealistically large bond orders.) `Bond` instances can increase or decrease their bond order as needed to represent a new electron configuration in the molecule.

**Electrons** — The Software only supports lone pair electrons, so each instance of

Electrons has an electron count of 2. Electrons also track the Atom instance that they are affiliated with.

### 3.5.2 Overview of Interactive Molecule Classes

As noted in the previous subsection, 3.5.1, crucial prefabs representing the molecule have their own model classes for relevant data. The same prefabs have component classes responsible for handling visuals and interaction, which have access to the corresponding model class instance.

For Atom, Bond, and Electron model classes, there are corresponding AtomInteractive, BondInteractive, and ElectronInteractive classes, which inherit some common constants and methods. Additionally, a GameObject with an Atom script may have child GameObjects with FormalCharge or ImplicitHydrogen component classes, which also inherit from the same abstract class as the other interactive classes. All of these interactive classes control the appearance of the GameObject they are attached to, and handle the appropriate selection-related logic triggered on user interaction. Any extra responsibilities of these classes are explained below.

**AtomInteractive** — This class positions and rotates any ElectronInteractive, FormalCharge, and ImplicitHydrogen child instances, taking into account the number and placement of BondInteractive instances it is involved with. It also changes appearance if there is an error in its Atom instance when called upon to do so, and controls how and when its child instances appear.

**BondInteractive** — This class maintains a list of Unity LineRenderer objects, increasing and decreasing their number to reflect the bond order of its Bond model instance. This class also lays out the LineRenderer objects appropriately for visualizing the bonds.

**ElectronInteractive** — In response to user interaction, this class passes the interaction on to the AtomInteractive instance on the same GameObject as the Atom instance that its Electrons instance stores.

**FormalCharge** — In response to user interaction, this class passes the interaction on to the AtomInteractive instance in its parent GameObject.

**ImplicitHydrogen** — In response to user interaction, this class also passes the interaction on to the AtomInteractive instance in its parent GameObject.

The Molecule model class is a special case: it is attached to an empty GameObject with no visual component. However, another empty GameObject contains a Puzzle component class that all interactive classes have access to, and the Puzzle instance



has access to all the interactive class instances through the Molecule instance. The Puzzle instance contains booleans that let interactive class instances know whether the user can select parts of the molecule (e.g. practice mode vs. solution review mode). This class will be discussed in more detail in the next subsection, 3.5.3

Additionally, there is a **CurvedArrow** interactive class that does not inherit from a common abstract class with the other interactive classes. It draws a curved arrow between a source (tail) and a target (head) by animating the head from source to target. It removes the arrow by animating the head going from target to source. It also pushes the arrow with an animation that moves the tail from source to target and spawns a short-lived prefab with a lone pair electrons visual. When the user interacts with a CurvedArrow in practice mode (as determined by a Puzzle instance's boolean), the CurvedArrow instance animates the arrow's removal and then self-destructs.

### 3.5.3 Overview of Manager Classes

The following classes instantiate and control other classes.

**Puzzle** — One instance of this class is created per problem, and is destroyed after the user solves or skips the problem. This class contains data about the visual and interactive state of the current problem, such as the current mode, whether an error should be displayed, what type of error should be displayed, and scaffolding settings for the current problem. A Puzzle instance also stores the initial state data of the Molecule before the user interacts with it, so it can verify that the user's submission is not the initial resonance structure. An instance of this class also stores the CurvedArrow instances that the user has created; it starts their animations when the user submits a solution and reverses their effects when the user tries again. This class orders all Atom instances, through their parent Molecule instance, to check themselves for orbital errors and report back with a boolean. This class validates the user's solution, though for the resonance-specific significance determination, it calls on a function in a static ResonanceUtilities class.

**LevelManager** — There is only one instance of this class. This class creates Puzzle and Molecule instances for each problem, and links them together; it destroys them when the user solves or skips the problem. It also manages all of the non-molecule user interface components, such as buttons and dialogs, which are their own prefabs. The LevelManager connects these components to the appropriate function in the current Puzzle instance. The LevelManager handles switching between practice mode and solution review mode, which includes triggering the Puzzle's solution check

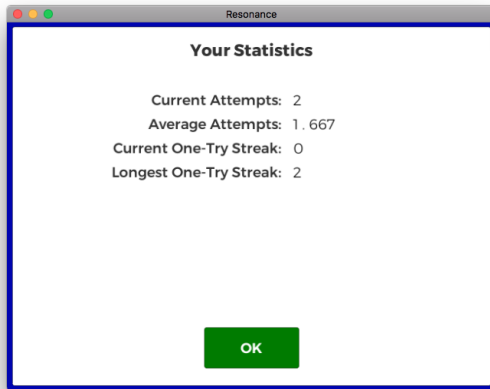


Figure 3.21: The user statistics dialog that appears when the user taps the button in the upper left-hand corner. The `ProgressManager` passes the required user data to the `LevelManager` who passes it on to the dialog's script.

function and waiting on it to finish before triggering the appropriate success or fail state response.

**ProgressManager** — There is only one instance of this class, and it manages the classes responsible for user data and problem selection, and also determines when to move the user to the next scaffolding tier. The `ProgressManager` receives data about the outcome of each submitted solution from the `LevelManager`, and updates relevant user data. The `ProgressManager` outsources user data maintenance to a `UserData` class instance that updates performance statistics, problem history, and certain non-performance statistics. The `ProgressManager` checks the `UserData` instance after each problem to decide whether to promote the user to the next scaffolding tier, and if so, updates the tier in the `UserData` instance and triggers a tutorial dialog about the promotion. The `ProgressManager` passes on user data to the `LevelManager` whenever the user interface needs it (for displaying statistics, streaks, attempted solution count, etc.). When the `LevelManager` is ready to move on to the next problem, the `ProgressManager` passes on the new problem data and the user's current scaffolding tier. The `ProgressManager` outsources selecting the next problem to a `ProblemSelector` class instance, which weights problems so that the user never sees the same problem twice in a row, so that problems solved in one try are weighted less than those that took more tries, and so that skipped problems are weighted higher to likely reappear again soon (the `ProblemSelector`'s initial weights per problem can be set to favor easier problems first, and problems can be set as "beginner problems").

### 3.5.4 Representing Molecule Problem Data

Problem data is represented inertly as strings that correspond to a resonance structure of a molecule. Though existing software and file types may use text representations of chemical compounds, some are proprietary, and though there are dozens of open formats [63], other representations are not very human-writable because they are intended for purposes such as cheminformatics that require much more data than a student practice problem does [64, 65].

For the purposes of allowing a human to quickly extract structure information from drawings to add them to the Software, I developed a simplified encoding scheme that can represent a smaller subset of molecules; this encoding scheme and its decoder could be expanded upon to include more types of molecule substructures. The main benefit of the encoding scheme is that it only encodes information that is explicitly available in the bond-line diagram—atoms, bonds, and formal charges—and lets the decoder interpret the information to add implicit hydrogens and lone electron pairs, rather than rely on an error-prone human interpreter. These string representations are also much more compact than other representations.

The Software's decoder can handle two types of substructures: rings (where atoms are bonded together as a regular polygon) and chains (where atoms are bonded in a linear sequence). The Software's decoder only recognizes one atom as the "anchor" of the substructure—determining how the substructure should be positioned relative to other substructures by attaching it to that atom—so rings that share a bond (rather than just one atom) and chains that are anchored at both ends are not supported by the decoder. For chains, the decoder also does not support having an anchor somewhere other than an endpoint. The decoder does not support decoding formal charges less than -1 or greater than +1, but the encoding scheme can support multiplying the symbol to represent those cases (e.g. "N--" for a negative formal charge of 2 on a nitrogen atom).

The encoding scheme has a few simple rules for ordering the substructures that represent a molecule:

1. Encode all rings before encoding any chains
2. If substructure A is anchored by an atom in substructure B, substructure B must be encoded before substructure A
3. Always encode chains as continuous structures (so that the decoder draws them correctly as a continuous zigzag)

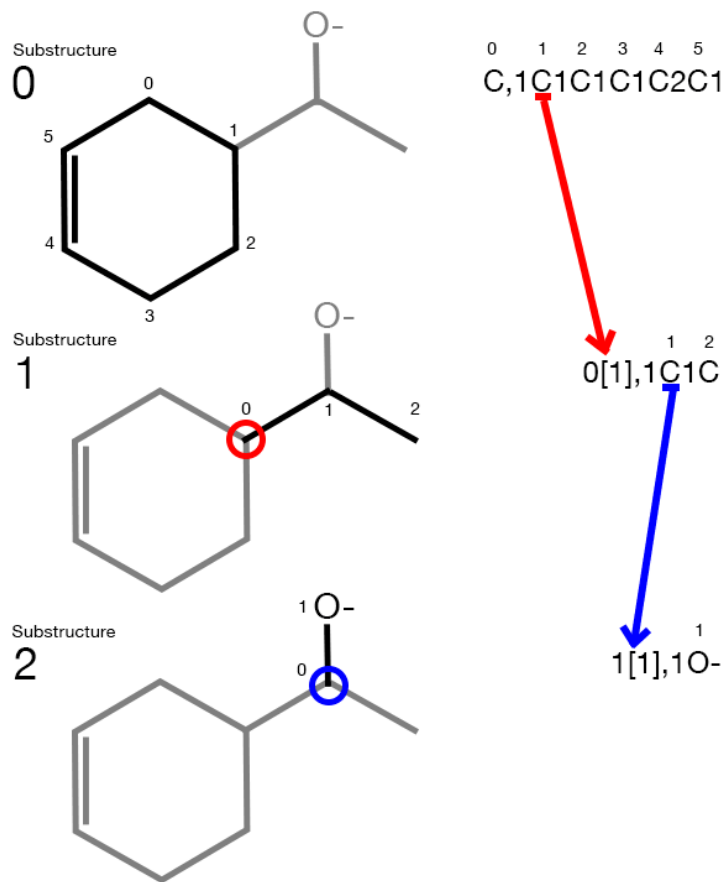


Figure 3.22: An example of the encoding scheme, showing how each substructure is identified and encoded.

4. Separate substructure encodings with a newline character

A substructure is encoded as follows:

1. Begin with the anchoring atom:

- If the substructure is unanchored (first substructure), simply use the element symbol, followed by a formal charge if any, e.g. "C" or "N<sup>+</sup>"
- If the substructure is anchored, put the 0-indexed index of the substructure containing the anchor followed by square brackets containing the 0-indexed index of the atom within that substructure, e.g. "1[3]" indicates the fourth atom of the second substructure
- Whether anchored or unanchored, add a comma (",") after the first atom

2. From the anchor, trace the substructure in a continuous path through each atom by following the bonds, adding the bond order number to the string for each bond (e.g. "1" for a single bond) and adding the element symbol followed by a formal charge if any for each atom (e.g. "O-")

Figure 3.22 illustrates the encoding process with an example.

The encoding scheme implies the substructure: for a ring substructure, the last character will be an integer for the bond connecting the last atom back to the first, while for a chain the last character is part of an atom (alphabetic) or its formal charge ("+" or "-"). For a ring substructure, the decoder assumes that the anchor is at the top of the ring and builds the ring clockwise from there. The decoder can determine how many lone pair electrons to add to each atom based on whether it has a neutral or formal charge; for carbon, it will first add implicit hydrogens (there are set patterns for the number of bonds vs. number of lone pairs for C+ and C- carbon [9]).

The set of encoded problem strings are stored in a list, such that the index identifies the problem. When the LevelSelector chooses an index based on its weighting scheme and passes it to the ProgressManager, the ProgressManager uses that index to represent the problem in relevant user data. When the LevelManger joins the Puzzle and Molecule instances to create a problem, the Puzzle instance passes the problem string to the Molecule instance, which calls the static decoder function to parse the string and add substructures to the Molecule instance.

Once the Molecule instance has all of the substructures represented as lists of Atom instances, the static class MoleculeLayoutMaker positions each substructure in order, using the position of the first Atom instance in the substructure's list to anchor that substructure. The MoleculeLayoutMaker positions the Atom instances within each substructure based on substructure type (chain if the first and last Atom instance in a substructure list are not bonded to each other, with fixed angles to create the zigzag of bond-line drawings; ring if they are bonded, with angles determined by number of Atom instances in the substructure). Conventionally, bonds should be shown as far from each other as possible [9]. So, for example, if an Atom instance that anchors a chain already has two bonds which are not in the chain substructure, the chain should have its first bond along the bisector of the convex angle formed by the two bonds so that it is as far apart from the other two bonds as possible. After all substructures are laid out, the MoleculeLayoutMaker centers the entire structure.

## 3.6 Chapter Summary

This chapter covered major decisions in design and technical implementation that contributed to the Software, including the impact of a formal usability study.

# Chapter 4

## Utility Study

This chapter provides an overview of the utility study conducted on the Software. For more detailed information about the utility study setup, procedure, questions, and results, see Appendix B (which repeats some of this chapter). For details about how margin of error (MOE) was calculated using two-tailed t-values, please see Appendix F.

### 4.1 Procedure Overview

To evaluate the Software, 21 participants were recruited for a study about perceived usefulness of the Software and other types of academic resources. The anticipated user for the Software would be someone learning or reviewing chemical resonance, but the study used proxy users who had the prerequisite domain knowledge. All participants self-reported that they taken first-semester organic chemistry or equivalent at an accredited four-year college or university in the last four years in the United States.

Participants filled out an approximately 15-minute questionnaire. The questionnaire had four parts, listed here sequentially as they appeared:

1. Questions about prior experiences with different types of academic resources and media
2. Evaluation of different types of academic resources
3. A page with an embedded version of the Software; the button to continue in Qualtrics appeared after five minutes on the page and participants were invited to use the Software however they wanted to
4. Evaluation of the Software and follow-up questions

For the two evaluation sections, eight dimensions of utility were presented as follows (adapted from [66]):

1. Useful
2. Efficient (speed, time cost-effectiveness)
3. Effective (improvement to future performance and learning)
4. Accessible (availability, monetary cost, ability to use)
5. Convenient (ease to use, ease to learn how to use, time/place/context convenience)
6. Motivating (encouragement, confidence, frustration)
7. Clarifying (explanations, information, assistance, confusion, comprehension)
8. Appealing (comfort, stress, intimidation, aesthetics, experience)

For the sake of space, the parenthetical parts of the above dimensions will be omitted from tables in this thesis, but the labels in the questionnaire included the parenthetical parts, which attempted to make abstract terms more precise [66].

For the section evaluating academic resources, a checkbox table had rows labeled with academic resource types. Its first column was labeled with "Did Not Use or N/A" and subsequent columns were labeled with each of the previously mentioned dimensions of utility. Participants were instructed to check any boxes that described their experiences while studying organic chemistry with the resource of each row. For the section evaluating the Software, the column labels stayed the same (except "Did Not Use or N/A" since all participants had used the Software), and participants checked boxes if they thought that the Software would meet that description for the type of user labeled in each row (e.g. "For someone revisiting organic chemistry while studying for the MCAT"). The checkbox table appeared to be more efficient for participants than attaching each checkbox to a Yes/No multiple-choice question, and keeping the questionnaire within 15 minutes was a concern for recruitment purposes. The binary evaluation of each dimension avoids letting participants select a neutral or "neither" response and thereby forces a decision [66].



