7-19-2022

The Process of Establishing a Blooming Chemistry Tool for Use in Undergraduate Chemistry Education and Research

Emryse Geye
Portland State University

The Process of Establishing a Blooming Chemistry Tool for Use in Undergraduate Chemistry Education and Research

by

Emryse Geye

A thesis submitted in partial fulfillment of the
requirements of the degree of

Master of Science
in
Chemistry

Thesis Committee:
Jack Barbera, Chair
Erin E. Shortlidge
Gwen Shusterman

Portland State University
2022

ABSTRACT

While it is still common in college chemistry to assess student learning and skill with summative assessments, the CER community does not currently have a simple tool to determine and communicate whether an assessment is actually aligned with the outcomes of interest. In particular, as so-called evidence-based teaching practices and active learning strategies gain a foothold in college chemistry classrooms, the ability to communicate whether those (often labor-, cost-, and time-intensive) interventions are not only aligned with course outcomes, but also provide measurable benefit to students becomes more imperative. While college chemistry has made some strides in the area of categorizing chemistry-specific cognitive skills, this work is largely disjointed and repetitive, making it difficult for community members to know what resources they have available, and how useful they are. This study developed a taxonomy of college chemistry problems, the Blooming Chemistry Tool (BCT), modeled after a successful discipline-specific translation of Bloom's taxonomy in college biology, but based on previous work in chemistry-specific cognitive skills. Once developed, the Blooming Chemistry Tool (BCT) was used in a qualitative study with members of the CER community to gain insight on how they engage with learning taxonomies and how they saw the BCT fitting into the greater CER

landscape. Interviewee suggestions for uses of the BCT spanned all components of the CER community–there were proposed uses for researchers, practitioners, and students, though there was the most consensus that due to its accessibility and ubiquity, that the BCT had the most potential as a tool for instructor training. The most significant finding from this study was that interview data suggests college chemistry assessment items do not have a single inherent sorting within the BCT dimensions–it really is all about context–which also means that the BCT could better establish another use for learning taxonomies: to structure the context necessary to compare classroom environments. However, while it was clear that all interviewees saw value for more consistent use of *a* learning taxonomy in CER, it's possible that the BCT may not meet that need for the CER community. While Bloom's taxonomy's quality of "brand recognition" was considered to be a positive attribute when selecting a base learning taxonomy for this project, it also means there would be "bad press" surrounding the still-widely-held criticisms of Bloom's original publication with which to engage when convincing CER community members that the BCT is both a useful and an acceptable tool.

DEDICATION

This thesis is dedicated to my mother, Mycroft, and Grandpa Max.

"Almost!"

ACKNOWLEDGEMENTS

This thesis would not exist without the kindness, consideration, and dedicated effort of several people. Beyond the contributions of my interviewees, the first of these is Dr. Jack Barbera Jr. who said he didn't want to drag me by my ear through this program and only kind of had to. Thank you for not giving up on me and not letting me give up.

Secondly, I would like to proudly acknowledge, for her help in coding and guidance, and being the first person it felt natural and easy to call my mentor, Dr. Nicole James of Reed College. Thank you for always reminding me that what I am experiencing is normal, and sometimes, not even my fault.

I want to thank Adrian Olives for being my partner in life and for growing up with me. I have known you half of my life, and half of that time was in graduate school! Thank you (and Langley Lou Whidbey Waffle) for choosing to see, love, and back who I ended up being.

For logistical support, I must thank my parents, Christian and David Geye, for generously supporting me through the last (couple of) years of my graduate work, as well as the rest of my family, Paul Geye and Paola Arbogast, for their unwavering belief in my potential. Thank you for your love. In this area, the Garlands also deserve thanks, specifically for hosting and feeding me while I wrote the last 100 pages of this thesis. Thank you for the never-ending sweet tea.

For emotional support, Dr. Cory Henson deserves a special shout-out for his gentle assurances that this would all be over one day. I also want to thank my gal pals, Melissa Garland, Rebecca Ator & Jocelyn Thompson for talking me into coming back to Texas (it's just not the same since I've been away.) Plus the rest of my cheering section, which includes but is certainly not limited to Eric Olivas, Bianca Phipps, Kimberly Durkin, Beth Grumer, Pooja Karnamadakala, Jeremiah Strong, Maxwell Peoples, Dr. Cody Clary, and Anders Horn, for being proud of me but also reminding me that I am worth more than my productivity. Further, I am grateful to Portland's Purrington's Cat Lounge, for providing the most wonderful breaks from thesis writing. Thank y'all for celebrating all of the (tiniest) milestones with me.

Finally, I would not exist without the hard work and diligence of Jessica Blashke and Dr. Keith Conant for providing me mental health care throughout my graduate program. Thank you for literally saving my life.

CONDENSED TABLE OF CONTENTS

COMPREHENSIVE TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE: INTRODUCTION

At its most basic, chemistry education research (CER) is interested not only in evidenced-based chemistry teaching practices, but in the act of measuring the effect any changes to teaching practices have on a multitude of student outcomes. In any endeavor of measurement, chemistry or otherwise, there are parameters which must be set: the investigator must know what they are intending to measure, how they are intending to measure it, and at what level of resolution. Measuring the effect an experiment has on mass or temperature is relatively simple; chemistry laboratories have tools specifically designed for these purposes. However, measuring the effect that changes to teaching practices have on students is much more abstract: how does one determine if a student or a class has shown evidence of improvement? What would improvement look like? What is a measurable outcome that can be used as a proxy for unmeasurable concepts such as improvement?

Like much of education, CER has a history of using student exams and course grades as proxies to measure student “improvement,” “success,” AND “understanding” especially when it comes to evaluating novel methods of teaching, such as active learning. Whether exam scores and class averages are reasonable measures of success and efficacy remains to be seen, and is beyond the scope of this project. However, as long as the CER community continues to use these proxies, we can at least endeavor to assure that these measures are aligned with our outcome of interest.

Statement of Problem

While it is still common in college chemistry to assess student learning and skill with summative assessments, the CER community does not currently have a simple tool to determine and communicate whether an assessment is actually aligned with the outcomes of interest, whether that be the course learning outcomes, or an inventory of concepts. Communicating course outcomes at a resolution higher than proxies such as final course grade would, of course, be a valuable ability for any practitioner to compare between classrooms, but being able to present evidence that assessments are aligned with the outcomes of interest has implications for other college chemistry educational measures, such as concept inventories and standardized exams. In particular, as so-called evidence-based teaching practices and active learning strategies gain a foothold in college chemistry classrooms, the ability to communicate whether those (often labor-, cost-, and time-intensive) interventions are not only aligned with course outcomes, but also provide measurable benefit to students becomes more imperative. In order to do that, CER needs a tool that not only captures what assessments are trying to elicit, but also which skills the CER community deems relevant to learning chemistry.

These tools, usually referred to as learning taxonomies, exist in education circles, often drafted by education generalists, though there are many examples of Discipline-Based Education Researchers specifying these measures to their own fields. However, this translation has not occurred yet in college chemistry, and the amount of time and effort it takes to appropriately apply generalist language and constructs to a particular disciplinary context is largely prohibitive to the segments of the Chemistry Education (CE) community that need it most, namely new instructors. In addition to reducing the time and effort needed by beginning teachers in the CE community to start metacognitively engaging with their own learning outcomes and assessments, such a taxonomy could also assist the CER community to better characterize learning environments and the nature of assessments. While college chemistry has made some strides in the area of categorizing chemistry-specific cognitive skills, this work is largely disjointed and repetitive, making it difficult for community members to know what resources they have available, and how useful they are.

Purpose of Study

It is the purpose of this study to develop a taxonomy of college chemistry problems based on previous work in chemistry-specific cognitive skills, and then to collect and analyze qualitative data about the use of that tool by members of the CER community. This qualitative data will provide insight on how members of the CER community engage with learning taxonomies and how they see a tool such as this fitting within the CER landscape. The research will answer the following questions.

Research Questions

1. What is the structure of, and what are the categories of skill necessary to construct a taxonomy of college chemistry problems?
2. How do members of the CER community, with prior experience sorting items by cognitive level, engage with such a taxonomy? What value do they anticipate this tool having in the CER community?

Significance of Study

This study developed a taxonomy of college chemistry problems, the Blooming Chemistry Tool (BCT), modeled after a successful discipline-specific translation of Bloom's taxonomy in college biology, but based on previous work in chemistry-specific cognitive skills. Beyond being the first discipline-specific translation of Bloom's taxonomy in college chemistry, the BCT was unique to the CER community in two ways: (1) rather than building a new tool whole-cloth, this taxonomy was based off of a well-known educational construct with which most educators have at least a passing familiarity, commonly known as Bloom's taxonomy, and (2) the BCT was developed with a goal of accessibility for all members of the CE community, not just for research members, to better align with the original intention of Bloom's taxonomy. Once developed, the Blooming Chemistry Tool (BCT) was used in a qualitative study with members of the CER community to gain insight on how they engage with learning taxonomies and how they saw the BCT fitting into the greater CER landscape. This approach was novel in that evaluations of learning taxonomies in college chemistry are usually limited to a proscriptive instructing of community members in how to use a given learning taxonomy, rather than a descriptive investigation into how taxonomies are actually used by community members. The resulting interviews provided valuable insight from CER community members regarding chemistry-specific language and concepts as well as assumptions about chemistry student reasoning and learning contexts. Interviewees also provided feedback on how the BCT could be improved.

The most significant finding from this study was that interview data suggests college chemistry assessment items do not have a single inherent sorting within the BCT dimensions–it really is all about context–which runs counter to previous forays into sorting college chemistry items by cognitive level, which have presented rankings for items as if they were inherent to the item itself and not based on the context in which it was asked. This is a substantial research finding that has implications for how the BCT might fit into the greater CER landscape. Interviewee suggestions for uses of the BCT spanned all components of the CER community–there were proposed uses for researchers, practitioners, and students, though there was the most consensus that due to its accessibility and ubiquity, that the BCT had the most potential as a tool for instructor training. Virtually all interviewees saw value in a Bloom's-like resource as a tool for *people who are starting to teach*–not to train them how to determine the "right" sorting for particular types of chemistry items, but rather to consciously engage with their teaching and assessment and develop an internal understanding of what sorting cognitive items looks like in their own classroom.