## 4.2 Results

### 4.2.1 Results of Prior Experience Questions

The first question asked whether participants had ever used an online tool to supplement academic learning, and the majority of participants reported having used an online tool (15 responses/71.4% "Yes" vs. 6 responses/28.6% "No", MOE 21.1%).

The next question gave participants the opportunity to expand upon their answer, either sharing which resources they used and why, or explaining why they had never used an online tool.

Of the participants who had used an online tool and chose to elaborate, a majority had used Khan Academy and one in three had used YouTube for video explanations of academic material in some subject. One in five of the participants who had used an online tool specifically mentioned watching videos for help with organic chemistry on either Khan Academy or Youtube, either to clarify confusion about a concept or to prepare for an exam. One participant noted that they had used Khan Academy but found the length of the videos unappealing (unenthusiastic about watching a video "8 minutes or more").

A caveat: the wording of the question listed Khan Academy, Duolingo (which 7 of the 15 "Yes" participants mentioned), and Coursera as examples of online tools, so resources other than those three examples may be more likely to have been omitted by participants when elaborating.

Beyond Khan Academy, YouTube, and Duolingo, other resources received no more than one mention in the "Yes" elaborations.

Of those who answered "No" and chose to elaborate, participants thought that normal coursework and course materials sufficiently met their learning needs, had concern about a mismatch between outside resources and course expectations (which one "Yes" elaboration also noted was a drawback), had difficulty finding appropriate material, or perceived books as having greater detail and depth than online tools.

Participants were then asked:

In general, what makes a digital learning tool useful? Are there particular types of learners, circumstances, or subjects/skills for which digital tools are especially useful or not very useful? If so, why?

Characteristics of digital tools cited as useful by one or more participant: easy to use, easy to navigate/well-organized, visual, provides feedback, user-friendly, video

playback options, has quizzes or practice problems with rapid feedback. Characteristics cited by no more than one participant: images and graphs, easy to access, animations, typically cheaper costs, can take notes directly on page, updates/corrections immediately rather than waiting for next edition, progress bars if done in a satisfying way.

For types of learners, ten responses cited appealing to different learning styles, especially visual and audio, as an advantage to digital tools. One caveat to note is that the phrase "types of learners" in the question may have evoked the idea of sensory learning styles (which science does not support [67]), biasing responses. Other learner-related benefits cited by more than one participant include self-pacing and extra exposure to learn material. One participant noted that digital tools can be less stressful than a classroom setting for students who may be intimidated by their peers or professor when initially learning the material.

Circumstances cited by more than one participant in which digital tools are useful: reviewing material, replacing or supplementing inadequate course materials or experiences, providing alternative explanations. Circumstances cited by no more than one participant: rote memorization, demonstrating problem-solving or procedures, exploring nuance, high-level overviews, filling in gaps in foundational knowledge when course moves too quickly or assumes prior knowledge, finding more examples and problems, theoretical reading at leisure.

Subjects and skills cited by more than one participant in which digital tools are useful: STEM subjects, spatial reasoning or systems. Math, chemistry, and physiology received specific mentions among the STEM subjects.

Reasons or cases where digital tools are not useful:

1. inadequate for problem-based courses including organic chemistry (1 participant flat-out described as not useful, 1 participant described as not as good as traditional ways of doing solo practice) (2 participants)
2. can replace or discourage professor-student interactions (1 participant)
3. poor substitute for in-person learning experiences (1 participant)
4. if learner prefers physical book in hand (1 participant)
5. if learner is already very familiar with material (1 participant)
6. inadequate for primary instruction (1 participant)

The next question asked about participant preference for digital media or paper media when doing academic tasks, including whether that preference changed depending on the circumstance. The majority of participants had mixed preferences for different types of subjects and tasks, although a few participants had always or near-always preferences for one type of media.

Reasons or cases where digital media preferred over paper media, as reported by more than one participant:

1. search functionality
2. note-taking capabilities, especially for speed or essay preparation
3. efficiency: faster and easier to access content
4. mobility
5. preferred for writing papers
6. preferred for humanities subjects

Though only 6 reasons or cases are listed above, participants shared a total of 19 reasons or cases in favor of digital media. The reasons or cases mentioned by no more than one participant: preferred for subjects involving 3D space, preferred for reviewing material, preferred to mix media for reviewing and summarizing, preferred for self-study of languages, preferred for programming and computer science, digitizing paper notes for easy and mobile access, typically lower cost, preferred for biology, shareable, preferred for presenting specific aspects of material, better for one's back (vs. carrying around textbooks), accessing databases, preferred for reinforcing activities, preferred for feedback and identifying what to work on.

Reasons or cases where paper media preferred over digital media, as reported by more than one participant:

1. preferred physicality of writing and/or thought it helped them remember better
2. preferred physicality of reading and/or thought it helped them remember better
3. easier to draw, which may be necessary for some subjects
4. preferred for math, for some because digitally writing equations is difficult

5. preferred for chemistry (including organic), for some because digitally drawing compounds and writing chemical equations is difficult
6. less distracting and more focused
7. can see own thinking better
8. preferred for longer readings

Though only 8 reasons or cases are listed above, participants shared a total of 21 reasons or cases in favor of paper media. The reasons or causes mentioned by no more than one participant: liked physical control, disliked highlighting and typing on computer, easier on the eyes, liked physicality of books, felt more personal, preferred to avoid screen time, preferred for hard sciences in general, preferred to mix media for reviewing and summarizing, more mobile (e.g. not needing a power outlet), graphing and tangibly interacting with data, printing out slides, preferred for assignments and tests.

#### **4.2.2 Results of Evaluation of Other Resources**

Table 4.1 shows the number of participants who did not check "Did Not Use or N/A" for each resource as participants who did use the resource, by count and by percentage. In this subsection, "users" and "participants" are not interchangeable: users have used a resource while not all participants have.

Not all 21 participants used all of the resources, therefore Table 4.2 gives the percent, out of those who had used it, who checked the box for each utility metric.

Table 4.1: Percent of users of each type of resource for organic chemistry coursework and learning, from the utility study. Margin of error included.

<b>Resource</b>	<b>Number of Users (Max. 21)</b>	<b>Percent Did Use</b>
Textbook	18	85.7% $\pm$ 16.3%
Workbook	13	61.9% $\pm$ 22.6%
Teaching Assistants	11	52.4% $\pm$ 23.3%
Professor Office Hours	12	57.1% $\pm$ 23.0%
Formal Study Group	9	42.9% $\pm$ 23.0%
Informal Study Group	17	81.0% $\pm$ 18.3%
Tutor	4	19.0% $\pm$ 18.2%
Attended Class	21	100.0% $\pm$ 0.0%
Attended Lab	20	95.2% $\pm$ 10.0%
Online Lectures	8	38.1% $\pm$ 22.6%
Website	10	47.6% $\pm$ 23.3%
App	5	23.8% $\pm$ 19.9%
Other	7	33.3% $\pm$ 22.0%

Table 4.2: Evaluation of academic resources from the utility study. Percentages indicate the percent of users who did check the checkbox for the utility dimension, out of all users who did not check "Did Not Use or N/A" for the resource.

Resource	Number of Users (Max. 21)	Useful	Efficient	Effective	Accessible	Convenient	Motivating	Clarifying	Appealing
Textbook	18	61.1% ± 24.9%	11.1% ± 16.1%	44.4% ± 25.4%	50.0% ± 25.6%	38.9% ± 24.9%	11.1% ± 16.1%	55.6% ± 25.4%	16.7% ± 19.1%
Workbook	13	53.8% ± 31.4%	23.1% ± 26.5%	53.8% ± 31.4%	53.8% ± 31.4%	38.5% ± 30.6%	30.8% ± 29.0%	30.8% ± 29.0%	15.4% ± 22.7%
Teaching Assistants	11	90.9% ± 20.3%	18.2% ± 27.2%	36.4% ± 33.9%	54.5% ± 35.1%	36.4% ± 33.9%	54.5% ± 35.1%	63.6% ± 33.9%	36.4% ± 33.9%
Professor Office Hours	12	66.7% ± 31.3%	25.0% ± 28.7%	58.3% ± 32.7%	66.7% ± 31.3%	25.0% ± 28.7%	41.7% ± 32.7%	83.3% ± 24.7%	33.3% ± 31.3%
Formal Study Group	9	33.3% ± 38.9%	22.2% ± 33.9%	44.4% ± 40.5%	11.1% ± 25.6%	33.3% ± 38.4%	55.6% ± 40.5%	33.3% ± 38.4%	33.3% ± 38.4%
Informal Study Group	17	58.8% ± 26.1%	29.4% ± 24.1%	58.8% ± 26.1%	41.2% ± 26.1%	35.3% ± 25.3%	76.5% ± 22.5%	58.8% ± 26.1%	41.2% ± 26.1%
Tutor	4	50.0% ± 80.1%	0.0% ± 0.0%	50.0% ± 80.1%	0.0% ± 0.0%	0.0% ± 0.0%	0.0% ± 0.0%	25.0% ± 69.4%	0.0% ± 0.0%
Attended Class	21	90.5% ± 13.7%	28.6% ± 21.2%	61.9% ± 22.7%	66.7% ± 22.0%	47.6% ± 23.3%	23.8% ± 19.9%	71.4% ± 21.1%	19.0% ± 18.3%
Attended Lab	20	55.0% ± 23.9%	0.0% ± 0.0%	35.0% ± 22.9%	35.0% ± 22.9%	25.0% ± 20.8%	45.0% ± 23.9%	40.0% ± 23.5%	15.0% ± 17.1%
Online Lectures	8	87.5% ± 26.6%	62.5% ± 43.4%	50.0% ± 44.7%	62.5% ± 43.3%	62.5% ± 43.4%	12.5% ± 29.6%	50.0% ± 44.7%	12.5% ± 29.6%
Website	10	90.0% ± 22.6%	60.0% ± 36.9%	70.0% ± 34.6%	60.0% ± 36.9%	60.0% ± 36.9%	20.0% ± 30.2%	60.0% ± 36.9%	20.0% ± 30.2%
App	5	80.0% ± 51.4%	60.0% ± 63.0%	60.0% ± 63.0%	60.0% ± 63.0%	60.0% ± 63.0%	20.0% ± 51.4%	20.0% ± 51.4%	20.0% ± 51.4%
Other	7	14.3% ± 35.0%	0.0% ± 0.0%	14.3% ± 35.0%	14.3% ± 35.0%	0.0% ± 0.0%	0.0% ± 0.0%	0.0% ± 0.0%	14.3% ± 35.0%

A conservative margin of error (MOE) shows that the statistical data for Tutor and Other tells little, other than that for dimensions of utility that received at least one endorsement, there exist former students of organic chemistry who perceive that dimension in that resource. The same is true for App for most dimensions, except Useful.

Resources that were "Useful" for a majority of participant users and, even taking the lower bound of the 95% confidence interval, likely would be considered "Useful" by other former students of organic chemistry who used them:

1. Teaching Assistants (90.9% of participant users, lower bound 70.6%)
2. Attended Class (90.5% of participant users, lower bound 76.8%)
3. Website (90.0% of participant users, lower bound 67.4%)
4. Online Lectures (87.5% of participant users, lower bound 60.9%)

This does not mean that these resources are more "Useful" than other resources, simply that a majority of users in the study who had used them found them "Useful" and there is reason to believe that this would likely hold in a broader population of users of that resource who are recent former students of organic chemistry in the United States.

Resources that were "Useful" for at least half of participant users and whose lower bound of the 95% confidence interval is greater than 20%:

1. App (80.0% of participant users, lower bound 28.6%)
2. Professor Office Hours (66.7% of participant users, lower bound 35.4%)
3. Textbook (61.1% of participant users, lower bound 36.2%)
4. Informal Study Group (58.8% of participant users, lower bound 32.7%)
5. Workbook (53.8% of participant users, lower bound 22.4%)
6. Attended Lab (55.0% of participant users, lower bound 32.1%)

It is likely that at least one in five users of these resources would classify them as "Useful" provided that the users are like the users who participated in the study (former organic chemistry students who took the course in the United States). Note

that digital media (Website, Online Lectures, App) excluding Other appears to be "Useful" for at least one in five users.

Even looking just at the users among participants, no resource earned "Efficient" endorsements from more than 70% of users. The only resources that earned "Efficient" endorsements from more than half of users in the study are digital media (Website, Online Lectures, App) and of those, only Website has a confidence interval lower bound greater than 20%:

1. Website (60.0% of participant users, lower bound 23.1%)

Eight resources received an "Effective" endorsement from at least half of participant users, including digital media (Online Lectures, Website, App), though only five had a confidence interval lower bound greater than 20%:

1. Website (70.0% of participant users, lower bound 35.4%)
2. Attended Class (61.9% of participant users, lower bound 39.2%)
3. Informal Study Group (58.8% of participant users, lower bound 32.7%)
4. Professor Office Hours (58.3% of participant users, lower bound 25.6%)
5. Workbook (53.8% of participant users, lower bound 22.4%)

"Accessible" similarly had eight resources receiving endorsements from at least half of participant users, including digital media (Online Lectures, Website, App), with only five holding a confidence interval lower bound greater than 20%:

1. Attended Class (66.7% of participant users, lower bound 44.7%)
2. Professor Office Hours (66.7% of participant users, lower bound 35.4%)
3. Website (60.0% of participant users, lower bound 23.1%)
4. Workbook (53.8% of participant users, lower bound 22.4%)
5. Textbook (50.0% of participant users, lower bound 24.4%)

The only resources that earned a "Convenient" endorsement from more than half of participant users were digital media (Online Lectures, Website, App), though Attended Class barely fell short (47.6%). Only two resources had a confidence interval lower bound greater than 20%:



1. Website (60.0% of participant users, lower bound 23.1%)
2. Attended Class (47.6% of participant users, lower bound 24.3%)

Informal Study Group was by far the most frequently "Motivating" resource, among three resources rated as "Motivating" by more than half of participant users; all three resources involved people who might be considered peers (Informal Study Group, Formal Study Group, Teaching Assistants). However, of the three, only Informal Study Group had a confidence interval lower bound greater than 20%, as did Attended Lab:

1. Informal Study Group (76.5% of participant users, lower bound 54.0%)
2. Attended Lab (45.0% of participant users, lower bound 21.1%)

"Clarifying" had seven resources with endorsements from at least half of participant users. Two of these resources (Professor Office Hours and Attended Class) had a confidence interval lower bound greater than 50%, and four other resources had a confidence interval lower bound greater than 20%:

1. Attended Class (71.4% of participant users, lower bound 50.3%)
2. Teaching Assistants (63.6% of participant users, lower bound 29.5%)
3. Website (60.0% of participant users, lower bound 23.1%)
4. Informal Study Group (58.8% of participant users, lower bound 22.7%)
5. Textbook (55.6% of participant users, lower bound 30.2%)

No resource received "Appealing" endorsements from half of its participant users. Informal Study Group had the highest percentage of participant users who considered it "Appealing" at 41.2% and it had the highest lower bound at 15.1%.