While an individual practitioner can be trained to use a tool such as the BCT to develop or improve a course by iteratively engaging in a cycle of (1) developing learning objectives (LOs), (2) using those LOs to develop their assessments, and then (3) evaluating what level their students are engaging at on those assessments, having evidence that items do not have an "inherent" sorting also means that the BCT could better establish another use for learning taxonomies: to structure the context necessary to compare classroom environments. A discipline-specific item sorting tool could serve as both the framework upon which to hang course outcomes, as well as the lens through which other CER community members can view those reports. Consistent use of such a tool could remove the burden for each CER community member to independently build the language necessary to communicate their outcomes, similar to how the Blooming Biology Tool has affected the Biology Education Research community. However, while it was clear that all interviewees saw value for more consistent use of *a* learning taxonomy in CER, it's possible that the BCT may not meet that need for the CER community. While Bloom's taxonomy's quality of "brand recognition" was considered to be a positive attribute when selecting a base learning taxonomy for this project, it also means there would be "bad press" surrounding the still-widely-held criticisms of Bloom's original publication (regarding its unidimensionality, assumption of cumulative hierarchy, and lack of theoretical basis/validity as a taxonomy,) with which to engage when convincing CER community members that the BCT is both a useful and an acceptable tool.

Limitations

Qualitative research is by nature interpretive. Therefore, the researcher being the instrument by which the data are collected and analyzed is the most general limitation to any qualitative study. For this study, a lot of the information elicited from the interviewees depended on the follow-up questions asked by the researcher. Thus, not all of the information that could have been pertinent to understanding the participant's sorting would have surfaced. Furthermore, in order to process the interview transcripts, the researcher needed to interpret the interviewees' responses and make judgements about their understanding. Therefore, the breadth and depth of interviewee understanding will not come across entirely in these interpretations. In fact, due to the limited number of items presented to interviewees and the limited number of interviewees themselves, the item sortings that were captured are by no means exhaustive. Also, interviewees were members of the CER community chosen for their experience sorting items by cognitive level either through the research or teaching; chemistry practitioners without experience sorting items were not interviewed.

List of Abbreviations and Glossary of Selected Terms

ACS: American Chemical Society

BBT: The Blooming Biology Tool

BCT: The Blooming Chemistry Tool

BER community: Biology Education Research Community

Bloom's Taxonomy: a complex and collaborative conglomerate of question classification taxonomies, starting with the original, subject-generalist Taxonomy of Educational Objectives conceptualized by Benjamin Bloom in 1956

CE community: Chemistry Education Community

CER community: Chemistry Education Research Community

Cognitive Process Dimension: One axis of the bidimensional Revised Taxonomy referring to cognitive abilities, skills, and behaviors

EF: The Expanded Framework for Analyzing General Chemistry Exams

HOCS: Higher-Order Cognitive Skills

Item Sorting: Placing an assessment item along a question classification taxonomy

Knowledge Dimension: The second axis of the bidimensional Revised Taxonomy organizing the discipline-specific types of knowledge students must know

LO: Learning Objective

Learning Objective: A discipline-specific statement of intention which contains both a verb and a noun

LOCS: Lower-Order Cognitive Skills

OT: the Original Taxonomy, referring to Volume I: the Cognitive Dimension of The Taxonomy of Educational Objectives written by Benjamin Bloom in 1956

Question Classification Taxonomies: Tools used to assess assessment items for the skills necessary to successfully produce an answer

RT: the Revised Taxonomy, referring to the most current subject-generalist version of Bloom's taxonomy written by Lorin Anderson and David Krathwohl in 2001

CHAPTER TWO: LITERATURE REVIEW

Frequently in education and educational research, the Taxonomy of Educational Objectives is referred to simply as “Bloom’s Taxonomy.” However, the term “Bloom’s Taxonomy” belies a more complex and collaborative conglomerate of metacognitive structures.

The Original Taxonomy (OT)

The original, subject-generalist Taxonomy of Educational Objectives resulted from several years of educator conferences consisting of primarily college and university examiners, chaired by Benjamin Bloom. The committee attempted to facilitate the exchange of assessment items among higher-education institutions and concluded that the skills that instructors may wish to assess in students consisted of three domains: cognitive, affective, and psychomotor. The cognitive domain refers to intellectual abilities and skills and categorizes the knowledge-based skill level necessary to accomplish a given task; while the affective domain refers to emotion-based skills and typically targets awareness and growth in attitudes, identity, and interest; and the psychomotor domain refers to physical or tactile skills. A volume on the cognitive domain, was published in 1956 by Bloom, and a second volume on the affective domain was published by Krathwohl, Bloom and Masia in 1964. While Bloom and colleagues did not publish a volume on the psychomotor domain, soon after, other educators proposed their own psychomotor taxonomies (Harrow, 1972; Simpson, 1972). The cognitive domain has been the primary focus of most traditional education and educational research, so most references to “Bloom’s Taxonomy” only allude to the first volume. To differentiate it from its subsequent iterations, in this project, the first volume corresponding to the cognitive domain will be referred to as the original taxonomy (OT.)

The OT is divided into six levels (with varying numbers of subcategories therein) referred to as knowledge, comprehension, application, analysis, evaluation, and synthesis. Frequently in educational literature, the OT is represented as a pyramid composed of six stacked levels, suggesting that knowledge is the foundation for all other skills, and that comprehension precedes application, and so forth. A representation of this pyramid format is provided in Figure 2.1. In this way, the OT was assumed to have a cumulative hierarchy, such that mastery of less complex behaviors (i.e., lower levels) was assumed to be a necessary but not sufficient condition for learning the more complex (i.e., higher levels) behaviors. As the OT was intended primarily for use in higher education, it makes sense that evaluation, defined as assessment of another’s work according to the criteria of their discipline, would be considered the ultimate cognitive behavior. The first category, knowledge, was considered to be somewhat distinct from the following five levels, as knowledge is necessary for all cognitive abilities thereafter to act upon, though it was still presumed to participate in the cumulative hierarchy.

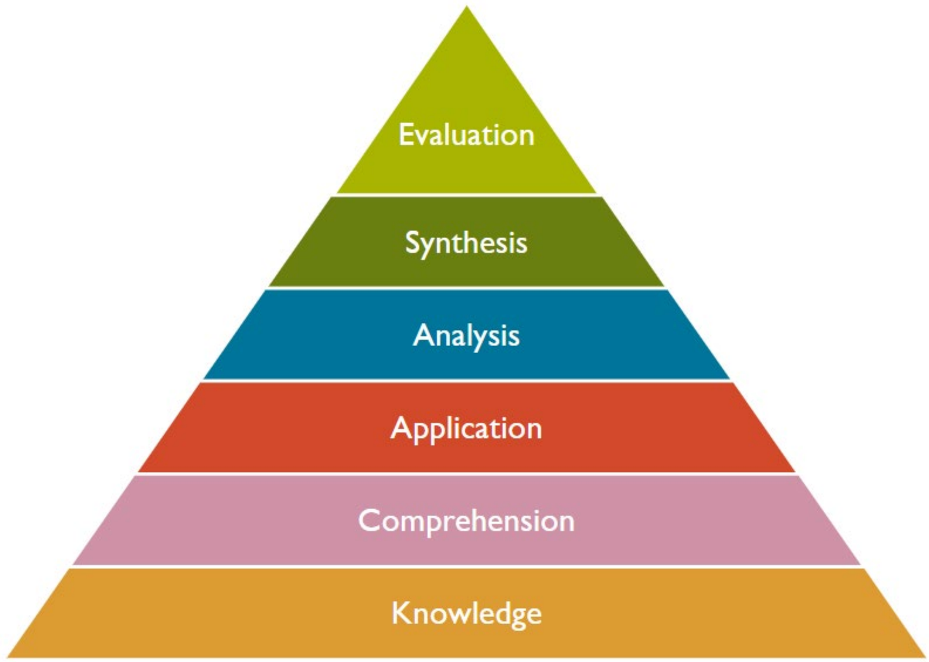

Figure 2.1. A Common Representation of the OT.

The OT was intended to assist an instructor in two ways: to assess student responses on assessments and to assess assessments. An instructor could use the OT when evaluating student responses on assessments in order to determine which cognitive skills the student has achieved. For example, if an assessment requires analysis, but the student displays difficulty accomplishing comprehension tasks, the instructor can use the OT as a framework to interpret increases or gaps in their cognitive skills. Alternatively, instructors can use the OT as a tool to develop assessments that might elicit a given level of cognitive skill from a student. By using the OT to jointly develop and evaluate student assessments, an instructor could, in theory, stratify students in a course by cognitive skill.

The layout of the OT makes evident that it was developed to assist college educators in creating exam items, since the handbook provided extensive banks of test items for each of the six categories. However, since its publication, the OT has become highly widespread in K-12 education (Flinders, Anderson, & Sosniak, 1996). Late in his career, Bloom even wrote, "unexpectedly, [the original taxonomy] has been used by curriculum planners, administrators, researchers, and classroom teachers at all levels of education" (Flinders et al., 1996, p. 1). K-12 educators primarily use the OT to develop learning objectives and class activities rather than assessment items. However, the ubiquity of the OT in K-12 education, in conjunction with limitations of its framework, has led to a few popular criticisms, primarily related to the validity of its structure—whether the original categories and subcategories constitute a sound taxonomy, whether the OT can be considered unidimensional, and whether the assumption of cumulative hierarchy is reasonable.

Criticisms of the OT

In 1965, soon after the publication of the affective volume of Bloom's Taxonomy, Morshead criticized the taxonomy as not being properly constructed. It was suggested that in order to be a taxonomy, (i.e., a scheme of classification,) there must be some systematic rationale of construction, and that Bloom's cognitive and affective domain taxonomies possessed "ontology without logic" (p. 165). However, Morshead also acknowledged the promise of such a project and noted that his criticisms were constructive. In this way, Morshead's more pressing concern was not that the OT itself lacked rigor, but rather that a system with such potential for the educational community should have a construction unassailable by misinterpretation.