### **4.2.3 Results of Software Evaluation and Feedback**

After at least five minutes spent with the Software, participants evaluated whether the utility dimensions could apply to different groups of organic chemistry learners. The results appear in Table 4.3.

Table 4.3: Evaluation of the Software from the utility study, with 21 participants.

User Scenario	Useful	Efficient	Effective	Accessible	Convenient	Motivating	Clarifying	Appealing
For a student in an organic chemistry course at a college or university	18 (85.7% ± 16.3%)	17 (81.0% ± 18.3%)	15 (71.4% ± 21.1%)	16 (76.2% ± 19.9%)	19 (90.5% ± 13.7%)	15 (71.4% ± 21.1%)	10 (47.6% ± 23.3%)	16 (76.2% ± 19.9%)
For someone revisiting organic chemistry while studying for the MCAT	18 (85.7% ± 16.3%)	13 (61.9% ± 22.7%)	14 (66.7% ± 22.0%)	12 (57.1% ± 23.1%)	13 (61.9% ± 22.7%)	14 (66.7% ± 22.0%)	8 (38.1% ± 22.7%)	14 (66.7% ± 22.0%)
For someone learning organic chemistry through an online course	19 (90.5% ± 13.7%)	16 (76.2% ± 19.9%)	14 (66.7% ± 22.0%)	16 (76.2% ± 19.9%)	16 (76.2% ± 19.9%)	12 (57.1% ± 23.1%)	9 (42.9% ± 23.1%)	13 (61.9% ± 22.7%)
For someone learning organic chemistry through self-study online	16 (76.2% ± 19.9%)	13 (61.9% ± 22.7%)	12 (57.1% ± 23.1%)	18 (85.7% ± 16.3%)	16 (76.2% ± 19.9%)	14 (66.7% ± 22.0%)	10 (47.6% ± 23.3%)	15 (71.4% ± 21.1%)

Key takeaways from the checkbox data, with 95% confidence interval in mind:

1. A majority of participants perceived the Software as "Useful" for some members of all given learner groups (student in an in-person academic course, learner reviewing for MCAT, student in an online course, learner self-studying online), and the confidence interval implies that a majority former organic chemistry students like the participants would most likely also perceive it as Useful for all given learner groups.
2. For a student in an organic chemistry course at a college or university, a majority of participants perceived the Software as Useful, Efficient, Effective, Accessible, Convenient, Motivating, and Appealing. The majority of former organic chemistry students would endorse each of these dimensions if they had briefly used the Software.
3. For a someone revisiting organic chemistry while studying for the MCAT, a majority of participants perceived the Software as Useful, Efficient, Effective, Accessible, Convenient, Motivating, and Appealing. It is likely that at least one in three people like the participants would have held the same opinion for each of these dimensions. It is likely that a majority would have perceived the Software as Useful for this learner group.
4. For someone learning organic chemistry through an online course, a majority of participants perceived the Software as Useful, Efficient, Effective, Accessible, Convenient, Motivating, and Appealing. It is likely that at least one in three people like the participants would have held the same opinion for each of these dimensions. It is also likely that people like the participants would have a majority endorse the Software as Useful, Efficient, Accessible, and Convenient.
5. For someone learning organic chemistry through self-study online, a majority of participants perceived the Software as Useful, Efficient, Effective, Accessible, Convenient, Motivating, and Appealing. It is likely that at least one in four people like the participants would have held the same opinion for each dimension. It is also likely that a majority of people like the participants would have perceived the Software as Useful, Accessible, Convenient, and Appealing.

Note that simply because the participants or people like the participants would perceive the Software in a particular way for different learner groups, it should not be expected that users of the Software in each learner group would necessarily have

the same trends in their opinions. Still, all of the participants were at some recent point in time in the category of the first learner group (student taking an in-person course at a university) and some may have also fit other categories. The participants generally believe that the Software could have usefulness to at least some learners in each of the categories, and the data suggests that people like the participants would also believe so under similar circumstances; together, this suggests that it is likely that learners of organic chemistry exist who would perceive usefulness in using the Software or a digital tool like it after an initial five-minute exposure.

Subsequent open-ended questions asked for elaboration on checkbox reasoning, as well as perceived advantages, perceived disadvantages, disliked aspects, and liked aspects of the Software.

Disadvantages and disliked aspects raised by participants:

1. Inability to see correct resonance structure before moving on via skip (5 participants) - could be frustrating especially if that is the end of the practice session (1 participant)
2. Lack of in-depth explanations (4 participants)
3. Perception of repetitive problems (similar solutions) and/or dislike seeing the same problem again (4 participants)
4. Potentially confusing without a tutorial or more direction (2 participants)
5. Actually drawing out the structure has learning benefits (2 participants) - including preparing for situations that require drawing the molecule and visualizing electron movements in your head (1 participant)
6. Not useful after user feels mastery over material or for advanced work (2 participants)
7. Not clarifying for someone self-teaching organic chemistry, due to lack of detail and assumption of prior knowledge (2 participants)
8. Lack of context (e.g. embedded in a lesson or with more of an introduction would be more useful to those less familiar with resonance) (2 participants)
9. Need prior basic understanding of resonance (2 participants)
10. Little direction given if a wrong answer submitted/Won't help learner understand their errors the same way instructor feedback does (2 participants)

11. Hint button might be useful after two or three attempts, which could make it more "clarifying" (2 participants)
12. Would prefer links to review relevant information (1 participant)
13. Didn't think it was very useful for learning (1 participant)
14. Problem difficulty varied but most seemed not very challenging (1 participant)
15. Problem variety could be improved (1 participant)
16. Concern about how accessible digital tools are generally as in most people won't use it if they have to do several steps to get to it (1 participant)
17. Disliked "Learn More" button and would prefer seeing more obvious explanation (1 participant)
18. Inefficient for someone reviewing for MCAT due to volume of material they need to review, as too time-intensive for those purposes (1 participant)
19. Concern about potential to give false sense of achievement (1 participant)
20. Didn't like sound effects (1 participant)
21. Problem order didn't seem like order of difficulty (1 participant)
22. Didn't really understand how to use it at first (1 participant)
23. Didn't understand where to click at first (1 participant)
24. The term "significant" was unfamiliar (1 participant)

Advantages and liked aspects raised by participants:

1. Good practice for pushing electrons/resonance (6 participants)
2. Enjoyable/fun (5 participants)
3. Liked immediate error-catching and feedback (5 participants)
4. Good for reinforcement or review/Good for someone who is somewhat familiar with resonance (4 participants)
5. Liked sound for getting a problem right (4 participants)

6. Visual (3 participants)
7. Liked detailed error messages explaining source of error and thought they would help with learning (3 participants)
8. Easy to use (3 participants)
9. Good for practicing many problems/finding new examples (3 participants)
10. Liked tracking streak (3 participants)
11. Liked gold stars for performance (2 participants)
12. Lack of detailed clarification may foster critical thinking and independent problem-solving / avoid giving away the answer (2 participants)
13. Self-paced (2 participants)
14. Liked sound effects (unspecified) (2 participants)
15. Liked visual assistance in seeing what is happening with the electrons (3 participants)
16. Liked option to turn on/off sound (1 participant)
17. Good for learning and practicing the basics (1 participant)
18. Liked lack of discouragement for getting a problem wrong (1 participant)
19. Liked motivating text when getting a problem right (1 participant)
20. Liked wording in message shown when skipping a problem: doesn't discourage returning to the problem (1 participant)
21. Interactivity helps with focus on task (1 participant)
22. Good for self-checking knowledge (1 participant)
23. Good for anxious learner to gain some comfort with the material before approaching others (1 participant)
24. More appealing than non-interactive digital tools e.g. videos, text, images (1 participant)
25. Would use on mobile during idle moments (1 participant)

26. Liked lack of introductory time/steps; can start practicing right away (1 participant)
27. Liked trial-and-error style (1 participant)
28. Could reduce frustration with wrong answers by not emphasizing wrongness and by helping people find the issues with their solution (1 participant)
29. Liked instant gratification (1 participant)
30. Faster feedback than homework which may take days to grade (1 participant)
31. Liked being able to play around with the molecule (1 participant)
32. Liked assurance that Software will prevent learner from spending lots of time doing something incorrectly as might happen during independent problem-solving on paper (1 participant)
33. Liked animation (1 participant)
34. Straightforward and appealing graphics (1 participant)
35. Arrows easy to use after some initial trial-and-error (1 participant)
36. Liked that user's solution is shown again after completing the problem (1 participant)

Of the 21 participants, 20 answered the following multiple-choice question:

Thinking back to when you studied organic chemistry, would you have preferred to use the tool over paper media (textbook problems, workbook problems, etc.)?

Participants gave the following answers:

1. Yes, always — 3 participants (15%)
2. Sometimes, when first learning about resonance — 5 participants (25%)
3. Sometimes, after getting very familiar with resonance — 5 participants (25%)
4. Sometimes, depending on the context — 6 participants (30%)
5. No, never — 1 participant (5%)

(Always 15%, Sometimes 80%, Never 5%,  $\pm$  10.5%, with "Yes, always" as 1, all "Sometimes..." as 0.5, "No, never" as 0 when calculating the sample mean)

Participants had a chance to explain their answer to the previous question. The "No, never" participant explained that the Software seemed inefficient and that they could have done the problems on paper more quickly. Some of "Sometimes..." participants noted that this tool was purely for practice rather than instruction, which influenced their answers. Others noted that while the Software is helpful for a very specific skill, or for practicing quickly without time-intensive drawing, or for getting immediate feedback, other types of resources may build other important skill sets such as drawing. Participants differed in their opinions about whether the Software was more suitable for learning (given basic level of problem-solving and immediate feedback) or review (given lack of detailed explanations). One participant compared the Software to their physical molecule model kit in its potential to help trigger an "aha!" visualization moment, in conjunction with other types of resources. Two "Yes, always" participants cited the instant feedback for self-correcting, while the third cited the paper-based inefficiency of having to refer back to original problems if not writing answers in textbook or workbook because of resell value.

### 4.3 Limitations

The small number of participants (21) and nonrandom sampling are two key limitations to the utility study. The number of participants introduces wide margins of error, which for some quantitative metrics limit whether the data can be generalized to a broader population. The small number of participants increases the risk of sampling artifacts and also extreme results due to random chance [68], and while the margin-of-error attempts to account for the latter, it cannot adequately account for the former. Though participants were not exclusively recruited from the same university as happens in many studies, the sampling of former organic chemistry students was still nonrandom: they were recruited either by the investigator or by friends and family of the investigator through social media, email, and other means of communication, and they opted to take the time to participate. That said, even statistical studies with a high number of participants carry risks that important variables are not adequately controlled for [15].

A caveat about participant bias, for both quantitative and qualitative results: in general, people fail to accurately evaluate innovative products because they cannot objectively forecast how they would actually use the product in their lives, or even if



they would use it [26]. The Software has similarities to existing products, so there is no reason to entirely disregard the forecast of participants in the study, but this type of study in general warrants some skepticism.

# Chapter 5

## Discussion and Conclusions

This chapter discusses and analyzes the results of a utility study that examined perceived usefulness of a variety of academic resources and of the Software. This chapter also discusses future work and highlights the challenges posed and questions raised by this thesis project.

### 5.1 Analysis of Utility Study Results

This section discusses the results of the utility study as described in Chapter 4.

#### 5.1.1 Discussion of Quantitative Results

Overall, the primary findings of the quantitative results are:

1. Participants use different resources for different purposes and reasons, with individual variation and preferences
2. Participants have mixed opinions about the usefulness of different resources across different utility dimensions in their own personal experiences studying organic chemistry, and this likely holds for the general population of recent former students of organic chemistry in the United States
3. A majority of participants perceived the Software as useful in multiple utility dimensions for some members of four groups of learners: students in an in-person academic course, people studying for the MCAT, students in an online course, and a self-studying online learner; recent former students of organic chemistry in the United States likely also would have a majority endorse the Software as "Useful" for all four groups under similar circumstances

4. A majority of participants indicated that they would have preferred to use the Software at least some of the time over paper media, and, under similar circumstances, recent former students of organic chemistry in the United States likely also would have a majority indicate this preference

Though recruitment and setup did not completely minimize participant bias, the utility study strongly suggests that a nontrivial number of recent former organic chemistry students would perceive the Software as comparably or even favorably useful to other resources in various circumstances, though discrepancies in exposure during the study and differing expectations about different media complicate direct comparisons [69]. Taking recent former organic chemistry students as a proxy for the intended audience of current students who are in the process of learning resonance, there is reason to suspect that the intended audience may also include a nontrivial number of people who would perceive the Software as useful to their learning.

Moreover, participant evaluations of the Software for "Motivating" and "Appealing" dimensions suggest that a digital practice tool like the Software might benefit learners in different ways than other types of resources—as more consistently motivating than a website resource, for example—though further inquiry is required. Given that questionnaire findings put teaching assistants and study sessions with peers as the most consistently motivating resources, a motivating digital resource may benefit students who are unwilling or unable to connect with fellow students, especially early on in their organic chemistry studies.

### 5.1.2 Discussion of Qualitative Results

Both quantitative and qualitative data support the conclusion that many people studying organic chemistry use a variety of academic resources and tools to fill different niches in their learning needs and personal circumstances. In this academic resource ecosystem, it seems plausible that a practice tool like the Software can fit into at least one niche.

At a bare minimum, mixing media for practice problems without relying on multiple choice may lessen the effects of cues on retrieval events, which leads to better learning outcomes (as discussed Section 1.2.2). Not only are multiple-choice cues eliminated with a tool like the Software, but if it can pull some learners away from doing all of their practice problems in the same notebook by offering an alternative with sufficient appeal, then those learners will build more decontextualized knowledge that they can draw upon more readily in the absence of external cues [15, 17,

21]. Simply switching the medium reduces the risk of context dependence, which can be so insidious that even a small format difference on exams can destroy student performance if their study habits unintentionally promoted overspecific knowledge tainted by superficial cues [17, 21]. One study participant noted that they liked to mix media when reviewing and summarizing material, which qualifies as an effective study strategy.

Yet it seems that at least for the participant sample, using digital media for learning organic chemistry usually means watching videos and seeking out explanations, rather than engaging in practice. Both a participant who frequently sought out supplementary digital resources and one who disavowed them agreed that supplementary digital resources may differ from course expectations and conventions, and that the discrepancy was a drawback. If the learner has not learned the material well or only relies on the external resource, that may be a drawback on an exam, but in the long term decontextualizing knowledge is often beneficial. However, the risk that practice material may not be at sufficient difficulty for preparing for an exam can exist within course materials, if there is a large discrepancy between homework and exams (as one participant noted in their experience with the homeworks assigned through Pearson's Mastering Chemistry vs. much more difficult exams). Still, the risk is magnified for external resources, and that does pose a challenge to developing digital practice tools outside of a specific course context.