Another significant criticism of the OT is its assumed unidimensional structure. In a unidimensional structure, knowledge cannot be both the basis of, and somewhat separate from the remaining levels. These criticisms argue that no thinking or cognitive process can exist without knowledge—in other words, one cannot think about nothing (Bailin, 2002; Booker, 2007.) In this way, the assumption of a unidimensional framework is unacceptable for some educators who feel that knowledge needs to be more closely associated with each individual level, rather than just with the first. Soozandehfar and Adeli (2016) suggest that by only partially removing the knowledge level from the rest of the structure in the OT, the knowledge category then possesses both the *noun* and *verb* forms of the term. While parts of speech may seem trivial in chemistry education research (CER), the assumption of unidimensionality in the OT asks students to both know the knowledge *and* remember it whereas none of the other levels explicitly contain this duality. In their 1994 review of additional studies investigating the cumulative hierarchy of the OT, Kreitzer and Madaus found that placement of knowledge anywhere within the framework of the OT was problematic, which provides evidence that the OT does not function unidimensionally.

Combining both the concerns of the lack of a systematic rationale and a unidimensional structure, the OT has been criticized for its assumption of cumulative hierarchy. A cumulative hierarchy would require that there can be no overlap between categories: one skill must be fully achieved before the next one can be attempted. This is problematic for three reasons. First, it would follow that even the most difficult questions of a lower category must require less cognitive skill than the simplest questions of the next-highest. For example, the most difficult comprehension question must elicit less cognitive skill than the simplest application question. However, within the CER community, it has been frequently noted that "plug-and-chug" items–which are often considered to exist at the application or analysis levels of the OT–are not necessarily more demanding than all types of comprehension questions (Nurrenbern & Pickering, 1987,) because any item can be made harder or easier at the behest of the examiner. Second, a cumulative hierarchy fails to acknowledge that students perform at varying levels of proficiency within each category of the OT. As Soozandehfar and Adeli (2016) note, novice learners still conduct tasks using all cognitive skills, they just do so at a more novice level. Third, a cumulative hierarchy implies that students that are successful on higher-level items must inherently be proficient at lower-level. However, Booker argued in 2007 that the structure of the OT encourages the dismissal of its lower levels as unworthy of teaching and has been used to devalue basic skills in education to the point of exclusion of anything other than higher-level thinking. Soozandehfar and Adeli (2016) went on to observe that some educators interpret the OT as having lower levels for teaching lower-level courses, while higher-level skills are only appropriate for advanced- or graduate-level courses. This leads to an extraordinary misreading of the OT, where educators may eschew lower-level items for higher-level items and those higher-level cognitive skills might be taught without the requisite foundation. Booker was concerned that students deprived of the fundamentals could not reasonably be expected to succeed in courses that emphasize only higher-level skills.

It is important to note that the OT did not suggest that any of the cognitive levels were less important, only that they followed a hierarchical structure. Kreitzer and Madaus’ 1994 review of additional studies investigating the cumulative hierarchy of the OT concluded that while there was evidence for ordering the lower levels of the OT (i.e., comprehension, application, and analysis) there was not as much evidence for ordering the higher levels of the OT (i.e., synthesis and evaluation.) However, eventually the educational community commuted the existence of "higher" and "lower" levels into higher-order cognitive skills (HOCS,) generally referring to the last three cognitive skills (analysis, synthesis, and evaluation,) and lower-order cognitive skills (LOCS,) generally referring to the first three cognitive levels (knowledge, comprehension, and application) (Allen & Tanner, 2002; Crowe et al., 2008; Zoller & Tsaparlis, 1997). This arbitrary delineation has had repercussions through all levels of teaching, including in CER.

The Educational Legacy of the OT in CER

In 1995, Zoller et al., combined their conception of the OT with Nurrenbern and Pickering's work in CER formalizing the dual conceptual and algorithmic nature of chemistry which resulted in their proposal of four recognized types of chemistry questions: algorithmic, conceptual, LOCS and HOCS. In 1997, Zoller and Tsaparlis defined HOCS in line with the OT as

> "Quantitative problems or qualitative conceptual questions, unfamiliar to the student, that require for their solution more than knowledge and application of known algorithms: they require analysis, synthesis, and problem solving capabilities, the making of connections and critical evaluative thinking including the application of known theory or knowledge to unfamiliar situations or situations with an unusual element or dimension. Such an application may further require (partially or fully) the abilities of reasoning, decision making, analysis, synthesis, and critical thinking," (p. 118).

Zoller and colleagues' work to investigate these categories (Zoller, et all., 1995, Zoller & Tsaparlis, 1997; Zoller, 2001; Zoller et al., 2002, Zoller &

Tsaparlis, 2003; Zoller & Scholz, 2004; Zoller, 2011) has been highly influential in the CER community. Unfortunately, in 2003, when Zoller and Tsaparlis found that students who did very well on HOCS did not necessarily score higher on LOCS, this evidence for a lack of cumulative hierarchy in the OT in CER led Zoller and Tsarpalis to reject the OT, and much of the CER community followed. This means many members of the CER community are unfamiliar with a revision of the OT that addressed most of its major criticisms, including that of an assumed cumulative hierarchy.

The Revised Taxonomy (RT)

In 2001, a major revision of the OT was undertaken by Lorin Anderson and David Krathwohl. Krathwohl worked with Bloom on the OT, and Anderson was one of Bloom's students. In their revision, Anderson and Krathwohl attempt to align the OT with more modern conceptions of cognition and to address some of the criticisms and limitations of the OT while maintaining a connection to Bloom's previous work. They believe that the OT's "continuous and widespread citation attests to its perceived value over time" (p. 302). Despite its critics, the effect that the OT has had on the educational landscape has been argued to outweigh any perceived lack of rigor in its construction (Allen & Tanner, 2002; Soozandehfar & Adeli, 2016.) By addressing these criticisms, Anderson and Krathwohl hoped to capitalize on the ubiquity of the OT while creating a revision that is more serviceable and user-friendly. In this project, this updated version of the cognitive dimension taxonomy will be referred to as the revised taxonomy (RT.)

Some of the changes Anderson and Krathwohl made to the OT are based in structure, some in terminology. The first of which is the matrix structure of the RT. A representation of the matrix form of the RT can be found in figure 2.2. Anderson and Krathwohl clarify that the structure of the OT did not appear to have an empirical basis at the time it was made, whereas the RT presents the six cognitive domain levels as one axis of a matrix, arranged by cognitive complexity. This change was made in order to separate out the skills and abilities from the OT (i.e., the verb component of knowledge) from the content those abilities and behaviors are used to process (i.e., the noun component of knowledge.) This axis of abilities and behaviors is now referred to as the cognitive process dimension, while the added axis is referred to as the knowledge dimension, and includes four categories of knowledge—three of which are a combination of the OT's 12 knowledge subcategories–with the addition of metacognitive knowledge, which reflects modern conceptions of cognition. This axis is arranged in order of abstractness, with factual knowledge being the least abstract and metacognitive knowledge being the most abstract. By separating out the knowledge dimension from the cognitive process dimension, Anderson and Krathwohl attempted to address two of the OT's main criticisms: its lack of systematic rational of construction (Morshead, 1965) and its problematic positioning of knowledge (Bailin, 2002; Booker, 2007; Kreitzer & Madaus, 1994; Soozandehfar & Adeli, 2016).

Knowledge Dimension	Cognitive Process Dimension					
	Remember	Understand	Apply	Analyze	Evaluate	Create
Factual Knowledge						
Conceptual Knowledge						
Procedural Knowledge						
Metacognitive Knowledge						

Figure 2.2. A Representation of the RT.

Other changes in the RT include renaming the cognitive levels. To start, all levels were switched to their verb rather than noun form (i.e., analysis became analyze.) Further, the second level, comprehension, was changed to the broader term understand, and synthesis was changed to create to emphasize the generation of original ideas and work. Finally, the two highest categories were switched in order—in the OT, synthesis precedes evaluate, while in the RT, evaluate precedes create. Anderson and Krathwohl argue that create involves induction, and induction is inherently a more complex cognitive process than deduction. Deduction involves breaking a whole into its parts, evaluating them, and determining whether criteria are met. Induction, on the other hand, involves finding things that could fit together, judging their appropriateness, and assembling them to best meet criteria. As an example, scientific writing is both created and evaluated, and the construction of new and unique writing is generally considered to be more cognitively taxing than the evaluation of another's writing. A comparison of the structures of the OT and RT is provided in table 2.1.

Table 2.1. A comparison of the structure of the OT and RT.

	Number of Dimensions	Presented as	Intended Use	Order of Cognitive Levels	Order of Knowledge Levels
OT	Unidimensional	A stacked pyramid of six levels	Help writing college exams	Knowledge Comprehension Application Analysis Synthesis Evaluation	–
RT	Bidimensional	A 6x4 matrix	Help creating K-12 learning objectives	Remember Understand Apply Analyze Evaluate Create	Factual Knowledge Conceptual Knowledge Procedural Knowledge Metacognitive Knowledge

Unlike the OT, which was developed by and for college and university examiners, the RT was designed to be of use for teachers of all levels, particularly keeping K-12 teachers in mind. Therefore, examples of use in the RT were shifted away from sample assessment items. Instead, the RT is mainly intended to assist in the planning of curricula, instruction, assessment and the alignment of these three. Such an alignment follows from the careful construction of learning objectives. Acknowledging the grammatical criticisms of the OT, Anderson and Krathwohl define a learning objective as “a statement of an objective [which] contains a verb and a noun” (p. 4). The verb describes the intended cognitive process and the noun describes the knowledge students are expected to acquire or construct. Thus, a learning objective can be assessed by placing it at the intersection of its two dimensions. For example, an item asking students to “apply the ideal gas law,” might exist at the intersection of apply and procedural knowledge, as would many plug-and-chug items (Nurrenbern & Pickering, 1987) whereas an item that asks students to “design a flow-chart for all the carbonyl reactions discussed in the previous chapter,” might be placed at the intersection of create and metacognitive knowledge.

While the knowledge dimension is considered to be arranged in order of abstractness and the cognitive process dimension is considered to be arranged in order of cognitive complexity, these categories are not subject to the same hierarchy as the OT. While Anderson and Krathwohl stress that the RT places a much greater emphasis on teacher usage than on developing a strict hierarchy, they do posit that the center of the presumed range of each of the six major categories of the cognitive process dimension should be successively greater in complexity. With this change in perspective and intended use, Anderson and Krathwohl attempt to address the criticisms that stem from the OT’s assumption of a cumulative hierarchy (Soozandehfar & Adeli, 2016; Zoller & Tsaparlis, 1997).