Though the questionnaire did advise participants at one point that the problem content of the Software was not under study, several expressed opinions about the problems. They seemed repetitive, out of order of difficulty, too easy, too similar, or too basic, depending on the participant. For more proficient learners, the version of the Software in the study may not have had sufficient quantity of problems for the time participants spent using it. Additionally, manually curating content for a digital practice tool requires educational domain knowledge and extensive user testing to set the order; given a large set of problem, algorithmic reordering based on aggregate user data may be beneficial in a publicly available deployed tool.

As several participants pointed out, the Software without context seemed incomplete. Some participants advised adding a more detailed initial setup, links to other material, more detailed instruction about resonance, or more detailed explanations. Given the limited focus of this Software on chemical resonance, and given the difficulty of discovering a standalone tool whose purpose is short-lived for most learners, further development of this Software would depend heavily on the context for deployment, which should be within a larger learning environment, whether that is a suite

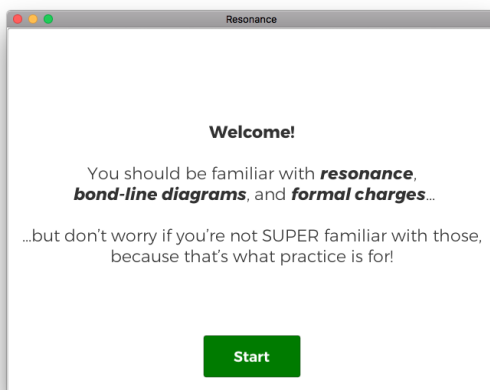


Figure 5.1: This welcome dialog is the extent of the Software's learning context, and as utility study participants pointed out, this context does not feel complete.

of organic chemistry practice tools or part of a site that includes non-practice content. As one participant emphasized, if it is not easy to access a digital tool, including easy to find, the tool will not get used. One participant noted that they could use it on a mobile device during idle moments; again, the context, including platform and access, determines whether or not the Software would be a quick and easy time-filler. The right design is the right design for the context and the task, and without context, much of future development would hinge on arbitrary decisions [55].

Context would also inform how to motivate users to find all of the resonance structures, rather than just the first one that they learn how to make from the initial structure; one way might involve making the user find all of the structures to solve the problem, and another might have the user unlock structures through a collection mechanic over long term use, and other approaches exist as well. Finding all resonance structures of a molecule and recognizing their relative significance is the eventual goal of resonance mastery [9]; moreover, the requirement of finding all of the structures would add another way to build difficulty and variation into the problem content.

The issue of multiple resonance structures impacted the qualitative feedback in other ways, however. Participants expected to be able to see a correct structure if they could not figure out the answer, as one participant also brought up in the usability study. One participant also expressed a desire for a hint system that would lead the user to an answer, rather than just skip or see the solution.

Organic chemistry students typically check their answers against an answer key, and not being able to do so in the Software frustrated some. Other participants mentioned an appreciation for trial-and-error approaches, with two noting that the

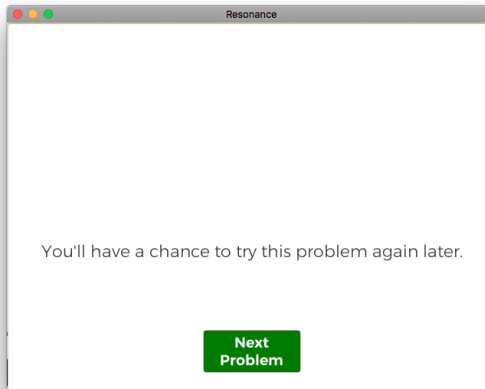


Figure 5.2: The skip dialog did not show a correct solution to the skipped problem, which some participants would have preferred.

lack of detailed clarification in the error messages meant that the answer was never given away and forced the user to think for themselves.

When the problem goal is to find one new resonance structure, as it was in the Software, there are many solutions, and deciding how to guide the user through a hint system becomes a difficult problem or an arbitrary decision, which may or may not deal with what the user actually got stuck on. Simply showing the solution should require showing all possible solutions, but the user might not take the time in that case to make sure that they understand how each is a valid solution. Again, handling multiple resonance structures depends on the type of learning environment: how much it saves user progress and expects users to return for multiple sessions. Moreover, the context may inform the best way to prevent abuse of the hint system, as some learners will rely too heavily on progressive hints leading to a correct answer [25].

On a related note, one participant disliked the "Learn More" button accompanying brief error messages and wanted a more obvious explanation immediately available. Keeping the error detail message one click away adds just a small amount of friction to discourage relying too heavily on it, and to encourage more independence with error analysis. This design does not prevent users from getting more guidance about their errors, but it does nudge them to better practice habits of self-checking [58].

Several participants liked the feedback and error messages that they received from the Software, which helped them identify errors and correct their thinking in the moment rather than wait for homework to be graded.

Some participants liked the sound for correctness. Some participants noted the

wording of text in parts of the Software, including the emphasis on trying again rather than getting discouraged when the solution is incorrect. Several participants liked tracking their streak and getting gold stars, as did users in the usability study. These reinforcements of success may encourage greater persistence as problems and material get more challenging [21]. One participant noted that the Software could benefit an anxious learner who avoided approaching others until they had some initial comfort with the material, as can other digital tools. Though people generally perform better when other people are involved, the reverse often happens for a novice trying to learn unfamiliar skills [4].

Multiple participants found visual aspects of digital tools helpful in general, and also specifically liked the visual aspect of the Software, with some noting that it helps to see the electrons move with the arrow. One participant compared the Software to how a physical model kit can help organic chemistry students better understand and visualize the molecule, though in a different way.

Some participants found the Software fun to use, which is not required of a practice tool, though it suggests that the Software may have some appeal which would encourage use.

Two participants identified the self-pacing aspect as an advantage. One participant noted that the Software would prevent someone from practicing incorrectly on their own without noticing it.

One participant noted that they had not heard the term "significance" used with chemical resonances. Textbooks and chemists appear to use major/minor and significant/insignificant terminology interchangeably, sometimes mixing them (e.g. [9] uses "major" and "insignificant"). Moreover, some resources do not stress significance or discuss it with specific terms [8]; some resources do not discuss it at all [1]. In general, subjects where learners may use different conventions and vocabulary depending on the course they are enrolled in is a major challenge for supplementary digital tools.

Some participants noted that the Software was helpful for a specific skill but incomplete without other types of resources for other skills, which is not surprising. The same can be said for the other resources that participants evaluated, and reinforces the significance of context. How the Software or any digital practice tool relates to the resources in its immediate learning environment and to those available to the learner will determine what niche it fills and how it can best support the learner's learning.

### 5.1.3 Generalizable Challenges and Opportunities

For supplementary digital practice tools aimed at benefiting students who are enrolled in an academic course on the subject, without being developed for a specific course, there are particular challenges that this study's results highlight:

1. Providing practice of sufficient difficulty to prepare the learner for exams, where exam difficulty is unknown to the tool
2. Providing mastery feedback without harming learner's assessment of how well they are doing compared to course expectations and exam difficulty level, both of which are unknown to the tool
3. Providing sufficient variety and quantity of practice content
4. Effective communication and coverage in domains where courses and resources may not agree on terminology and conventions

The last point attests to a common design goal: avoid designing material so that only the target audience can use it, especially around vocabulary [55]. Of course, with domain-specific knowledge there is a need to use precise vocabulary, but a supplementary learning tool can and should empower learners of different backgrounds to overcome discrepancies in vocabulary and conventions to understand its content.

There are also potential opportunities for further inquiry:

1. Can digital practice tools motivate students better than other types of digital resources?
2. Are digital practice tools a more accessible way to motivate certain types of students than peer study groups?
3. How can digital resource ecosystems better integrate practice tools, especially in organic chemistry?

People attribute social attributes to interactive technology and respond emotionally, which is a manifestation of anthropomorphism [4, 26]. The utility study data about certain social resources as consistently "Motivating" and the perception of the Software as "Motivating" suggest that in the latter case, anthropomorphism is at work, and that users may respond to congratulatory text, celebratory sounds, or other aspects of the Software as a kind of social interaction [4]. Unlike more passive digital



media, practice tools respond strongly to user actions, and thus, social anthropomorphism is more intense [26]. Therefore, investigating how digital practice tools might shape user motivation and integrate into learning environments may help students who struggle with motivation or receive less social feedback.

## 5.2 Future Work

Any future development of the Software or similar digital tools beyond an initial proof-of-concept stage will require collaborating with educators to curate content and establishing the deployment environment. A practice tool for an audience that primarily practices in order to meet coursework and testing objectives should be embedded in an environment that includes other resources, such as practice tools for other topics and reference material, in order to attract users and to redirect users with inadequate prior knowledge to other resources. Moreover, with online learning, learners often fail to distinguish between key topics and minor topics unless the material is well structured, well marked, and well made, all of which point to the importance of the context in which a practice tool is situated [70]. The context may even matter more than the features [71, 72].

That said, for future development, here are advanced features to consider, depending on the context:

1. **Hint System, Solution Reveal, and Multiple Structure Solutions** — As discussed with the results of the qualitative study data in 5.1.2, learners expect some kind of resolution other than simply skipping a problem when they give up on solving it independently. Depending on the context in which the Software exists, the hint or reveal system might lead towards a single resonance structure, or the full set of resonance structures that belong to the molecule in the problem. Whether or not the typical user returns for multiple sessions will influence the types of designs that may best support learning goals, such as a collection completion mechanic which rewards but does not require finding all structures, or demanding all structures for every problem.
2. **Intermixed Contrasting Cases** — One of the more advanced scaffolding tiers might intermix problems where resonance does not occur; learners would have to recognize this and opt to not offer a solution to the problem. Contrast promotes better abstract thinking: learners who must differentiate between material seem to build schemata that include the salient features rather than simply

superficial features [15, 17]. Intermixed contrasting cases prevent learners from approaching every problem the same way, even when such an approach may be inappropriate, and promote a more attentive assessment of what strategy to use. Outside of a classroom, most problems people encounter, whether in organic chemistry or not, require people to first diagnose the problem and only then choose a strategy to solve it.

3. **Interleaving** — As part of a suite of organic chemistry practice problem tools, rather than a single-topic tool, resonance problems could be interspersed with other types of organic chemistry problems, much like how Khan Academy's "World of Math" Mastery Challenges interleave different types of math problems [57]. Though conventional educational practices parcel out knowledge and practice in discrete, sequential chunks, a technique called blocking, research on learning calls for changing such practices. Interleaving, which intermixes problems that use related skills after some blocking for initial learning, proves much more effective across many disciplines, especially in the long term [17, 73]. Interleaving not only provides the benefit of increasing exposure to contrasting cases, as noted above, but also increases how often learners have to recall their knowledge [73]. Switching between knowledge areas and skills requires shifting the contents of working memory, which forces more recall events and thereby strengthens learning and memory [73].
4. **Saved State, Progress Visualizations, and Choose-Your-Level** — Allowing the user to save their progress, and for a new user to start at a more advanced level than beginner, are essential features for repeat use. The context in which the Software exists determines the technical nature of saving progress and may influence an appropriate approach to letting users start as non-beginners. Progress visualizations beyond the simple statistics already in the Software may interest users, but depending on how the Software integrates with other features of its learning environment, the visualizations may not be specific to the Software itself and instead visualize progress across multiple topics.

Smaller technical changes that could increase extendability and potential problem content:

1. **Support for radicals** — Many practice problems involve lone pairs of electrons, but sometimes a learner may encounter a radical, a nonbonding electron

on an atom that is not part of a pair. Some resources neglect radicals at the stage in which learners encounter resonance [9, 53]; others introduce them sparingly as an unusual [8]. The idea in the latter case is that if a learner understands lone pairs and other concepts, the learner should still be able to apply their resonance knowledge to work with a radical.

- 2. Support for triple-bond geometry** — In bond-line drawings, triple bonds are straight rather than zigzag as a nod to 3D linear characteristics of triple bond geometry [9]. For the occasional resonance problem in which a bond may change between double and triple across two resonance structures, triple-bond geometry support would be useful. Some care would have to be taken to avoid showing the drawing geometry change to beginners, as that might confuse a learner to see the shape of the molecule change (it would resemble a reaction somewhat; some resources do not change the geometry of the drawings for that reason e.g. see answer key for problem 2.67 in [9]).
- 3. Calculate all resonance structures** — For a problem requiring all solutions, or for a hint system, calculating all valid structures would be necessary, rather than simply verifying that the user's solution does not violate any rules. Different resources give human strategies for systematically identifying resonance structures; looking for patterns of adjacent features in the bond-line diagram [9, 53], or thinking more critically about available orbitals in three-atom bonded groupings [8]. Using any of the textbook-recommended strategies, it would be better to pre-process problem molecules and store these calculations, because the strategies must be run on each new structure found to ensure complete coverage of all resonance structures.
- 4. Support attaching functional groups** — Bond-line drawings sometimes replace a substructure with simply writing the functional group, such as replacing a nitrogen bonded to a hydrogen with "NH"; treating these functional groups as if they are atoms in the problem data would be one straightforward technical solution with minimal additional work.
- 5. More complex significance rules** — The extent to which different resources emphasize significance varies; some do not mention significance at all while others stress that it is part of the ultimate goal of learning about resonance, and still others mention that not all resonance structures contribute equally without getting into specifics [1, 8, 9, 53]. Adding some more edge-case rules, as well

as providing some instructional coverage for learners coming from backgrounds where significance is not stressed, may be a useful addition.

6. **More complex acceptable electron movement rules** — To support highly advanced resonance practice problems, very rare exceptions to the rules exist and could be supported [9].
7. **Support for more 2D substructures** — Though many molecules can be represented by the existing Software, supporting the parsing and layout functions of additional substructures would expand the variety of molecules that can be used for resonance problems or other types of problems.

A simple example of extending the Software to other early topics in organic chemistry would be a tool for practicing formal charges. Rather than drawing arrows, users would assign charges to atoms in the molecule displayed, based on the number of electrons each atom has. Displaying multiple molecules instead of one, a more complex extension might provide practice with reaction mechanisms, where curved arrows represent actual movement of electrons rather than imaginary movement. Both of these can be achieved with mostly superficial changes to user interface; the biggest difference is the rules for enabling and validating the user's actions and, for multiple molecules, small changes to problem data structures and layout.

### 5.3 Conclusion

The utility study showed the variety of ways that organic chemistry students use academic resources to support their learning goals. The study supported the idea that a digital practice tool like the Software can meet different needs for certain learners, such as reinforcing concepts visually, allowing learners to pace themselves, encouraging learners, and providing immediate feedback. Ultimately, designing for a specific deployment context determines the appropriate direction for future development and extension of the Software.

This thesis project succeeded in its goal to develop a proof-of-concept digital practice tool for organic chemistry that eschewed multiple choice and user drawings. The results of the study do not definitively show that a tool like the Software would add value for some learners, but strongly support that conclusion. Lastly, design challenges that tools like the Software face include the need for a known deployment context, curating content for diverse learner backgrounds, and managing appropriate

difficulty for diverse learner backgrounds and needs, especially if the deployment context is a supplementary external resource not part of a particular course. From a development standpoint, an additional consideration is that while parts of a tool like this are extendable and repurposable, the Software and tools like it cannot effectively cover many different subtopics, whereas multiple choice platforms can easily accept problems from different subtopics.

Multiple choice is not the only option for targeted organic chemistry practice on digital platforms. As this thesis project demonstrates, specialized digital practice tools for subtopics in organic chemistry offer new approaches to helping learners teach themselves through practice.

# Appendix A

## Formal Usability Study

While informal user testing guided some development, this formal study examined how prospective users and proxy users learned and used the interface and features. Content was not under study.

For more information about the version of the Software used by participants in this study, including screenshots, see Appendix C.

### A.1 Procedure

Five participants were recruited among adults (age 18 or older) who self-reported having taken first-semester organic chemistry or equivalent at an accredited four-year university within the last four years, who were in the Hanover, NH area at the time of the study. Each participant had an individual session, without overlap with other participants, over the course of three distinct days. All participants used the same version of the tool.

The user testing took no more than 30 minutes for each participant, with approximately half of that time spent interacting with the Software in some way on an iPad, and half of that time spent filling out a questionnaire on the same iPad.

The procedure was as follows:

1. Investigator asks participant to "think aloud" and use the Software for five minutes without guidance. Investigator may verbally prompt participant to share their thoughts if the participant falls silent. Investigator takes notes.
2. Investigator asks the participant to show how to do particular tasks with the Software:

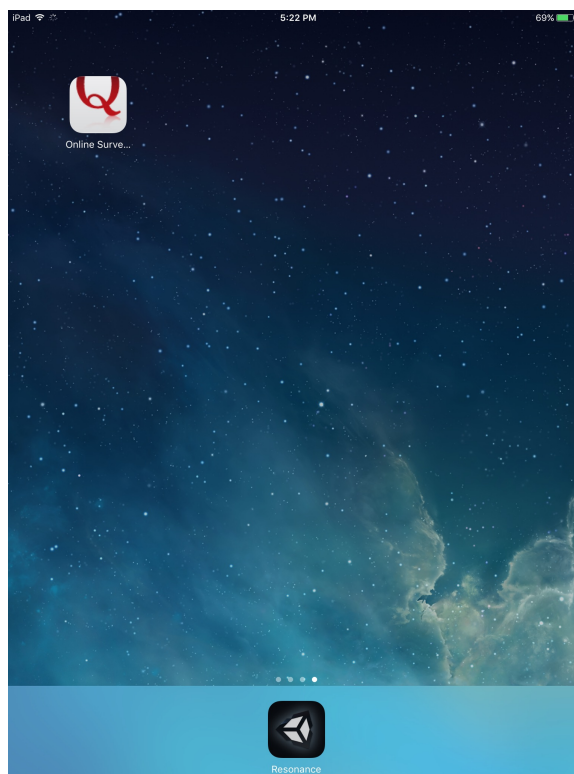


Figure A.1: The iPad setup for the usability study.

- (a) Can you show me how to push electrons from an atom to a bond?
- (b) Can you show me how to undo that arrow?
- (c) Can you show me how to check a solution?

This tests that the participant has a basic understanding of how to use the Software: recognizing atoms and bonds, drawing an arrow, undoing an arrow, and checking a solution.

3. Investigator interacts with Software and asks participant about the results:
  - (a) Investigator makes two arrows, and points to the one made first. Asks, "Can you tell me how to undo this arrow without undoing the other one?"
  - (b) Investigator makes an arrow whose source is a single bond and taps "Check Solution" button. Asks, "Can you tell me why this arrow is red?" and follows up with: "How did you know?" then "What could you do if you didn't understand why it was an error?"
  - (c) Investigator makes an arrow whose source and target are too far apart. Asks, "Can you tell me why my arrow disappeared?" and follows up with:

"How did you know?" then "What could you do if you didn't understand why it disappeared?"

- (d) Investigator makes many arrows. Asks, "Can you tell me how to start over without undoing each arrow individually?"
- (e) Investigator points to solution checks count. Asks, "Can you tell me what this green number represents?" and follows up with: "Can you tell me what the number beneath it represents?" (the user's solution checks per level average).
- (f) Investigator solves the problem to move to the next problem, and then solves the new problem in one go and taps "Check Solution" button. Asks, "Can you tell me what the star represents?"

This tests whether the participant understands more advanced features, including getting more information about an error. It also provides an opportunity to introduce the participant to the one-try star, if the participant did not previously solve a problem in one try.

- 4. Investigator closes the Software and opens a Qualtrics questionnaire on the iPad. Investigator leaves the room while participant fills out the questionnaire. Participant exits the room when done.

## A.2 Observations

Issues observed during initial unguided use:

- 1. A majority of participants first attempted to drag to draw arrows. Two participants did not immediately then try tapping.
- 2. A participant mistook "H" (hydrogen atoms) for "I" due to their orientation and got confused, briefly.
- 3. A participant tried to use the "Undo" button to deselect their most recent selection, rather than undo an arrow.
- 4. Two participants verbally noted that the lone pair electrons looked somewhat messy.
- 5. Two participants verbally indicated that they prefer "Submit" to "Check Solution" as the latter implied to them the idea of giving up and seeing someone



else's solution. One of these participants kept changing their arrows because they thought that the tool would automatically inform them that they had arrived at a correct solution.

6. A majority of participants verbalized that the error messages and explanations were helpful.
7. A participant became frustrated with a problem and verbally expressed a desire to be able to skip it.
8. Some participants tapped non-tappable parts of the screen (whitespace, formal charges, hydrogens bonded to carbon) at first; two were particularly persistent about tapping the hydrogens.
9. One participant verbally indicated that being able to draw an arrow from a single bond (a forbidden move) seemed odd.

Results of demonstration and explanation sessions:

1. All five participants demonstrated basic proficiency when asked to show how to push electrons from an atom to bond, undo an arrow, and check a solution.
2. To undo the arrow when asked, two participants tapped the arrow, one participant tapped the arrow's source, and two participants used the "Undo" button. All approaches were correct and successful.
3. All participants successfully explained how to undo an arrow without undoing a subsequent arrow, although one participant noted that it was a guess (this participant had only used the "Undo" button previously).
4. All participants successfully explained the cause of red arrows and disappearing arrows, referencing the error message and the "Learn More" detailed message as resources for someone who did not recognize the cause of the error.
5. All participants successfully chose the "Reset" button to clear a problem of multiple arrows at once.
6. All participants correctly explained the significance of the solution checks count; one participant incorrectly interpreted the average solution checks as referring to the average checks on the problem at hand across all users, rather than referring to their own average number of checks across all prior problems.

7. All participants who had solved a problem in one try explained the significance of the gold star as representing solving a problem without any incorrect solution checks; one participant, who had not solved a problem in one try on their own, agreed that it signified that the solver had done a "good job" on the problem, but did not explicitly connect it to solving the problem in one try.

### **A.3 Questionnaire Results**

A user identification number linked observed participant use of the tool with questionnaire responses.

First, participants scored aspects of the Software on how learnable, usable, and useful they seemed. For each question, the integer score response options ranged from "1 - Not at all" to "4 - Very" with no neutral response. The aggregate data for the first section of the questionnaire is reproduced below:

Table A.1: Usability study multiple choice responses.

Question	1 - Not at all	2	3	4 - Very	Mean Score
1. How difficult is learning to use the tool?	4 responses	1 response	—	—	1.20
2. How difficult is using the tool?	5 responses	—	—	—	1.00
3. How useful are the tool's instructional messages?	—	—	3 responses	2 responses	3.40
4. How useful are the tool's error messages?	—	—	3 responses	2 responses	3.40
5. How useful are the tool's error visuals?	—	—	4 responses	1 response	3.20
6. How intuitive is using the tool?	—	—	3 responses	2 responses	3.40
7. How understandable are the tool's bond-line diagrams, vs. textbook sources?	—	1 response	—	4 responses	3.60
8. How understandable are the tool's curved arrows, vs. textbook sources?	—	—	1 response	4 responses	3.80
9. How understandable was the information about solution checks on the current level?	—	—	1 response	4 responses	3.80
10. How useful was the information about solution checks on the current level?	—	1 response	—	4 responses	3.60
11. How understandable was the information about average solution checks?	—	—	2 responses	3 responses	3.60
12. How useful was the information about average solution checks?	—	—	3 responses	2 responses	3.40

Though these numerical results should not be considered statistically significant, they were an opportunity to flag discrepancies between the investigator's observations and participant experiences. However, despite variation in observed participant use of the tool and in the open-ended questions on the questionnaire, few scores stuck out from those of other participants. The two questions where four participants gave a score of 4 and one participant gave a score of 2 (Questions 7 and 9) had the same lone participant give the lower score. For Question 7, about how understandable the bond-line diagrams were, a likely source of the discrepancy is the participant's initial confusion about the hydrogen atoms, as the participant mistook some rotated "H" symbols as "I" symbols. As for Question 9, about the usefulness of the solutions check count, the same participant later in the survey observes that such a feature may be helpful for a student trying to track what problems they may need to work on (as did other participants); nothing in the questionnaire, nor observed use, provides any insight into why that participant gave that score.

After the page of multiple-choice answers, the questionnaire moved into open-ended questions:

1. Describe what it was like to learn how to use the tool.
2. Describe what it was like to use the tool.
3. Describe what the information about the number of solution checks on the current level meant to you as you used the tool, if anything.
4. Describe what the information about the average number of solution checks meant to you as you used the tool, if anything.
5. Describe what the congratulatory messages about solving problems in one try meant to you as you used the tool, if anything.
6. What parts of the tool's user interface were confusing, misleading, ambiguous, or otherwise hard to understand, if any?
7. Were there any aspects of the interface or feature set that surprised you? If so, what were they?
8. Were there any aspects of the interface or feature set that you expected would be there, but were not? If so, what were they?
9. The information about number of solution checks, average solution checks, and one-try streak information could be considered a kind of "score" or evaluation

for self-study. Did those features seem useful to you, and if not, why not? Are there any other types of information or evaluation that would be useful for self-study with the tool?

10. What problems did you have, if any, when learning or using the tool?
11. What problems do you think someone else might have when learning or using the tool? (Note: assume someone else has at least started covering resonance in a course setting)
12. What additional feedback or opinions about the tool's interface and feature set do you have, if any?

Issues raised by the open-ended questions:

1. One participant noted that it was frustrating to not be able to skip a difficult problem, and that they expected to be able to do so and to view someone else's solution (as one can do with a textbook's problem). Another participant also briefly described a wish to have the option to see a solution after trying an incorrect solution.
2. The participant who had confused the hydrogen "H" symbols with "I" noted that issue. Another participant noted that showing the hydrogens bonded to carbon made the molecule look more complicated and confusing (indeed, that is why organic chemistry favors implicit hydrogens in bond-line diagrams, though it requires greater fluency).
3. Three participants noted that they initially sought to drag rather than tap to draw arrows.
4. Two participants noted that they expected or preferred a "Submit" button to a "Check Solution" button.
5. Two participants noted that they were dissatisfied with how lone pair electrons looked.
6. The participant who had expected automatic confirmation of a correct solution noted that they had been confused by the "Check Solution" button.

Other insights from the open-ended responses:

1. A majority of participants noted that the current solution checks count could tell a user how well they were doing on a particular problem, as a way to self-check understanding or flag problems for follow-up. One participant noted that this was especially important for preparing for an exam, when the only way to check your answer was to self-check.
2. A majority of participants noted that the average solution checks could assess overall success or mastery.
3. A majority of participants wrote that the gold stars brought some satisfaction or confidence when they solved a problem in one try. Two participants noted that when they realized they could get more stars for streaks of one-try successes, they challenged themselves to catch their own errors.
4. One participant expressed a wish to be able to pick problems.
5. One participant expressed a wish to make the user identify all contributors rather than just one contributor to the resonance hybrid.
6. One participant noted that this tool seemed like it would save time as an alternative to having to draw the molecule and then the solution.
7. One participant briefly expressed a wish to be able to take notes, without much detail about the type of note.
8. One participant expressed a wish to save a difficult problem for later and come back to it.

# Appendix B

## Utility Study

To evaluate the Software, a utility study collected data from participants about perceived usefulness of the Software and other types of academic resources. Participants filled out an approximately 15-minute questionnaire in Qualtrics on a non-mobile device. The questionnaire had three parts, listed sequentially as they appeared:

1. Evaluation of prior experiences with different types of academic resources
2. A page with an embedded version of the Software; the button to continue in Qualtrics appeared after five minutes on the page
3. Evaluation of the Software

There were 21 participants, all of whom self-reported that they taken first-semester organic chemistry or equivalent at an accredited four-year college or university in the last four years. Participants were not family or friends of the investigator, and were recruited through family and friends' social networks as well as through posts in semi-anonymous online academic forums.

The anticipated user for the Software would be someone learning or reviewing chemical resonance, but the study used proxy users who had the prerequisite domain knowledge.

No performance data was collected other than timing data from using the Software and filling out the questionnaire.

See Appendix F for information about how margin of error (MOE) was calculated.

## B.1 Procedure

See Appendix D for details about the version of the Software used for the study and its deployment.

The procedure was as follows:

1. Prospective participants receive an anonymous Qualtrics link, with a request to use a non-mobile platform when participating (the web build of the Software used for the study was not compatible with mobile browsers).
2. Upon viewing an information sheet in Qualtrics, participants confirm that they are 18 years of age or older, consent to participate anonymously, and have taken first-semester organic chemistry or equivalent coursework at an accredited four-year university or college in the United States within the last four years (including postgraduate pre-health programs). Prior to continuing in the questionnaire, they also see a note instructing them to continue only if they were not using a mobile device and if they do not consider the researcher to be a personal friend or family member. They are also advised to have volume turned on if possible, though this is not required.
3. Participants answer questions about prior experiences and preferences for digital and non-digital academic resources; participants then evaluate resources that they used while studying organic chemistry.
4. Within Qualtrics, participants are presented with a web-based version of the Software and asked to use it for at least 5 minutes, in whatever manner they choose. They are informed that they will be able to proceed to the rest of the questionnaire after 5 minutes have passed. The page prior to the embedded Software contained the following description:

You are about to see and use a digital tool for practicing resonance problems in organic chemistry. The tool could be a standalone product, or part of a suite of digital tools for self-study of organic chemistry.

Please note that the particular practice problems you will see are not being studied at this time.

Please note that you may encounter bugs, as this software is in development.



5. Participants evaluate the Software's utility and give opinions about scenarios in which the Software may be useful or not useful.

## B.2 Questionnaire Results

### B.2.1 Prior Experiences

In this first section, participants were asked about their prior experiences and opinions of learning tools and resources.