The Educational Legacy of the RT

Despite Anderson and Krathwohl's careful construction of the RT and their thoughtful responses to the original framework's critics, there are relatively few examples of the educational community using the RT or even acknowledging the revisions to the structural assumptions underlying the RT's framework. This lack of recognition can be at least partially attributed to the way the RT has been proliferated–the revisions and structure of the RT are often misrepresented, and therefore the criticisms of the OT are perpetuated and assumed to have been commuted to the RT.

Despite its explicit bidimensional nature, the RT is still most typically arranged, like its predecessor, in a pyramid. Such structures do not support the findings of studies such as those mentioned in Kreitzer and Madaus' 1994 review which only supported the hierarchical ordering of the lower levels of the OT. These misrepresentations perpetuate the idea that the OT and RT are interchangeable. In fact, many comparisons of the OT and RT summarize the revisions as the switching of the order of the highest two cognitive processes, such as in (the otherwise encouraging) Soozandehfar and Adeli (2016)! Both of these oversimplifications ignore that the RT was made bidimensional to assist in writing learning objectives, which according to Anderson and Krathwohl, require a verb (from the cognitive process dimension) and a noun (from the knowledge dimension).

It is important to recognize the proliferation of these misrepresentations when examining the research that has utilized the RT, as any conclusions may be based on a poorly structured depiction. It is also important to keep the various misinterpretations in mind when considering taxonomies suggested as more suitable alternatives to the RT due to its perceived failings, since these claims of inadequacy may be specious.

The Educational Legacy of the RT in CER

Tools used to assess assessment items for the skills necessary to successfully produce an answer are considered *question classification taxonomies*. The earliest and most well-known example of these is the OT. In the years since its publication, other taxonomies have been developed with the intention to supplement, clarify, or improve upon the OT but a recent review in CER of 41 question classification taxonomies for the purpose of analyzing verbal questions asked in class by chemistry teachers found that all the other taxonomies were reflections of earlier work by Bloom.

> "...[T]here were no new developments in terms of how classroom questions have been conceptualized in recent years that are qualitatively different from the way questions were characterized in much older studies. Instead, I noted a large extent of replication of question categories and an introduction of new terms to describe the same kinds of question categories or types that were originally identified. Though some authors introduced new terms to describe questions that occur in specific subjects, the descriptions accompanying those new terms implied similar meanings as those attached to related/similar question categories in earlier studies" (Kayima, 2018, p. 47.)

The most direct descendant of the OT is the RT, and while both have been used to assess HOCS on college chemistry exams (Bergendahl & Tibell, 2005; Tsaparlis & Zoller, 2003), they are also used to assess other aspects of chemistry courses, such as textbook items (Dávila & Talanquer, 2010; Gillette & Sanger, 2014; Pappa & Tsaparlis, 2011) and classroom response system questions (Bruck & Towns, 2009; Rodrigues, Taylor, Cameron, Syme-Smith, & Fortuna, 2010), and frequently these studies evaluate K-12 assessments (Edwards, 2010; Gillette & Sanger, 2014; Karamustafaoğlu, Sevim, Karamustafaoğlu, & Çepni, 2003; Ramirez & Ganaden, 2008; Rodrigues et al., 2010; Tikkanen & Aksela, 2012; Upahi et al., 2015; Upahi & Jimoh, 2016). Few of these studies applied the OT or RT in a congruent way with other researchers, as there is no comprehensive guide for the classification of college chemistry exam items. Of particular note, none of these studies attempt to adapt either the OT or RT to a chemistry-specific form.

According to Anderson and Krathwohl, the most successful discipline-specific applications of the OT were those who “adjusted the breaks between categories to fit their subject matter fields and created subcategories to highlight important discipline-related distinctions” (p. 301). Such a discipline-specific blooming tool would be useful for several reasons. First, a discipline-specific blooming tool could take into account the nuances of its community. For instance, the terms "analyzing" and "synthesizing" have discipline-specific meanings in chemistry, so a chemistry-specific blooming tool might use conceptions of these terms in the community to draw parallels with the appropriate cognitive levels, or it might opt for different language. Secondly, discipline-specific blooming tools would provide better resolution for users when trying to apply cognitive levels to items. For example, an organic chemistry assessment item asking a student to retroactively synthesize a molecule is not inherently an item that requires synthesizing or evaluation cognitive level skill, and a chemistry-specific blooming tool could address this dissonance. Finally, the subject-generalist OT and RT are strongly associated with the education communities where they were developed, and therefore may be seen as too broad to be useful to individual disciplines. Narrowing Bloom's taxonomy to discipline-specific blooming tools gives ownership to the community, and then becomes a construct that they can further adapt to discipline-specific needs. This has certainly been the case of the Blooming Biology Tool.

The Blooming Biology Tool (BBT)

The Blooming Biology Tool (BBT) was developed by Crowe, Dirks, and Wenderoth in 2008 to specify the OT for college biology assessments. The BBT translates each level of the OT to biology-specific constructs, concepts and skills, while also giving example question stems, applicable question types, and multiple-choice question suggestions. They furthermore developed additional tools that helped contextualize and root the BBT in their field. The Bloom's-based Learning Activities for Students, or BLASt helps both students and instructors choose studying or classroom activities at specific OT cognitive levels. For example, if a student prefers to study using flash cards but finds that classroom assessments are written at a higher cognitive level, they can consult the BLASt to find study activities that would push them to study at an appropriately higher cognitive level.

The response of the college biology education research community to the BBT suggests that if other discipline-specific translations of the OT or RT were developed, they would be well-utilized. Other college biology educators have used the BBT to assess concept inventories (Kalas, O'Neill, Pollock, & Birol, 2013), course exams (Newton & Martin, 2013; O'Neill, Birol, & Pollock, 2010; Zheng, 2008), and differences in student cognitive skill after implementing alternative learning environments (DeRuisseau, 2016; Elliott, 2010; Terry et al., 2012). Additionally, the BBT has been further translated into sub-discipline-specific Blooming tools, such as the Blooming Anatomy Tool (Thompson & O'Loughlin, 2015), the Visualization Blooming Tool (Arneson & Offerdahl, 2018), and the Bloom's Taxonomy Histology Tool (Zaidi et al., 2017).

The Expanded Framework for Analyzing General Chemistry Exams (EF)

The first foray in CER to adapt the RT to discipline-specific needs came from Smith et al., in 2010. This study generated a comprehensive guide for classifying general chemistry exam items by developing an Expanded Framework for Analyzing General Chemistry Exams (EF) based on the RT, as well as Zoller's work on HOCS in chemistry, and other influential research in the CER community over the last three decades about the types of skills that should be taught and assessed in chemistry classrooms (Tsaparlis & Zoller, 2003; Zoller, 1993; Zoller et al., 1995; Zoller & Tsaparlis, 1997; Niaz & Robinson, 1991; Nurrenbern & Pickering, 1987; Stamovlasis, Tsaparlis, Kamilatos, Papaoikonomou, & Zarotiadou, 2005.) As chemistry education is not a purely algorithmic discipline, most of this research calls for balance with respect to algorithmic and conceptual items (Zoller, Dori, & Lubezky, 2002). However, in order to structure exams with more balance, Smith et al., built a mechanism to provide detailed characterizations of chemistry questions and to imbue that mechanism with more utility than just classifying items as conceptual and algorithmic.

The EF links conceptual and algorithmic item types to the various terms suggested by members of the CER community. As stated earlier, Zoller et al., (1995) identified four types of chemistry items. In 1998, Robinson and

Nurrenbern described three broad categories of chemistry questions: recall, algorithmic and higher-order. Stamovlasis et al., (2004, 2005) categorized four types of chemistry questions: knowledge-recall, simple algorithmic, demanding algorithmic and conceptual. Each of these levels can be aligned with the RT, so Smith et al., arranged conceptual, algorithmic, and the additional terms along the dimensions of the RT to establish congruence of language. A reproduction of this table can be found in figure 2.3.

Cognitive process dimension	Knowledge dimension				
	Remember	Apply	Analyze	Evaluate	Create
Factual knowledge	LOCS[2], recall[3], knowledge-recall[4], information	LOCS[2], knowledge-recall[4]			
Conceptual knowledge		Conceptual[2,4], HOCS[2], higher order[3], problem solving[6], conceptual understanding[6]			
Procedural knowledge		Algorithmic[2,3], LOCS[2], HOCS[2], higher order[3], simple algorithmic[4], stoichiometric transformations[5], algorithmic application[6], problem solving[6]	HOCS[2], higher order[3], demanding algorithmic[4], problem solving[6]		

[1]Anderson and Krathwohl, 2001; [2]Zoller *et al*., 1995; [3]Robinson and Nurrenbern, n.d.a; [4]Stamovlasis *et al*., 2004; [5]Dori and Hameiri, 2003; [6]Hanson and Wolfskill, n.d.a.

Figure 2.3. Cognitive Skills in Chemistry Organized by Smith et al., Along Both Dimensions of the RT.

However, there are some inconsistent elements in the construction and assumptions of Smith et al's attempt at language congruence. First, both the OT and RT have six cognitive skill levels, but Smith et al.,'s reproduction of the RT only has five—Smith et al., did not include the cognitive process of understand, and did not give an explanation for its absence. Likewise, the knowledge dimension of the RT has four categories, but Smith et al.,'s reproduction of the RT only possesses three, missing the most abstract metacognitive knowledge level. Further, the additional CER terms are restrictedly plotted along both dimensions of the RT. Inherent to the structure of the RT, each cognitive skill level can be applied to any level of the knowledge dimension. Algorithmic cognitive skills, according to Zoller (1995) and Nurrenburn and Robinson (1987) would be congruent with the RT's apply level, and therefore would be applicable for all levels of knowledge, not just procedural knowledge. For example, using Boyle's law to explain what effect a decrease in volume would have on the pressure of a gas might be considered an algorithmic question at the apply cognitive skill level, but at the conceptual knowledge level. Ultimately, it is not congruent with the RT for additional CER cognitive skill terms to be locked into a particular type of knowledge.

The Legacy of the EF in CER

The EF consists of three major levels: definition, algorithmic, and conceptual. Smith et al., (2010) further refined and tested their framework by sorting the cognitive skill levels of the 40 items on both the first- and second-term American Chemical Society (ACS) Exams Institute Special Exam from 1998. According to the ACS Exams Institute, the Special Exam is the original version of a paired questions exam. Paired question exams are a series of exams designed by an ACS committee of chemistry educators to contain half conceptual items and half algorithmic items that cover a range of topics and types of questions representative of the general chemistry curriculum (Holme

& Murphy, 2011). ACS exams are useful instruments to analyze, because they are often used by many institutions as testing instruments and Smith et al., refer to them as "the recognized gold standard in multiple-choice chemistry exams." (p. 149). Smith et al., constructed their full framework based on the presumed cognitive skill level employed by students to answer each exam item. This framework was then used to code both terms of the ACS Special Exams. After several iterations of refining the EF and then re-sorting the items on the ACS Special Exams, Smith et al., further combined their primary levels into traditional skills, which included the definition and algorithmic primary levels, and conceptual skills. The final EF consists of three major levels containing 10 of sublevels therein. A representation of the EF is provided in figure 2.4. From this final form, Smith et al., found that both Special Exams were composed of roughly half traditional items and half conceptual items.

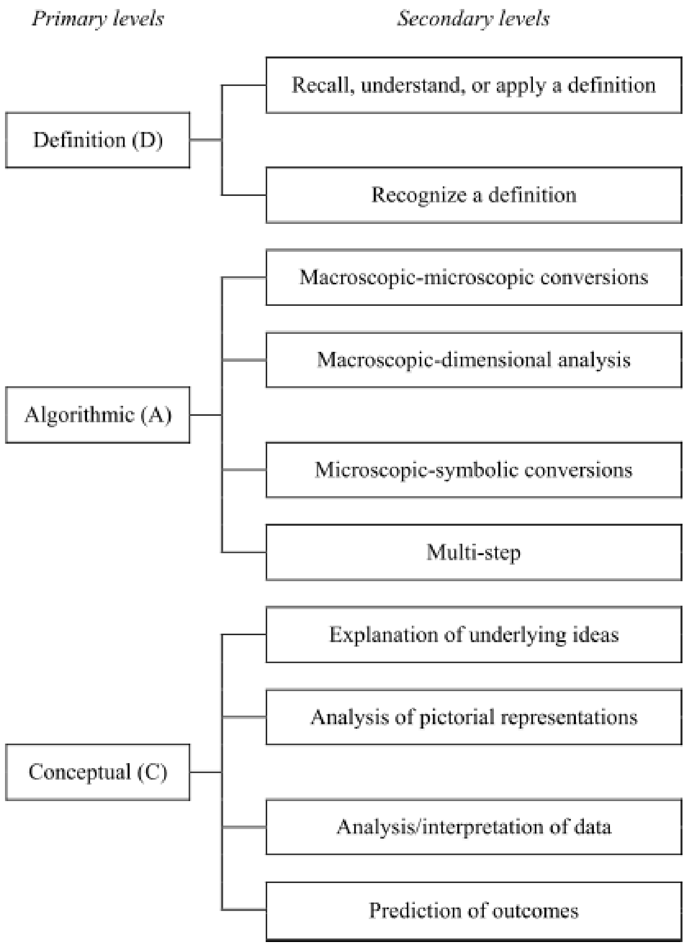

Figure 2.4. Depiction of the Expanded Framework for Analyzing General Chemistry Exams.

However, there are some concerning methodological choices in this study. First, the same coders developed the rubric as did all of the coding of exams, which introduces concerns that inter-rater reliability would be much lower for coders who were not so intimately familiar with the construction of the EF. Second, Smith et al., coded items based on the thought processes most likely employed by students to solve them, but did not appear to conduct any qualitative research to corroborate these with chemistry students' actual thought processes. Third, the EF appears to have been developed with the same items it was eventually used to assess–which might misrepresent the EF's applicability to other instruments. It would have been more compelling to test the EF on a separate exam, or else to develop the rubric with a subset of items from both exams, and then to test the EF against the remaining items. Finally, the researchers used a method to determine sufficient agreement between coders that involved, "dividing the total number of correct codes by the sum of the total number of correct and incorrect codes" (p. 149), but the researchers do not externally confirm an item's correct code beyond their own internal categorization. Alternatively, the researchers may have instead discussed a measure of inter-rater reliability that compares agreement of codes between coders, not a presumed correctness of an item that they cannot corroborate with an external source.

Beyond methodological issues of the construction of the EF, its structure is not grammatically consistent. The grammatical structure of the RT—verbs paired with nouns—makes it simple to discuss across levels. In the EF, both primary and secondary levels are a mixture of verb, noun, and adjective outcomes, which makes it difficult to compare primary or secondary levels. For example, the skill of recognizing a definition is not easily compared against a macroscopic-microscopic conversion—one is a skill, one is a type of item. Likewise, the primary level *definition* is a type of knowledge (noun), while the primary level *algorithmic* is a categorization of skills (adjective.) The EF was not designed to assist instructors in designing learning objectives, and therefore does not need to follow the strict arrangement of noun-versus-verb found in the RT, however, congruency in the language used in the EF would make it more straightforward to use. Finally, Smith et al., lose considerable resolution with which to discuss the highest HOCS—if an item requires the cognitive processes of evaluate or create, they cannot isolate those skills any further than being conceptual. Unfortunately, in light of this lack of resolution, their analysis of the ACS Special Exams is less compelling—roughly half the items were found to be conceptual, but according to the EF, the highest cognitive process the items could assess is analysis.

In a follow-up study, Sanabria-Rios and Bretz (2010) found that there was compelling evidence to use both the EF and the OT in tandem to categorize exam questions. This determination was made by applying the EF to a set of items written by college chemistry instructors in Puerto Rico, and then correlating each items' EF sorting with the items' OT sorting. They found that the EF's primary level of definition correlated significantly with the cognitive skill level of knowledge, but not with comprehension or application. This is inconsistent with the work of Smith et al., where application and comprehension's corollary in the RT are used to describe skills of the EF. Further, this also means that none of the EF primary levels correlated strongly with the OT cognitive level comprehension, while the EF primary level of algorithmic correlated significantly only with the OT cognitive level application, and the EF primary level of conceptual correlated significantly with the OT cognitive levels analysis, synthesis, and evaluation. By suggesting that the OT and EF should be used in tandem, Sanabria-Rios and Bretz (2010) recognize the lack of resolution in the conceptual primary level of their tool, however, they do not provide a justification as to why this follow-up study used the OT to compare against the EF, as opposed to the RT upon which the EF is based. The inconsistent construction and implementation of the EF may have led to it being proliferated throughout the CER community to a much lesser extent than the BBT was through the BER community.

A Blooming Chemistry Tool (BCT)

The BBT and the EF attempted similar research premises, using similar taxonomies, were published within two years of one another, and yet, the BBT has had much more depth of use and impact on its field. This project proposes that this disparity is due to two factors. First, the BBT is a specification of the OT to the discipline of biology—it takes a well-known subject-generalist educational construct that many educators have at least passing familiarity with, and considers it through the lens of biology.

Alternatively, the EF fit the RT into the language of the CER community and lost the familiarity of Bloom's Taxonomy in the process. Second, outside of the framework of the OT, the BBT is a much more structured and accessible tool—Crowe et al., followed the conventions of the OT, had a detailed structure at all levels, and used consistent language throughout the BBT. Conversely, the EF did not follow the conventions of either taxonomy, and its inconsistent underlying structure and assumptions may contribute to its lack of impact on the CER community. This project proposes that a structure can be made that combines both the RT and the EF into a single, easy-to-use tool. A Blooming Chemistry Tool (BCT) can be developed by combining the EF with the RT, translating it to chemistry-specific HOCS and conceptual skills (Holme & Murphy, 2011; Nurrenbern & Pickering, 1987; Tsaparlis & Zoller, 2003; Zoller et al., 1995), and modeling it after the more accessible structure of the BBT.

CHAPTER THREE: METHODOLOGY

This project consisted of two main phases–tool development, and tool testing–both of which were informed to different degrees by interviews with members of the Biology Education Research (BER) and Chemistry Education Research (CER) communities. Institutional Review Board (IRB) approval from Portland State University was received for all data collected within this study and appropriate consent was obtained from interviewees as required by the IRB. BBT insight interviews were conducted under IRB number 174336, while the pilot testing interviews were conducted under IRB number 207008-18.

BCT Development

Developer Insights about the BBT

Instrument development of the BCT started with investigating the BBT's construction and underlying framework. Two of the three authors on the original BBT publication (Crowe, et al., 2008) were interviewed to enquire about the development of the BBT: what inspired it, how its supplemental portions (such as the BLAsT) were developed, and how the BER community has responded to their work. All potential interviewees for this project were contacted via email; those who were interested in participating in an interview were directed to fill out a consent form and schedule an interview. All interviews were conducted by the same interviewer.

User Insights about the BBT

Recruitment Procedures

From the discussion of other researchers in the Developer Insight interviews, and from literature review of publications in BER that have cited the BBT, potential subjects for additional interviews were selected. All interviewees in the sample were doctorate members of the BER community who had applied the BBT to their teaching or research in college biology classrooms.

Data Collection

A total of 6 Insight interviews with 7 participants were conducted. All interviews for this phase were completed remotely over Zoom between 7-Feb and 28-Feb 2018. Interviews were conducted using a semi-structured interview approach. The participants were asked about how they had used the BBT in their teaching or research of college biology classrooms, and then for any critiques they had of the tool. Follow-up questions were asked as needed for further clarification. All interviews were audio and video recorded, and the interviewer took notes during the interview.

Data Analysis

None of the Insight interviews were transcribed for analysis. These interviews provided context on how members of the BER community used the BBT, and stressed the importance of usability and subject-specificity. Notes from the Insight interviews were referenced when developing a BCT Pilot.

Development of A BCT Pilot

The BBT Insight interviews informed the construction of a pilot version of the BCT, which drew its structure from the RT, its prioritizing of subject from the BBT, and its chemistry-specific language from the EF. Beyond that, a subject-specific blooming tool should be recognizable as a derivative of the RT, but distinct enough that there is no need to refer back to the subject-generalist RT. It was decided to keep the bidimensional matrix structure of the RT. The RT is a blank matrix, but the BBT provides example item stems at each cognitive skill level. It was decided to keep the BCT as a blank matrix, so that different “layers” could eventually be overlaid, such as item stems for specific topics or intersections of note for the CER community. It was decided to retain all levels of RT knowledge and cognitive skills, even though college chemistry classrooms may not regularly conduct assessments at all levels of the RT. For example, the assessment of metacognitive knowledge is still relatively rare in college chemistry classrooms, but it was decided for the integrity of the tool to keep these levels and let members of the CER community determine the boundaries of their own assessments.