They were asked the following questions:

1. Have you ever used any online learning tool to supplement your academic learning? (Khan Academy, Duolingo, Coursera, etc.)
2. If you answered yes, please describe which tools, why you used them, and what your experiences were like. If you answered no, are there reasons such tools do not appeal to you, are unnecessary, or are inappropriate for your needs? If so, please describe those reasons.
3. In general, what makes a digital learning tool useful? Are there particular types of learners, circumstances, or subjects/skills for which digital tools are especially useful or not very useful? If so, why?
4. Generally, do you prefer paper or digital media for academic tasks? Why? Are there particular circumstances or subjects that change your preference, and if so, what causes the change?

The first question had Yes/No radio button options:

- "Yes" — 15 responses (71.4%)
- "No" — 6 responses (28.6%)

MOE 21.1%.

The other questions asked for open-ended responses.

Of those who answered "Yes" and elaborated on their answer in the next question:

- 3 specifically mentioned using Khan Academy to supplement coursework in organic chemistry, although one discontinued use due to the unappealing length of videos (too long at "8 minutes or more"). Reasons cited: especially useful for explaining reaction mechanisms, useful for pre-exam review

- 3 specifically mentioned using YouTube videos to supplement coursework in organic chemistry. Reasons cited: useful for explaining mechanisms and other concepts, channel Leah4Sci especially useful
- 8 used Khan Academy for other kinds of coursework or didn't specify organic chemistry. Reasons cited: free, useful for understanding difficult concepts especially if confused in lecture, useful for basic explanations and review of medical sciences (e.g. studying for MCAT), useful for pre-exam review, mix of written and video content, mostly accurate and indicates any corrections to the videos, detailed, visual, useful for math, useful for biochemistry, useful for physiology, useful for physics
- 2 used YouTube for other kinds of coursework or didn't specify organic chemistry. Reasons cited: useful for science classes, useful explanations and procedure walkthroughs, useful for preparing for the next day's class lesson and for exams
- 7 used Duolingo for foreign language learning in some capacity (independent self-study or alongside academic work), with 1 discontinuing it soon after starting for being too basic for their needs.
- 1 used the mobile game app Majong Chem during idle time to brush up on chemistry basics while reviewing for the MCAT (participant noted later in another open-ended question that they forgot to mention it earlier)
- 1 used Wolfram Alpha to check math solutions and had a good experience with the comprehensiveness of the site
- 1 used MOOC courses in multiple advanced mathematic subjects
- 1 used Chegg for homework help
- 1 used Kaplan for MCAT preparation and was dissatisfied with their experience, due to expense and alleged error in one of the subjects, and thereafter used Khan Academy

Those who answered "No" explained why online tools failed to interest them:

- 2 felt that normal coursework and studying with course materials (textbooks, lecture, etc.) were sufficient

- 1 disliked mismatches between outside resources and course resources, concerned that using outside resources may backfire on an exam or in other coursework. (Another participant who answered "Yes" also identified this downside but still frequently used outside resources.)
- 1 had difficulties finding materials that covered content at an appropriate level
- 1 preferred books to online learning tools and felt that books delved deeper into details

Characteristics that make digital tools useful:

- 4 cited user friendliness, including specifically: easy to use, easy to navigate, aesthetics, organization, highly visual, provides feedback
- 4 cited video playback speeds, rewinding, rewatching, and pausing
- 1 cited images and graphs
- 1 cited easy to access
- 1 cited animations
- 1 cited generally cheaper than traditional media
- 1 cited ability to take notes directly on the page
- 1 cited updates/self-corrections, unlike paper textbooks with uncorrected errors in the current edition
- 2 cited quizzes or practice problems with rapid or immediate feedback
- 1 cited progress bars, when done in a satisfying way

The wording of the question may have biased some participants to hone in on the concept of sensory learning styles. Still, many referred to visual aspects as appealing or useful to them personally. Ways digital tools can benefit particular types of learners:

- 10 cited learning styles, especially visual and audio
- 4 cited self-pacing and extra exposure to learn material
- 1 cited as less stressful than classroom setting for students who may be intimidated by peers or professor when not yet comfortable with material

Circumstances where digital tools are useful:

- 3 cited reviewing material
- 1 cited rote memorization
- 3 cited as an alternative or supplement to inadequate or confusing lectures, textbooks, or other course materials
- 3 cited as an alternative explanation
- 1 cited demonstrations of problem-solving or procedures
- 1 cited exploration of nuance
- 1 cited high-level overview of unfamiliar subjects
- 1 cited filling in the gaps in foundational knowledge when course assumes prior knowledge or goes too quickly
- 1 cited as finding more examples and problems
- 1 cited for theoretical reading at leisure

Subjects for which digital tools are useful:

- 3 cited STEM subjects, including specifically: math, physiology, chemistry
- 2 cited subjects that use spatial reasoning or systems

Reasons digital tools not useful:

- 2 cited problem-based courses including organic chemistry (1 described as not useful, 1 described as maybe useful but not as good as traditional solo practice)
- 1 cited replacing or discouraging professor-student interactions, as students may avoid interacting with the professor as a result of access to online resources, a drawback that may harm both the student's learning and the professor's gauge of teaching efficacy
- 1 cited poor substitute for in-person learning experiences
- 1 cited a preference for physical book in hand
- 1 cited if already very familiar with material

- 1 cited as not for primary instruction

Reasons or cases where digital media preferred over paper media:

- 2 cited search functionality
- 5 cited note-taking capabilities, especially for speed or essay preparation
- 1 mixed media when reviewing and summarizing
- 1 cited subjects involving 3D space
- 4 cited efficiency: faster and easier to access content
- 2 cited mobility
- 1 preferred for self-study of languages
- 1 preferred for programming and computer science
- 2 preferred for writing papers
- 1 preferred for reviewing material
- 1 cited digitizing paper notes for mobility and easy access
- 1 cited cost (usually cheaper)
- 2 cited humanities subjects
- 1 cited biology
- 1 cited shareable
- 1 cited specific aspects of material can be better presented digitally
- 1 cited better for back (vs. carrying around big heavy textbooks)
- 1 cited accessing databases
- 1 preferred for reinforcing activities
- 1 preferred for feedback and identifying what to work on

Reasons or cases where paper media preferred over digital media:

1. 8 liked physicality of writing and/or thought it helped them remember better, with 2 directly referring to research about muscle memory benefits
2. 1 liked physical control
3. 1 disliked highlighting and typing on computer
4. 4 cited easier to draw, with 2 citing as necessary for subjects like organic chemistry
5. 1 cited easier on the eyes
6. 1 liked physicality of books
7. 4 cited sensory engagement (touch) even while reading seemed to promote better recall
8. 1 felt more personal
9. 1 avoids screen time
10. 3 preferred for math, with 1 only because of difficulty digitally writing equations
11. 4 preferred for chemistry (including organic), for ease of drawing and difficulty digitally writing equations and chemical structures
12. 1 preferred for hard sciences generally
13. 1 mixed media when reviewing and summarizing
14. 2 cited less distracting and more focused
15. 2 cited can see own thinking better, 1 specifically emphasizing helps to catch own errors
16. 2 cited for longer readings
17. 1 cited writing in textbook as routinely possible only with paper
18. 1 cited more mobile, not needing power outlet
19. 1 cited graphing and otherwise tangibly interacting with data
20. 1 cited printing out slides

21. 1 for assignments and tests

The majority of participants had mixed preferences for different types of subjects and tasks.

## B.2.2 Evaluating Other Resources

Participants were then shown a table with rows for academic resources and columns with descriptive attributes, and asked: "Did you use any of the following resources when studying organic chemistry? If so, please check any boxes that describe those resources in your experience."

The first column had the label "Did Not Use or N/A" to provide information to determine whether participants had used resources and were not checking boxes (suggesting that they did not feel like the label was a good fit), or hadn't used resources and so weren't checking boxes.

The other columns covered eight dimensions of utility that were presented as follows:

1. Useful
2. Efficient (speed, time cost-effectiveness)
3. Effective (improvement to future performance and learning)
4. Accessible (availability, monetary cost, ability to use)
5. Convenient (ease to use, ease to learn how to use, time/place/context convenience)
6. Motivating (encouragement, confidence, frustration)
7. Clarifying (explanations, information, assistance, confusion, comprehension)
8. Appealing (comfort, stress, intimidation, aesthetics, experience)

In tables reproduced in this appendix, the prompts in parentheses for each dimension are omitted, though they were present in the header above the checkboxes shown to participants.

The following table shows the aggregate data with number of checked boxes summed per column per resource:

Table B.1: Evaluation of academic resources in the utility study.

Resource	Did Not Use or N/A	Useful	Efficient	Effective	Accessible	Convenient	Motivating	Clarifying	Appealing
Textbook	3	11	2	8	9	7	2	10	3
Workbook	8	7	3	7	7	5	4	4	2
Teaching Assistants	10	10	2	4	6	4	6	7	4
Professor Office Hours	9	8	3	7	8	3	5	10	4
Formal Study Group	12	3	2	4	1	3	5	3	3
Informal Study Group	4	10	5	10	7	6	13	10	7
Tutor	17	2	0	2	0	0	0	1	0
Attended Class	0	19	6	13	14	10	5	15	4
Attended Lab	1	11	0	7	7	5	9	8	3
Online Lectures	13	7	5	4	5	5	1	4	1
Website	11	9	6	7	6	6	2	6	2
App	16	4	3	3	3	3	1	1	1
Other	14	1	0	1	1	0	0	0	1



Since not all 21 participants used all of the resources, the following table shows the number of participants who had used the resource, and gives the percent of those that had used it who checked the box for each utility metric.

Table B.2: (Repeat of Table 4.1) Evaluation of academic resources. Percentages indicate the percent of users who did check the checkbox for the utility dimension, out of all users who did not check "Did Not Use or N/A" for the resource.

Resource	Number of Users (Max. 21)	Useful	Efficient	Effective	Accessible	Convenient	Motivating	Clarifying	Appealing
Textbook	18	61.1% ± 24.9%	11.1% ± 16.1%	44.4% ± 25.4%	50.0% ± 25.6%	38.9% ± 24.9%	11.1% ± 16.1%	55.6% ± 25.4%	16.7% ± 19.1%
Workbook	13	53.8% ± 31.4%	23.1% ± 26.5%	53.8% ± 31.4%	53.8% ± 31.4%	38.5% ± 30.6%	30.8% ± 29.0%	30.8% ± 29.0%	15.4% ± 22.7%
Teaching Assistants	11	90.9% ± 20.3%	18.2% ± 27.2%	36.4% ± 33.9%	54.5% ± 35.1%	36.4% ± 33.9%	54.5% ± 35.1%	63.6% ± 33.9%	36.4% ± 33.9%
Professor Office Hours	12	66.7% ± 31.3%	25.0% ± 28.7%	58.3% ± 32.7%	66.7% ± 31.3%	25.0% ± 28.7%	41.7% ± 32.7%	83.3% ± 24.7%	33.3% ± 31.3%
Formal Study Group	9	33.3% ± 38.9%	22.2% ± 33.9%	44.4% ± 40.5%	11.1% ± 25.6%	33.3% ± 38.4%	55.6% ± 40.5%	33.3% ± 38.4%	33.3% ± 38.4%
Informal Study Group	17	58.8% ± 26.1%	29.4% ± 24.1%	58.8% ± 26.1%	41.2% ± 26.1%	35.3% ± 25.3%	76.5% ± 22.5%	58.8% ± 26.1%	41.2% ± 26.1%
Tutor	4	50.0% ± 80.1%	0.0% ± 0.0%	50.0% ± 80.1%	0.0% ± 0.0%	0.0% ± 0.0%	0.0% ± 0.0%	25.0% ± 69.4%	0.0% ± 0.0%
Attended Class	21	90.5% ± 13.7%	28.6% ± 21.2%	61.9% ± 22.7%	66.7% ± 22.0%	47.6% ± 23.3%	23.8% ± 19.9%	71.4% ± 21.1%	19.0% ± 18.3%
Attended Lab	20	55.0% ± 23.9%	0.0% ± 0.0%	35.0% ± 22.9%	35.0% ± 22.9%	25.0% ± 20.8%	45.0% ± 23.9%	40.0% ± 23.5%	15.0% ± 17.1%
Online Lectures	8	87.5% ± 26.6%	62.5% ± 43.4%	50.0% ± 44.7%	62.5% ± 43.3%	62.5% ± 43.4%	12.5% ± 29.6%	50.0% ± 44.7%	12.5% ± 29.6%
Website	10	90.0% ± 22.6%	60.0% ± 36.9%	70.0% ± 34.6%	60.0% ± 36.9%	60.0% ± 36.9%	20.0% ± 30.2%	60.0% ± 36.9%	20.0% ± 30.2%
App	5	80.0% ± 51.4%	60.0% ± 63.0%	60.0% ± 63.0%	60.0% ± 63.0%	60.0% ± 63.0%	20.0% ± 51.4%	20.0% ± 51.4%	20.0% ± 51.4%
Other	7	14.3% ± 35.0%	0.0% ± 0.0%	14.3% ± 35.0%	14.3% ± 35.0%	0.0% ± 0.0%	0.0% ± 0.0%	0.0% ± 0.0%	14.3% ± 35.0%

Participants were then asked an open-ended question about their use of websites, apps, or "other" resources if they had used them for organic chemistry. Some of the responses:

- 2 watched videos on Khan Academy for explanations when struggling with material
- 3 watched videos on YouTube for explanations when struggling with material or reviewing material
- 3 regularly searched the web as needed in quick, targeted searches for clarification and details about molecules and reactions
- 1 used Master Organic Chemistry (website with text and video explanations)
- 1 had coursework assigned through Pearson's Mastering Chemistry platform, and wished that the homework problems had been as difficult as the exam problems, noting that the discrepancy could create a false sense of mastery
- 1 used various university-affiliated websites (notably UCLA) for explanations and clarifications as needed
- 1 used their own university's website for research in an honors course
- 1 used website resources when needing clarification or explanation for a specific problem, but approached the professor for help with understanding an entire topic
- 1 used a website to practice naming compounds, with many more compounds to practice on than the textbook
- 1 used an app to memorize the amino acids
- 1 used an online organic chemistry course purchased through The Great Courses
- 1 used PubMed for research articles in an honors course
- 1 used ChemDraw to digitally draw out solutions to assignments
- 1 considered David Klein's *Organic Chemistry as a Second Language* in the "other" category (rather than workbook) and noted using it

### **B.2.3 Software Evaluation and Feedback**

After using the Software for at least five minutes, participants were asked to rate the tool using the same utility metrics while considering different types of users. The text above the checkbox table was as follows:

You have just used the digital tool. Imagine that it might be hosted on a website for desktop use or within an app for mobile, as a free self-study resource. Please check any boxes that describe how you perceive this resource for the following types of people learning resonance:

The aggregate data of participant responses is reproduced in the following table:

Table B.3: (Repeat of Table 4.3) Evaluation of the Software from the utility study, with 21 participants.