Rather than utilizing the verbatim level headings from the BBT, RT, or EF, level headings were created to prioritize chemistry-based skills and types of knowledge. This process consisted of sorting items from ACS study guides and general chemistry textbooks along the RT, and then inserting appropriate verbs and nouns from those items into headings in the style of the RT while making particular effort to provide synonyms for common homonyms in chemistry (analyze, synthesize, etc.) A representation of the BCT Pilot can be found in Figure 3.1.

The Blooming Chemistry Tool Last updated: December 28, 2020) *Knowledge Dimension* / *Cognitive Processes*	**Remember** (LOCS)* Students are asked to retrieve, recall, identify, locate, find, recognize, list, label, define, state, repeat, reproduce, name	**Understand** (LOCS)* Students are asked to clarify, conclude, describe, explain, generalize, illustrate, interpret, paraphrase, represent, restate, summarize, symbolize, or exemplify in order to draw meaning from	**Apply** (LOCS/HOCS)* Students are asked, within a given context, to employ, calculate, carry out, change, conduct, execute, implement, infer, modify, operate, predict, use, utilize, or substitute	**Analyze** (HOCS)* Students are asked, in order to break material into its constituent parts and ascertain how those parts relate to one another, to attribute, compare, contrast, deconstruct, examine, parse, distinguish, reconstruct, outline, integrate, or structure	**Evaluate** (HOCS)* Students are asked to, based on criteria and standards, judge, assess, assign value, critique, detect, examine, test, check, discriminate, justify, reccomend, appraise, or grade	**Create** (HOCS)* Students are asked, in order to form a coherent or functional whole, to construct, design, hypothesize, plan, produce, reorganize, model, develop, devise, simulate, or draft
Factual knowledge the fundamentals (terminology, technical definitions, and specific details) students must know to be acquainted with chemistry.						
Conceptual knowledge the interrelationships among the basic components (classifications, categories, principles, generalizations, theories, models, and systems) within chemistry that enable them to function together.						
Procedural knowledge the skills (algorithms, techniques, methods) specific to chemistry and their criteria for appropriate practice.						
Metacognitive knowledge the awareness of their own cognition, as well as cognition in general.						

Figure 3.1 The Blooming Chemistry Tool (BCT.)

A selection of the items referenced to build the BCT Pilot headings were used as the sorting items for the next round of pilot-testing interviews in the second phase of the project.

Selection of Items for Sorting During Pilot Testing Interviews

When selecting items for interviewees to sort in the Pilot Testing interviews, an effort was made to choose items that spanned across the tool. Some of these items were chosen from the items referenced to build the BCT pilot headings, while the remaining items were adapted from other sources of chemistry items. Items that were expected to be more easily sorted came from the items referenced to build the BCT pilot headings, while items from other sources were intended to be more ambiguous. Figure 3.2 shows where researchers initially sorted each item for sorting along the BCT Pilot. These sortings were not intended to be used as an “answer key” against which to grade interviewee sortings, rather as an effort to maximize the spread of cognitive processes and knowledge addressed by the items presented to interviewees.

	Remember	Understand	Apply	Analyze	Evaluate	Create
Factual Knowledge	Q2					
Conceptual Knowledge		Q1	Q5	Q8	Q4	
Procedural Knowledge		Q3	Q7		Q6	Q9
Metacognitive Knowledge						

Figure 3.2 Initial Researcher Sorting of Items to be Sorted by Interviewees During BCT Pilot Testing.

Pilot Testing

The second phase of the project consisted of presenting the pilot version of the BCT to members of the CER and BER communities and asking them to sort selected items using the tool.

Recruitment Procedures

The sample for the pilot-testing interviews consisted of members of the Chemistry and Biology Education Research communities who had experience with sorting items by cognitive level, either through research or teaching. Potential participants were initially identified via a literature search that cross-referenced CERP and JCE publications over the last two decades with "Bloom's taxonomy". Results of this literature search were reviewed for the magnitude of their interaction with Bloom's taxonomy. Publications with significant levels of use of Bloom's taxonomy were collated, and contact information for corresponding and lead authors was collected.

Data Collection

Potential interviewees were contacted via email, which included a link to a Qualtrics consent and demographics survey as well as scheduling software. Once an interview was scheduled, the participants were sent a copy of the BCT and asked (but not required) to familiarize themselves with it before the interview. All interviews were completed over Zoom with audio and visual recording. A total of 18 interviews were conducted in this phase–16 with members of the CER community and two with members of the BER community. All interviews took place between 1-Jan and 31-Mar 2021. Interviews were completed remotely over Zoom using a semi-structured interview approach.

Interview Protocol

Section 1: Interviewees' Prior Learning Taxonomy Experience

The interviews consisted of three distinct, but related sections. The first of which related to the interviewee's familiarity with and perceived value of learning taxonomies. The interviewer asked about the participants' familiarity and experience with Bloom's Taxonomy and other learning taxonomies as well as about any publications of theirs in which they used the

RT or the OT in their teaching or research. Then, the researcher would briefly summarize the research project of which their interview was part, and ask the interviewee to describe what, if any, value they saw in a discipline-specific blooming tool. Example questions from this portion of the interview include: What value/useful data does sorting items by cognitive level have/produce for your teaching/research? Do you see this as an important need for the chemistry education community?

Section 2: Item Sorting using the Pilot BCT

In the second portion of the interview, the interviewer asked if the interviewee had any questions about the BCT that was emailed to them. From here, the interviewer would add a sample item to the chat function in zoom, read the item out loud and sort the sample item along the BCT. The interviewer would explain how that item could be sorted along the axes of the BCT and answer any questions the participant had about the structure or usage of the current iteration of the BCT. The sample item was chosen from the bank of potential sorting items–in the first two interviews, this selection was random, but by the fourth interview, after multiple interviewees had sorted multiple items into multiple intersections, it was decided to use the same sample item thereafter, and item Q5 was chosen for this purpose due to its very interpretive nature (i.e., an argument could be made for sorting it into a wide swath of intersections.) For this reason, item Q5 has few interviewee sortings relative to the other items.

Then the interviewer would have the interviewee sort additional items along the BCT. The nature of the items varied by participant–an effort was made to choose subject matter relevant to the interviewee's specialties. For this reason, some items (e.g. item Q4) have few interviewee sortings. In order to keep interviews within the 45-60 minute range, interviewees were asked to sort less items if their responses were particularly verbose. For this reason, some interviewees sorted less [1-4 items ($\bar{x}$ = 2.5 questions)] items.

All items were sent one at a time using the chat function in Zoom. The participant was asked to read each item out loud and then explain their process and reason for sorting. Thus, the interviews were considered to use a response-process approach, but rather than looking for evidence of response process validity on the solving of the chemistry item, evidence of response process validity was sought on how the participant categorized an item using the tool. Based on the depth and breadth of their response, the interviewer probed for clarification or asked natural follow-up questions as needed to gain a thorough understanding. Interview participants were encouraged to expand upon split cases–items that they felt did not cleanly fit inside a single BCT intersection. Depending on how much an interviewee had to say about a given item determined how many items were given to each participant.

Section 3: Debrief of the Pilot BCT

Finally, the third portion of the interview was a debrief of the interviewee's experience with the BCT in the previous segment of the interview. The interviewer asked if the interviewee encountered any issues of concerns during the prior interview portion that would prevent them from applying BCT to their teaching or research. Participants were also asked if they knew of other members of their education research community they would recommend for interviews. Interviewees were given the opportunity to pass along contact information for this potential interviewee to the researcher, or to pass on the researcher's interview call to the potential interviewee.

Data Analysis

Pilot testing interviews (with the exception of Interviews N and O) were transcribed using a professional transcription service, then checked for fidelity against their audio recordings. Responses to the middle Item Sorting section of the interviews were coded using both a deductive (corresponding to the interviewee's item sortings along the BCT) and an inductive approach (corresponding to the interviewee's reasonings for these sortings) using conventional content analysis (Hsieh & Shannon, 2005) with the software MAXQDA (Version 20.4.1).

Content Analysis

Conventional content analysis is an inductive method of analysis as compared to the more deductive directed content analysis where interview data is approached with *a priori* codes. The process of data analysis was carried out both independently and collaboratively by the researcher and a collaborator (Dr. Nicole James–Reed College) trained in chemistry education research. First, both authors familiarized themselves with the transcripts. Three of the fifteen transcripts were independently read multiple times to gain familiarity with the data. During these reads, each researcher recorded notes on patterns among the three transcripts. Due to the multiphasic nature of the interview data, after reading the first three transcripts, it was decided to structurally code the interviews (Saldana, 2015) to segment them into their three sections. Before applying the structural coding, the remaining 12 interviews were read for familiarity by both researchers, who would come together every 2-4 interviews to discuss impressions, relevant codes, and themes from the interviews as they were evolving.

Developing the Codebook

Once the interviews were read and examined for themes, an initial codebook was constructed by the researchers. This first codebook included the segmentation of the interviews and the deductive identification of all the BCT intersections at which the interviewee thought each item could conditionally exist. The codebook then addressed themes from the middle section of the interviews, regarding how interviewees sorted items along the BCT and the reasons they gave for those sortings. This codebook was then used and refined by the researchers to independently, iteratively, inductively code the first three interviews plus one additional interview, revising the codebook between each interview.

When deductively coding interviewee sortings, the researchers agreed on guidelines for what counted as a sorting. The interviewee had to use the cognitive process name (whether in verb or gerund form) to count as sorting an item at that level. However, just saying the name of a cognitive process or a knowledge dimension level at some point in their consideration was not sufficient. An interviewee's final sorting depended on both their most recent answer and the answer they confirmed when asked by the interviewer. Taking into account the cumulative nature of the cognitive skills categories, an item sorted as equally requiring two adjacent cognitive processes may just be coded as the more rightward one, (i.e., if an item required remember and understanding, this was coded as just understanding,) as the more rightward cognitive skill was assumed to surpass the leftward ones. However, if an interviewee said that an item required two adjacent skills, but leaned more towards the leftward cognitive process, then the item would be double coded. This process was also followed with the knowledge dimension. Another complication with coding was that some interviewees provided sortings in the form of intersections, while others provided a list of headers. In these cases, interviewees were coded as sorting at all possible intersections of the provided headings.