User Scenario	Useful	Efficient	Effective	Accessible	Convenient	Motivating	Clarifying	Appealing
For a student in an organic chemistry course at a college or university	18 (85.7% ± 16.3%)	17 (81.0% ± 18.3%)	15 (71.4% ± 21.1%)	16 (76.2% ± 19.9%)	19 (90.5% ± 13.7%)	15 (71.4% ± 21.1%)	10 (47.6% ± 23.3%)	16 (76.2% ± 19.9%)
For someone revisiting organic chemistry while studying for the MCAT	18 (85.7% ± 16.3%)	13 (61.9% ± 22.7%)	14 (66.7% ± 22.0%)	12 (57.1% ± 23.1%)	13 (61.9% ± 22.7%)	14 (66.7% ± 22.0%)	8 (38.1% ± 22.7%)	14 (66.7% ± 22.0%)
For someone learning organic chemistry through an online course	19 (90.5% ± 13.7%)	16 (76.2% ± 19.9%)	14 (66.7% ± 22.0%)	16 (76.2% ± 19.9%)	16 (76.2% ± 19.9%)	12 (57.1% ± 23.1%)	9 (42.9% ± 23.1%)	13 (61.9% ± 22.7%)
For someone learning organic chemistry through self-study online	16 (76.2% ± 19.9%)	13 (61.9% ± 22.7%)	12 (57.1% ± 23.1%)	18 (85.7% ± 16.3%)	16 (76.2% ± 19.9%)	14 (66.7% ± 22.0%)	10 (47.6% ± 23.3%)	15 (71.4% ± 21.1%)

Subsequent open-ended questions asked for elaboration on checkbox reasoning, as well as perceived advantages, perceived disadvantages, disliked aspects, and liked aspects of the Software.

Disadvantages and disliked aspects raised by participants:

1. Inability to see correct resonance structure before moving on via skip (5 participants) - could be frustrating especially if that is the end of your session (1 participant)
2. Lack of in-depth explanations (4 participants)
3. Perception of repetitive problems (similar solutions) and/or dislike seeing the same problem again (4 participants)
4. Potentially confusing without a tutorial or more direction (2 participants)
5. Actually drawing out the structure has learning benefits (2 participants) - including preparing for situations that require drawing the molecule and visualizing electron movements in your head (1 participant)
6. Not useful after user feels mastery over material or for advanced work (2 participants)
7. Not clarifying for someone self-teaching organic chemistry, due to lack of detail and assumption of prior knowledge (2 participants)
8. Lack of context (e.g. embedded in a lesson or with more of an introduction would be more useful to those less familiar with resonance) (2 participants)
9. Need prior basic understanding of resonance (2 participants)
10. Little direction given if a wrong answer submitted/Won't help learner understand their errors the same way instructor feedback does (2 participants)
11. Hint button might be useful after two or three attempts, which could make it more "clarifying" (2 participants)
12. Would prefer links to review relevant information (1 participant)
13. Didn't think it was very useful for learning (1 participant)
14. Problem difficulty varied but most seemed not very challenging (1 participant)

15. Problem variety could be improved (1 participant)
16. Concern about how accessible digital tools are generally as in most people won't use it if they have to do several steps to get to it (1 participant)
17. Disliked "Learn More" button and would prefer seeing more obvious explanation (1 participant)
18. Inefficient for someone reviewing for MCAT due to volume of material they need to review, as too time-intensive for those purposes (1 participant)
19. Concern about potential to give false sense of achievement (1 participant)
20. Didn't like sound effects (1 participant)
21. Problem order didn't seem like order of difficulty (1 participant)
22. Didn't really understand how to use it at first (1 participant)
23. Didn't understand where to click at first (1 participant)
24. The term "significant" was unfamiliar (1 participant)

Advantages and liked aspects raised by participants:

1. Good practice for pushing electrons/resonance (6 participants)
2. Enjoyable/fun (5 participants)
3. Liked immediate error-catching and feedback (5 participants)
4. Good for reinforcement or review/Good for someone who is somewhat familiar with resonance (4 participants)
5. Liked sound for getting a problem right (4 participants)
6. Visual (3 participants)
7. Liked detailed error messages explaining source of error and thought they would help with learning (3 participants)
8. Easy to use (3 participants)
9. Good for practicing many problems/finding new examples (3 participants)

10. Liked tracking streak (3 participants)
11. Liked gold stars for performance (2 participants)
12. Lack of detailed clarification may foster critical thinking and independent problem-solving / avoid giving away the answer (2 participants)
13. Self-paced (2 participants)
14. Liked sound effects (unspecified) (2 participants)
15. Liked visual assistance in seeing what is happening with the electrons (3 participants)
16. Liked option to turn on/off sound (1 participant)
17. Good for learning and practicing the basics (1 participant)
18. Liked lack of discouragement for getting a problem wrong (1 participant)
19. Liked motivating text when getting a problem right (1 participant)
20. Liked wording in message shown when skipping a problem: doesn't discourage returning to the problem (1 participant)
21. Interactivity helps with focus on task (1 participant)
22. Good for self-checking knowledge (1 participant)
23. Good for anxious learner to gain some comfort with the material before approaching others (1 participant)
24. More appealing than non-interactive digital tools e.g. videos, text, images (1 participant)
25. Would use on mobile during idle moments (1 participant)
26. Liked lack of introductory time/steps; can start practicing right away (1 participant)
27. Liked trial-and-error style (1 participant)
28. Could reduce frustration with wrong answers by not emphasizing wrongness and by helping people find the issues with their solution (1 participant)



29. Liked instant gratification (1 participant)
30. Faster feedback than homework which may take days to grade (1 participant)
31. Liked being able to play around with the molecule (1 participant)
32. Liked assurance that Software will prevent learner from spending lots of time doing something incorrectly as might happen during independent problem-solving on paper (1 participant)
33. Liked animation (1 participant)
34. Straightforward and appealing graphics (1 participant)
35. Arrows easy to use after some initial trial-and-error (1 participant)
36. Liked that user's solution is shown again after completing the problem (1 participant)

Problem difficulty was not under study, as briefly mentioned to the participants prior to using the Software, but it was not expected that they would remember this.

Of the 21 participants, 20 answered the following multiple-choice question: "Thinking back to when you studied organic chemistry, would you have preferred to use the tool over paper media (textbook problems, workbook problems, etc.)?" with the following answers:

1. Yes, always — 3 participants (15%)
2. Sometimes, when first learning about resonance — 5 participants (25%)
3. Sometimes, after getting very familiar with resonance — 5 participants (25%)
4. Sometimes, depending on the context — 6 participants (30%)
5. No, never — 1 participant (5%)

(Always 15%, Sometimes 80%, Never 5%,  $\pm$  10.5%, with "Yes, always" as 1, all "Sometimes..." as 0.5, "No, never" as 0 when calculating the sample mean)

Participants had a chance to explain their answer to the previous question. The "No, never" participant explained that the Software seemed inefficient and that they could have done the problems on paper more quickly. Some of "Sometimes..." participants noted that this tool was purely for practice rather than instruction, which

influenced their answers. Others noted that while helpful for a very specific skill, or for practicing quickly without time-intensive drawing, or for getting immediate feedback, other types of resources may build other important skill sets such as drawing. Participants differed in their opinions about whether the Software was more suitable for learning (given basic level of problem-solving and immediate feedback) or review (given lack of detailed explanations). One participant compared the Software to their physical molecule model kit in its potential to help trigger an "aha!" visualization moment, in conjunction with other types of resources. Two "Yes, always" participants cited the instant feedback for self-correcting, while another cited the inefficiency of having to refer back to original problems in paper media (e.g. flip to back of textbook for answer).

# Appendix C

## Software Version for Formal Usability Study

### C.1 Description of Software Version

The prototype of the Software used for the formal usability study had the following key differences from the version described in Chapter 3:

1. Different non-molecule user interface components
2. No audio
3. Different graphics for molecule user interface components
4. Implicit hydrogen and formal charges not interactive
5. Text differences

The version of the Software for the formal usability study can be identified by its git commit: `4ec82b713862eb237bcba8c70532bb3292c57baa`.

### C.2 Deployment

In Unity, the Software was built for iOS and exported to Xcode, then installed on an iPad device.

### C.3 Screenshots

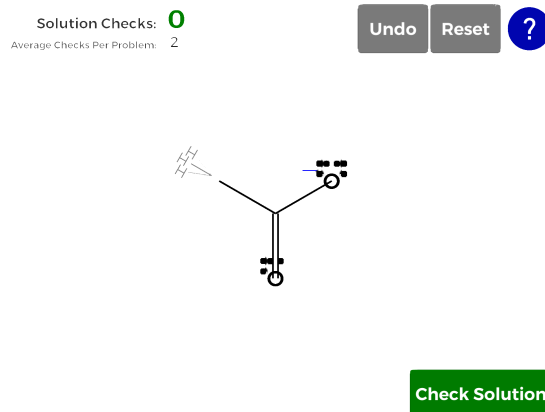


Figure C.1: Example of practice mode in the usability study prototype.

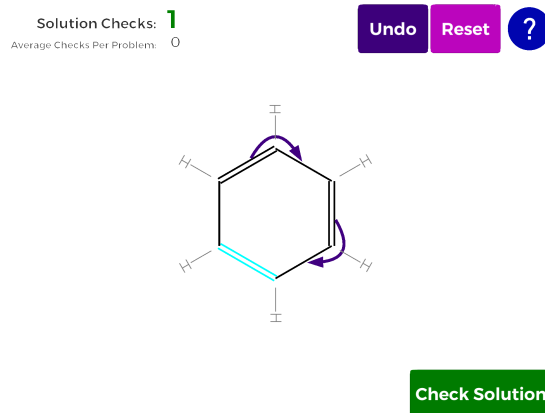


Figure C.2: Example of practice mode in the usability study prototype, with a bond selected.

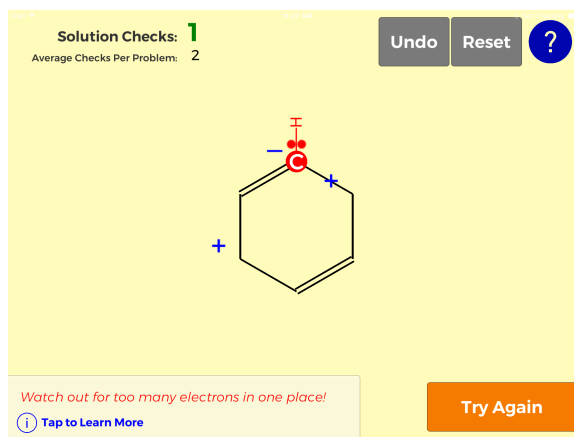


Figure C.3: Example of solution review mode in the usability study prototype.

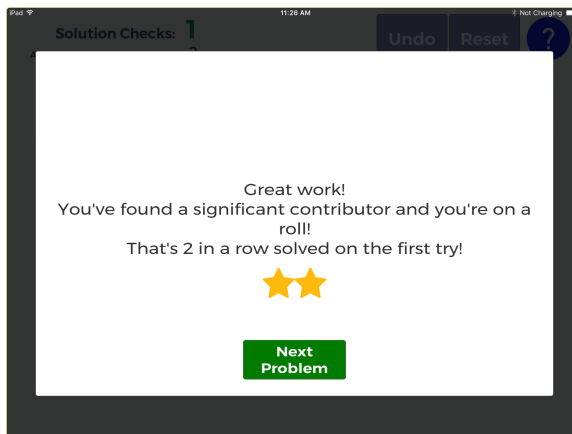


Figure C.4: Prototype's success dialog, after solving two problems in a row in just one attempt.

Welcome! You should be familiar with resonance, bond-line diagrams, and formal charges before using this tool...but don't worry if you're not SUPER familiar with those, because that's what this is for!

Your goal with each problem is to find a new significant contributor to the resonance hybrid.

Start

Figure C.5: Prototype's welcome message.

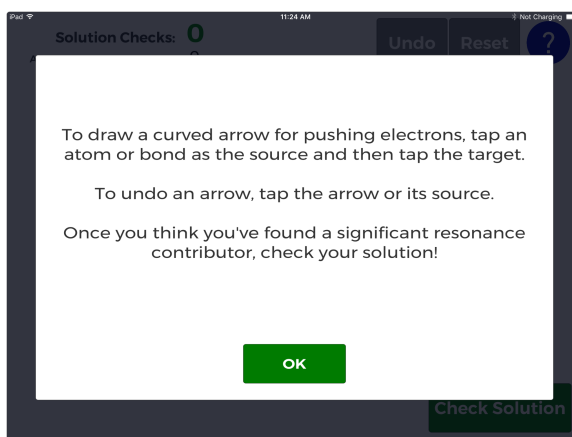


Figure C.6: Prototype's help dialog.

# Appendix D

## Software Version for Utility Study

### D.1 Description of Software Version

The version is the same as that described in Chapter 3. It can be identified by its git commit: 2ea52652f7c8a9a448ae4659b16694beb5081566.

### D.2 Deployment

The Software was hosted via Github Pages as a fixed-size Unity WebGL build. The web build of the Software used for the study was not compatible with mobile browsers. The build was embedded in a Qualtrics questionnaire, so the size was constrained to 480 x 320 for fit.

### D.3 Screenshots

Screenshots throughout this thesis use the OSX build rather than the WebGL build, which may render graphics differently. Here are two screenshots of the fixed-size WebGL build that was embedded in Qualtrics:

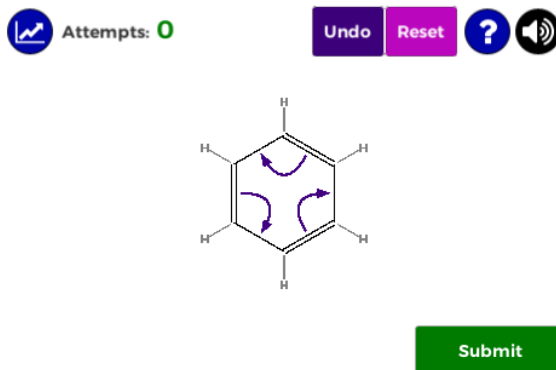


Figure D.1: WebGL build example of practice mode.

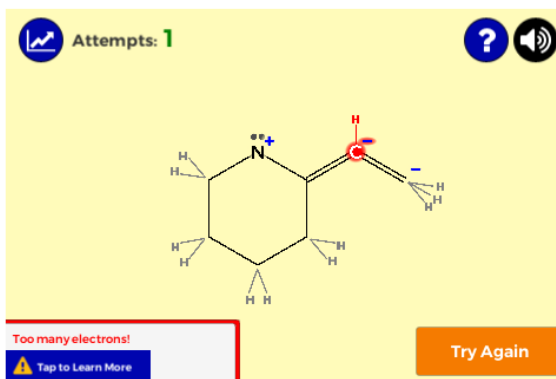


Figure D.2: WebGL build example of solution review mode.

# Appendix E

## External Assets and Resources

Versions of the Software that embed type files (such as mobile builds) do so under Open Font License 1.1 for the following type families:

1. Montserrat  
(<https://github.com/JuliettaUla/Montserrat>)
2. Font Awesome v4.7.0 by Dave Gandy  
(<http://fontawesome.io/>)

Versions of the Software that use sound do so under license:

1. Sounds from Octave library used under license  
(<https://github.com/scopegate/octave>)
2. Sound from shinephoenixstormcrow on Freesound used under Creative Commons Attribution license  
(<https://freesound.org/people/shinephoenixstormcrow/sounds/337049/>)
3. Public domain sounds from Freesound  
(<https://freesound.org/people/benagain/sounds/321083/>,  
<https://freesound.org/people/wagna/sounds/325805/>,  
<https://freesound.org/people/plasterbrain/sounds/397354/>)

All other assets are either created by the Software's creator or are provided defaults from Unity.