The codebook was considered complete when the two researchers reached a consensus that there were no more unique codes. This more final codebook was then used to code the next 9 interviews to consensus among the researchers. The primary researcher coded interviews N and O independently After all 15 transcripts were coded, the final codebook was analyzed for themes, and the codes themselves were revised and defined.

CHAPTER FOUR: RESULTS

Participants

A total of 15 interviews were conducted with members of the chemistry education research (CER) and biology education research (BER) communities. All participants had previous experience sorting items by cognitive level (though not necessarily along a version of Bloom's taxonomy) either through teaching or research.

Demographics

Interviews were conducted with 13 members of the CER community and 2 members of the BER community. All interviewees possessed doctorate degrees in either chemistry or biology, and were either faculty or staff at one of 13 institutions of higher education. The members of the BER community had further experience (beyond that noted above) using the Blooming Biology Tool for sorting items by cognitive level, and therefore were able to compare tools. Participant demographics including pronouns, field, title at the time of interview, and a qualitative description of their institution are provided in Table 4.1.

Table 4.1. Interviewee demographics, including pronouns, field, title, and a description of their institution.

Interviewee	Pronouns	Field	Title	Institution Description
A	he/him	CER	Assistant Professor	Large Midwestern state university
B	she/her	BER	Associate Professor	Midsize Midwestern state university
C	she/her	CER	Associate Professor	Large Southern state university
D	he/him	CER	Staff; Curriculum and Assessment Specialist	Large Midwestern state university
E	she/her	CER	Assistant Professor	Large Midwestern state university
F	she/her	CER	Associate Professor	Graduate level international university
G	he/him	CER	Professor Emeritus	Graduate level international university
H	he/him	CER	Associate Professor	Small Southeastern public university
I	he/him	CER	Assistant Professor	Large Midwestern state university
J	she/her	CER	Associate Professor	Small Southeastern private college
K	he/him	CER	Professor	Large Southern state university
L	he/him	CER	Assistant Professor	Small Southwestern private university
M	she/her	CER	Assistant Professor	Large Midwestern state university
N	she/her	CER	Associate Teaching Professor	Large Northwestern state university
O	she/her	BER	Teaching Professor	Large Northwestern state university

Items Sorted

The 15 interviewees provided 38 sortings for 9 items. Each interviewee sorted at least one item that was presented to at least three other interviewees. Figure 4.1 presents the items sorted by each interview in a stacked bar chart. Please note that Figure 4.1 does not present the items in the order they were presented to each interviewee.

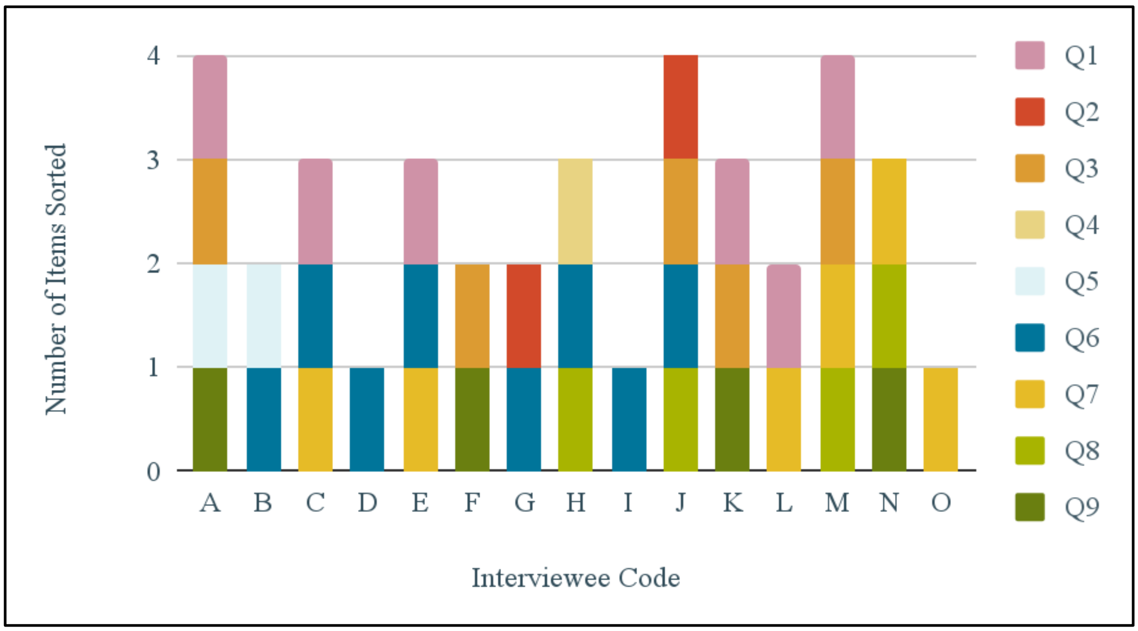

Figure 4.1. Items sorted by each interviewee. Each bar indicates the total number of items sorted by each interviewee.

Sortings and Codebook

A description of each of the nine items sorted by the interviewees is provided in Table 4.2. This table also denotes how many interviewees sorted each item.

Table 4.2. Items sorted along the BCT by interviewees as well as the number of interviewees who sorted each other.. The number of interviewees who sorted each item is presented with the item code.

Item Code	Item Wording
Q1 (n=6)	Which set of elements is arranged in order of increasing electronegativity? a. O < S < As < Ge b. Ge < As < S < O c. S < O < As < Ge d. As < O < Ge < S
Q2 (n=2)	Write the ideal-gas equation and give the units used for each term in the equation when R = 0.0821 L * atm / mol * K.
Q3 (n=5)	In sample Exercise 3.15, how do you explain the fact that the ratios C:H:O are 2.98:7.91:1.00 rather than exact integers 3:8:1?
Q4 (n=1)	A student standardizes a solution of aqueous NaOH against a measured mass of solid potassium hydrogen phthalate. She then uses this NaOH solution to titrate a measured mass of an unknown monocarboxylic acid to its phenolphthalein endpoint to determine its molar mass. Which errors will lead to a value of the molar mass that is too high? i. The potassium hydrogen phthalate is partially hydrated ii. The NaOH solution is allowed to stand after being standardized and absorbs some carbon dioxide from the air. a. i only b. Both i and ii c. ii only d. Neither i nor ii
Q5 (n=2)	One common test for water quality involves measuring dissolved oxygen, storing the water in a closed container at a constant temperature for five days, and then remeasuring the dissolved oxygen. If such a test shows a considerable decrease in dissolved oxygen over the five-day period, what can we conclude about the nature of the pollutants present?
Q6 (n=8)	Propose a method to separate each of the following mixtures: a) blood b) unrefined petroleum c) iron-sulphur (in powder form) mixture.
Q7 (n=6)	To the following containers that each contain an equal amount of water, the same amount of sugar is added. In which container is dissolution the fastest? a) cube sugar, 10 °C b) table sugar, 5 °C c) powdered sugar, 1 °C d) powdered sugar, 10 °C e) table sugar, 10 °C

Q8 (n=4)	A solution is prepared by dissolving 0.23 mol of hydrazoic acid and 0.27 mol of sodium azide in water sufficient to yield 1.00 L of solution. The addition of 0.05 mol of NaOH to this buffer solution causes the pH to increase slightly. The pH does not increase drastically because the NaOH reacts with the _________ present in the buffer solution. The Ka of hydrazoic acid is 1.9 x 10^-5 a) hydrazoic acid b) H2O c) This is a buffer solution: the pH does not change upon addition of acid or base d) H3O+ e) azide
Q9 (n=4)	By using a reaction flask, a manometer, and any other common laboratory equipment, design an experimental apparatus to monitor the partial pressure of $H_2(g)$ produced as a function of time.

Interviewee sortings of all items were deductively coded. Directed content analysis of the item sortings consisted of 19 *a priori* deductive codes corresponding to the six cognitive skill level categories of the BCT (Remember, Understand, Apply, Analyze, Evaluate, and Create), the four knowledge level categories of the BCT (Factual Knowledge, Conceptual Knowledge, Procedural Knowledge, and Metacognitive Knowledge) and the identification of the nine sorting items. A complete list of the inductive codebook and examples of each code from the interview transcripts can be found in Table 4.3. Interviewee transcripts have been lightly edited for brevity and clarity.

Table 4.3. The inductive codebook applied to interviewee sortings, definitions of the codes, and examples of codes within the text.

Inductive Code	Definition	Examples of Inductive Codes within the Text
Interaction with BCT	Interviewee refers to mechanics of using the BCT, asks clarifying questions for its use, or references wording of the BCT.	"So for, for procedural knowledge, um, is the concept, are the concepts that, like, underpin a procedure part of procedural knowledge or is that conceptual knowledge?" Interviewee M, on item Q3 "I'm more familiar with the cognitive levels and less familiar with the knowledge dimension. Um, and so that one, I am kind of reading these again and then for the cognitive processes, I am looking at the verbs." Interviewee J, on item Q3 "I saw…that explain, illustrate, [and] describe were next to each other and it bugged me." Interviewee I, on item Q6 "So I'm going to look across now knowing it's going to be a procedural knowledge and decide on the cognitive process dimension." Interviewee D, on item Q6
Internalized prior taxonomy experience	Interviewee's previous experience with a taxonomy (e.g. Bloom's, 3D-LAP) influences their interpretation of items/the BCT.	"I'm also relying a lot on…the sort of definitions that I've made for myself about what I think…students are doing. …[W]here I sort of did that is especially on…analyze. Um, the question I'm sort of asking myself is like, can our students, students are looking for what pieces are important. And for apply, I'm thinking about the, um, probably the, like, lower level of apply, which, which is we've done something similar before and now you're doing it, but with a different, like a different set of numbers." Interviewee J, on item Q3 "I find this much easier to put in a box with the 3D[-LAP] stuff, but it's what I'm used to. So that could be why." Interviewee I, on item Q6 "I don't think procedural has to involve numbers. I don't think it has to involve an algorithm or an equation or anything like that. I think it can be, um, uh, you know, 'I was shown this example of how to do this in class and can I apply that same process or thinking process.'" Interviewee E, on item Q6