Functions and algorithms adapted from Unity Forum or Stack Exchange are credited directly in the code where applicable.



# Appendix F

## Margin of Error Calculations

The margin of error (MOE) was calculated using two-tailed t-values to account for small sample sizes with conservative margins for a 95% confidence interval.

For a sample size of  $n$ :

$$MOE = t_{n-1} * \frac{s}{\sqrt{n}}$$

where  $s$  is the standard deviation of the sample:

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - x_{mean})^2}{n - 1}}$$

The t-values were obtained through MedCalc Software's two-tailed t-distribution lookup table, publicly available on their website [74].

# Bibliography

- [1] *Khan Academy*. 2017. URL: <https://www.khanacademy.org/>.
- [2] *Duolingo*. 2017. URL: <https://www.duolingo.com/>.
- [3] *Codecademy*. 2017. URL: <https://www.codecademy.com/>.
- [4] B. J. Fogg. *Persuasive Technology: Using Computers to Change What We Think and Do*. The Morgan Kaufmann Series in Interactive Technologies. Boston: Morgan Kaufmann Publishers, 2003.
- [5] Kamran Sedig and Paul Parsons. *Design of Visualizations for Human-Information Interaction: a Pattern-Based Framework*. Synthesis Lectures on Visualization. Morgan & Claypool Publishers, 2016.
- [6] Geoff F. Kaufman and Mary Flanagan. “High-Low Split: Divergent Cognitive Construal Levels Triggered by Digital and Non-digital Platforms”. In: CHI ’16. San Jose, CA: ACM Press, 2016, pp. 2773–2777.
- [7] Bethany Halford. “Is there a crisis in organic chemistry education?” In: *Chemical & Engineering News* 94.13 (Mar. 28, 2016), pp. 24–25.
- [8] John McMurry. *Organic Chemistry*. Hybrid Edition, Eighth Edition. Belmont, CA: Brooks/Cole, 2012.
- [9] David Klein. *Organic Chemistry as a Second Language: First Semester Topics*. 3rd ed. Hoboken, NJ: Wiley, 2012.
- [10] *OWLv2*. Cengage, 2017. URL: <http://www.cengage.com/owlv2/>.
- [11] *Connect for Carey, Organic Chemistry, 8e*. The McGraw-Hill Companies, 2012. URL: <http://connect.customer.mheducation.com/products/connect-for-carey-organic-chemistry-8e/>.
- [12] *Connect for Carey, Organic Chemistry, 9e*. The McGraw-Hill Companies, 2014. URL: <http://connect.customer.mheducation.com/products/connect-for-carey-organic-chemistry-9e/>.

- [13] *Connect for Smith, Organic Chemistry, 4e*. The McGraw-Hill Companies, 2014. URL: <http://connect.customer.mheducation.com/products/connect-for-smith-organic-chemistry-4e/>.
- [14] *Pearson Mastering Chemistry*. Pearson Education, Inc., 2017. URL: <https://www.pearsonmylabandmastering.com/northamerica/masteringchemistry/>.
- [15] E. J. Mastascusa, William J. Snyder, and Brian S. Hoyt. *Effective Instruction for STEM Disciplines: From Learning Theory to College Teaching*. The Jossey-Bass Higher and Adult Education Series. San Francisco, CA: Jossey-Bass, 2011.
- [16] Henry L. Roediger III. "Using Testing to Improve Learning and Memory". In: *Psychology and the Real World: Essays Illustrating Fundamental Contributions to Society*. Ed. by Kathleen McDermott and Morton Ann Gernsbacher. New York, NY: Worth Publishers, 2011, pp. 65–74.
- [17] Elizabeth Ligon Bjork and Robert A. Bjork. "Making Things Hard on Yourself, But in a Good Way: Creating Desirable Difficulties to Enhance Learning". In: *Psychology and the Real World: Essays Illustrating Fundamental Contributions to Society*. Ed. by Morton Ann Gernsbacher. New York, NY: Worth Publishers, 2011, pp. 56–64.
- [18] *MasteringChemistry for Organic Chemistry*. Pearson Higher Education, 2013-11-19. URL: <https://www.youtube.com/watch?v=rElXQiXrdtw>.
- [19] *Drawing Multistep Mechanisms in Mastering Chemistry*. URL: <https://www.youtube.com/watch?v=yBh4Qg0as74>.
- [20] Cecilia Henry Arnaud. "The Internet Has Shifted the Way Students Learn". In: *Chemical & Engineering News* 93.32 (Aug. 16, 2015). URL: <http://internet.cenmag.org/the-internet-has-shifted-the-way-students-learn/>.
- [21] Susan A Ambrose et al. *How Learning Works: Seven Research-Based Principles for Smart Teaching*. New York, NY: John Wiley & Sons, 2010.
- [22] Don Norman. *The Design of Everyday Things*. Revised and Expanded Edition. New York: Basic Books, 2013.
- [23] Arthur S. Reber. *Implicit Learning and Tacit Knowledge: An Essay on the Cognitive Unconscious*. Oxford Psychology Series no. 19. New York: Oxford University Press, 1993.
- [24] Neil Garg. "Chemistry Education for the Youtube Generation". In: *Chemical & Engineering News* 93.32 (Aug. 16, 2015). URL: <http://internet.cenmag.org/chemistry-education-for-the-youtube-generation/>.

- [25] V. J. Shute. “Focus on Formative Feedback”. In: *Review of Educational Research* 78.1 (Mar. 1, 2008), pp. 153–189.
- [26] Donald A. Norman. *Emotional Design: Why We Love (or Hate) Everyday Things*. New York: Basic Books, 2004.
- [27] *What is a Chunk?* University of California, San Diego: Learning How to Learn: Powerful mental tools to help you master tough subjects. Coursera, 2017.
- [28] Daniel Kahneman. *Attention and Effort*. Prentice-Hall Series in Experimental Psychology. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1973.
- [29] Neil W. Mulligan and Alan S. Brown. “Attention and implicit memory”. In: *Attention and Implicit Learning*. Ed. by Luis Jiménez. Advances in Consciousness Research v. 48. Philadelphia, PA: John Benjamins Publishing Company, 2003, pp. 297–334.
- [30] Susanne Vogel and Lars Schwabe. “Learning and memory under stress: implications for the classroom”. In: *npj Science of Learning* 1.1 (Dec. 2016).
- [31] Elliott M. Avedon and Brian Sutton-Smith. “Games in Education”. In: *The Study of Games*. New York: J. Wiley, 1971.
- [32] *Case Study Part I: Math Blaster*. MITx: 11.127x Design and Development of Games for Learning. edX, 2016.
- [33] *Lure of the Labyrinth: Design Process (Part 4)*. MITx: 11.127x Design and Development of Games for Learning. edX, 2016.
- [34] Peter G. Stromberg. *Caught in Play: How Entertainment Works on You*. Stanford, CA: Stanford University Press, 2009.
- [35] Kurt Squire. *Video Games and Learning: Teaching and Participatory Culture in the Digital Age*. In collab. with Henry Jenkins. Technology, Education–Connections: the TEC Series. New York: Teachers College Press, 2011.
- [36] Nicola Whitton. *Digital Games and Learning: Research and Theory*. Digital Games and Learning. New York: Routledge, 2014.
- [37] James Paul Gee. “Learning and Games”. In: *The Ecology of Games: Connecting Youth, Games, and Learning*. Ed. by Katie Salen. The John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning. Cambridge, MA: The MIT Press, 2008, pp. 21–40.

- [38] May-Li Khoe. *Chocolate-covered broccoli*. Khan Academy Long-Term Research. Aug. 17, 2016. URL: <http://klr.tumblr.com/post/149091244433/chocolate-covered-broccoli>.
- [39] Andy Matuschak. *Game designers vs. education researchers on unguided instruction*. Khan Academy Long-Term Research. Apr. 14, 2015. URL: <http://klr.tumblr.com/post/159583871898/game-designers-vs-education-researchers-on>.
- [40] Jeanne V. Russell. "Using Games To Teach Chemistry: An Annotated Bibliography". In: *Journal of Chemical Education* 76.4 (Apr. 1999), p. 481.
- [41] *Collisions*. 2017. URL: <https://www.playmadagames.com/>.
- [42] *ChemCrafter*. 2014. URL: <http://assets.chemheritage.org/ChemCrafter/>.
- [43] *Mahjong Chem*. Stetson University. 2015. URL: <https://play.google.com/store/apps/details?id=com.mahjongchem>.
- [44] *React! - The Organic Chemistry Board Game*. Kickstarter, 2017. URL: <https://www.kickstarter.com/projects/345110877/react-the-organic-chemistry-board-game>.
- [45] *React!* 2017. URL: <http://www.reactgame.com/>.
- [46] Steven C. Farmer and Molly K. Schuman. "A Simple Card Game To Teach Synthesis in Organic Chemistry Courses". In: *Journal of Chemical Education* 93.4 (Apr. 12, 2016), pp. 695–698.
- [47] Jeffrey M. Carney. "Retrosynthetic Rummy: A Synthetic Organic Chemistry Card Game". In: *Journal of Chemical Education* 92.2 (Feb. 10, 2015), pp. 328–331.
- [48] José Nunes da Silva Júnior et al. "Stereogame: An Interactive Computer Game That Engages Students in Reviewing Stereochemistry Concepts". In: *Journal of Chemical Education* 94.2 (Feb. 14, 2017), pp. 248–250.
- [49] Julia Winter, Michael Wentzel, and Sonia Ahluwalia. "Chairs!: A Mobile Game for Organic Chemistry Students To Learn the Ring Flip of Cyclohexane". In: *Journal of Chemical Education* 93.9 (Sept. 13, 2016), pp. 1657–1659.
- [50] *Molecules and compounds*. Khan Academy, 2017. URL: <https://www.khanacademy.org/science/chemistry/atomic-structure-and-properties/introduction-to-compounds/a/paul-article-2>.

- [51] *Bond-line structures*. Khan Academy, February 16, 2015. URL: <https://www.khanacademy.org/science/organic-chemistry/gen-chem-review/bond-line-structures/v/bond-line-structures-new>.
- [52] *Resonance structures*. Khan Academy, June 7, 2014. URL: <https://www.khanacademy.org/science/organic-chemistry/organic-structures/formal-charge-resonance/v/resonance-intro-jay>.
- [53] Timothy Soderberg. *Organic Chemistry with a Biological Emphasis Volume I*. 1-2016. Vol. 1. 2 vols. University of Minnesota, Morris Digital Well: Chemistry Faculty, 2016.
- [54] Nicholas E. Bodé and Alison B. Flynn. “Strategies of Successful Synthesis Solutions: Mapping, Mechanisms, and More”. In: *Journal of Chemical Education* 93.4 (Apr. 12, 2016), pp. 593–604.
- [55] Steve Krug. *Don't Make Me Think!: A Common Sense Approach to Web Usability*. Circle.com Library. Indianapolis: New Riders Publishing, 2000.
- [56] Nick Yee. *The Proteus Paradox: How Online Games and Virtual Worlds Change Us—And How They Don't*. New Haven, CT: Yale University Press, 2014.
- [57] *The World of Math*. Khan Academy, 2017. URL: <https://www.khanacademy.org/mission/math>.
- [58] Richard H. Thaler and Cass R. Sunstein. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New Haven: Yale University Press, 2008.
- [59] Nancy A. Hamilton and Rick E. Ingram. “Self-Focused Attention and Coping: Attending to the Right Things”. In: *Coping with Stress: Effective People and Processes*. Ed. by C. R. Snyder. New York: Oxford University Press, 2001, pp. 178–195.
- [60] C. R. Snyder and Kimberly Mann Pulvers. “Dr. Seuss, the Coping Machine, and ‘Oh, the Places You’ll Go’”. In: *Coping with Stress: Effective People and Processes*. Ed. by C. R. Snyder. New York: Oxford University Press, 2001, pp. 3–29.
- [61] Nigel Bevan et al. “The ‘Magic Number’ 5: Is It Enough for Web Testing?” In: *CHI '03 Extended Abstracts on Human Factors in Computing Systems*. ACM Press, 2003, pp. 698–699.

- [62] Aphra Kerr. “Player Production and Innovation in Online Games: Time for New Rules?” In: *Online Gaming in Context: The Social and Cultural Significance of Online Games*. Ed. by Garry Crawford, Victoria K. Gosling, and Ben Light. London: Routledge, 2011, pp. 25–39.
- [63] *Open Babel v2.3.0 documentation: Supported File Formats and Options*. 2010. URL: <http://openbabel.org/docs/2.3.0/FileFormats/Overview.html>.
- [64] Egon L. Willihagen et al. “The Chemistry Development Kit (CDK) v2.0: atom typing, depiction, molecular formulas, and substructure searching”. In: *Journal of Cheminformatics* 9.33 (June 6, 2017).
- [65] *JChemPaint*. In collab. with Egon Willihagen, Ralf Stephan, and Christoph Steinbeck. URL: <https://github.com/JChemPaint/jchempaint>.
- [66] Oliver Laitenberger and Horst M. Dreyer. “Evaluating the Usefulness and the Ease of Use of a Web-based Inspection Data Collection Tool”. In: *IEEE Comput. Soc*, 1998, pp. 122–132.
- [67] “Together, technology and teachers can revamp schools”. In: *The Economist* (July 22, 2017).
- [68] Daniel Kahneman. *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux, 2011.
- [69] Geoff F. Kaufman and Mary Flanagan. “Lost in Translation: Comparing the Impact of an Analog and Digital Version of a Public Health Game on Players? Perceptions, Attitudes, and Cognitions”. In: *International Journal of Gaming and Computer-Mediated Simulations* 5.3 (2013), pp. 1–9.
- [70] Caroline E. Morton et al. “Blended learning: how can we optimise undergraduate student engagement?” In: *BMC Medical Education* 16.1 (Dec. 2016).
- [71] Hans Christian Arnseth. “Learning to Play or Playing to Learn: A Critical Account of the Models of Communication Informing Educational Research on Computer Gameplay”. In: *Game Studies: The International Journal of Computer Game Research* 6.1 (Dec. 2006).
- [72] *Choice-Based Assessment (Part 1)*. MITx: 11.127x Design and Development of Games for Learning. edX, 2016.
- [73] Steven C. Pan. “The Interleaving Effect: Mixing It Up Boosts Learning”. In: *Scientific American* (Aug. 4, 2015). URL: <https://www.scientificamerican.com/article/the-interleaving-effect-mixing-it-up-boosts-learning/>.

- [74] *Values of the t-distribution (two-tailed)*. MedCalc Software, 2017. URL: <https://www.medcalc.org/manual/t-distribution.php>.