		“So that's almost like a classic analysis task in my mind.” Interviewee B, on item Q5
Item Forensics	Interviewee uses elements of the item (e.g. multiple choice question format, verbs in item stem, etc.) to direct their item sorting.	“And it's not asking for calculations, like, it's not saying after you've done this, you know, like, what's your, what's the stoichiometry here or anything.” Interviewee K, on item Q9 “But because it's, again, a multiple choice question, I would probably stick it in, remembering factual.” Interviewee H, on item Q8 “So I guess that's how I'm thinking about it and interacting with the tool is…honing in on the verb and then trying to suss out, like, what is, what is it that we're actually asking, right? Because we can use lots of verbs that are, that would fall at different levels, depending on…the task it's tied to. And so that's, that's how I'm seeing this is, here's the verb. Here's what we're asking the students to do. Here's the, the broader task. And that tells me, okay, this is not an absolute remember task. This is this other type of a task, I think.” Interviewee B, on item Q5 “I don’t really see how application could be at play here, given how the question is worded. It just asked what do you conclude? It's not probing any deeper for anything that I think is a higher-order cognitive engagement.” Interviewee A, on item Q5
Pattern-matching	Interviewee assigns category by matching a verb from the item prompt directly to a verb listed in the tool as an example, without further consideration (of e.g. the category definition) – also includes implied use of pattern-matching (e.g. verbally uses a word from the question stem to justify assigning	“It’s asking the learner to explain. And so then, because you so kindly gave me the word explain in understand, I would probably put it there” Interviewee M, on item Q3 “So ‘compare, contrast, deconstruct, parse, distinguish, reconstruct, um, attribute, compare.’ I mean, there's a comparison here. You're comparing temperatures.” Interviewee E, on item Q7 “Now along the other upside here it says the word explain which, you know by just matching words would key into understand…” Interviewee A, on item Q3

	to a category with that word).	
Assumptions of student reasoning or learning context	Interviewee describes the expected or anticipated reasoning a student would use to address the question, or the assumed context in which the students would have learned the material.	"Understand, you know, if it is what we call algorithmic type of problem, the students have, uh, uh, solved many, many, uh, questions, many exercises with the ideal gas equation, and you don't have to understand, you just have to proceed with the application. So in this case, you jump here from remembering you go to application, without having to, to, to understand the thing." Interviewee G, on item Q2 "Um, that would depend on the context, I think, um, because the, so for example, I'm, I'm sort of picturing this being as part of an analytical chemistry course, or maybe because of the blood example a biochemistry course." Interviewee J, on item Q6 "Students are going to be like, uh, 'either I repeat the thing that you've told me before, or I'm going to sort of make something up from something I've experienced in the past' and so their, their procedural knowledge is going to be potentially problematic." Interviewee I, on item Q6 "Students would have to be familiar with the term electronegativity. And so then they're comparing things and so they would have to have an idea of what electronegativity means. Right. And so to me, when a student reads the word electronegativity, then they should have like a little, I guess, a little set of ideas in their head. Okay. Electronegativity means dadadadada and something is going to be more electronegative because of this and something is going to be less electronegative because of that." Interviewee K, on item Q1
Complexity	Interviewee references multiple knowledge elements or threads in presumed student reasoning, number of steps or tasks in item	"I wanted to make sure that indeed these elements were not all in the same column or all in the same row. Because if they were, I would have considered that to be truly factual because you usually memorize periodic patterns via those ways. This now mixes that up." Interviewee A, on item Q1 "So they'd have to break what is blood into its various parts. What is iron sulfur and what is

	stem or that a student must conduct.	petroleum? So…each as a mixture…would need to be broken down. Once it's broken down, they need to know how to actually separate that." Interviewee D, on item Q6 "The iron and sulfur example, is a standard one…but…things like blood or, uh, unrefined petroleum, uh, mo–more complex, of course." Interviewee G, on item Q6 "It's more complex than just…understanding because you're, you're, you're doing multiple parts." Interviewee H, on item Q4
Difficulty	Interviewee references difficulty of the question to themselves or students.	"Um, and i- it's like not trivial for students. This is a difficult thing for- for students to reason through th- in my experience." Interviewee M, on item Q8 "I forget what we did in the lab, but it's something where, like, they use a magnet and it just comes right out. And I was like, 'okay, like, that's an easy mixture.' But if you do like blood, does that increase the difficulty, which also increases the amount of skills and that's where it's harder to do." Interviewee D, on item Q6 "Oh, powder form. That's tricky. Anytime you've got stuff in powder. That's crazy. I wouldn't know how to do that." Interviewee B, on item Q6
Elimination	Interviewee uses process of elimination as a tool to assign BCT category (often used to narrow down to a couple of choices, which they then choose based on assumptions of student reasoning and learning context).	"Metacognitive, nope. They're not being asked to think about their thinking." Interviewee L, on item Q7 "Yeah, because I don't see the apply coming in with, based on the definitions here, then they don't have to calculate anything or execute anything." Interviewee C, on item Q7 "I don't think I would put it on evaluate. I don't think I'd put it apply or below." Interviewee F, on item Q3 "So you're not really clarifying, you're doing more than that." Interviewee H, on item Q4
Conditional	Interviewee indicates that the categorization is dependent on	"And that might be some procedural knowledge that might be that, you know, depending on the situation that might be metacognitive, because they might have to go back and think about when

	something (e.g. a specific learning environment or student).	they failed in a lab and what went wrong and whatnot, and just depending on the situation. Right. So this could literally be anything depending on what, what, not only the student knows, but what is the purpose of this question?...What, what, what is this designed to measure? What is this designed to elicit?" Interviewee A, on item Q9 "I feel like if this was a test question, I'm still comfortable in my analyze, but okay. We're gonna break this in and I'm proposing like a hypothetical thing that may or may not work, but if this was like an actual, like 'design a lab experiment and go and get the materials and actually do this experiment,' I feel like that would be create." Interviewee D, on item Q6 "I guess then this is where I hedge and so it really depends on the context. So if they've already learned exactly how to do this, so if it's saying 'design an experiment,' but really they've actually seen the design, then I would put this at remember I guess." Interviewee F, on item Q9 "So I think initially when you're presenting it in class and explaining it, I think if you ask this question, like as the first question that students do on this topic, then I would say that it would be analyze, um, conceptual knowledge because students are kind of still thinking about electronegativity and how it all works together and everything. But I think maybe after a few questions like this, then maybe they wouldn't think so much about the conceptual, um, aspect of it. And maybe, um, will just look at the factual knowledge." Interviewee K, on item Q1
Unconfidence	Interviewee indicates some uncertainty or unconfidence in their categorization of the item.	"I'm thinking procedural because they ha- you know, but they're not really actually doing it. So this it says- so maybe it's not procedural because they, you asking them to explain the method, but they not, I don't actually have to do a method." Interviewee C, on item Q6 "I think I'm going to go with either analyze or evaluate, um, and, and kind of stick it in there. Uh, do I need- would you like me to choose one, or?" Interviewee L, on item Q1

		"So I will place to uh procedural and evaluate…the box there…but I'm talking for myself. Now, if I was more knowledge about blood…it could be, it could be top left box as well…Well, for a microbiologist, you know, it's…uh, everyday routine." Interviewee G, on item Q6

Nine of these ten codes (all except "Difficulty") were used in interviewee sortings of item Q1. A description of the coding process for the interviewees' sortings of item Q1 is provided in this section, while a description of the coding process for the other eight items is provided in the appendix.

Item Q1

Item Q1 was sorted by the 6 interviewees A, C, E, K, L, and M. Item Q1 was the first question sorted by all interviewees except Interviewee E, where it was their last item that was sorted.

Interviewee A

Interviewee A began sorting of item Q1 along the Knowledge Dimension, saying,

> "I think this is factual knowledge. I looked [away to check a periodic table] because I wanted to make sure that indeed these elements were not all in the same column or all in the same row. Because if they were, I would have considered that to be truly factual because you usually memorize periodic patterns via those ways. This now mixes that up, but I'm not entirely sure that that still moves it out of the factual knowledge, because…I don't believe you need a concept here. I

believe you need to have memorized patterns up and down and then just sort of use those memorized patterns. I don't see a chemistry concept that has anything to do with the structure of the atom or electrons or anything like that here needed to solve this problem.”

Because Interviewee A concluded that item Q1 required only factual knowledge, this excerpt was coded deductively with the *a priori* code Factual Knowledge using directed content analysis. The cognitive response process of Interviewee A’s defense of his sorting was coded inductively using conventional content analysis as follows.

Interviewee A sorted item Q1 as requiring factual knowledge based on the reasoning he assumed students would undergo when tasked with this item (lines 7-8; coded as Assumptions of Student Reasoning or Learning Context,) and how complex he assumed this item would be for them (lines 1-7; coded as Complexity.) While he was uncertain whether the lack of proximity of the elements in item Q1 was sufficient to elevate it from factual to conceptual knowledge (lines 5-6; coded as Unconfidence), he was also relying on his prior experience with Bloom’s taxonomy to determine what he considered to be factual knowledge (lines 3-6; coded as Internalized Prior Experience) and used that to eliminate conceptual knowledge (lines 5-10; coded as Elimination.)

Moving on to sort item Q1 along the Cognitive Process Dimension, Interviewee A said, “I find this very low on the…x-axis, either remember or understand. Now that is, it is true that these are elemental symbols and [the students] do need to know what they are and they do need to find them on the periodic table if they don't have that sort of in their brain. And so that is all sort of remembering.“

Interviewee A narrowed his potential sorting to between remember and understand (lines 11 and 14-15; coded as Unconfidence) along what he referred to as the “x-axis” of the tool (line 11; coded as Interaction with BCT.) Interviewee A once again referred to the reasoning he assumed students would use when approaching this item (lines 12-14; coded as Assumptions of Student Reasoning or Learning Context,) and tentatively determined that all of that reasoning could be considered a remember process.

Then Interviewee A considered the format of item Q1, saying,

“I think that since it's a multiple choice question and there can be other problem solving strategies here that don't require sort of an explanation or anything like that. I'm going to put this in the remember category.”

Because he thought this item was answerable through memorization or through test-taking strategies due to the nature of the item (lines 16-17; coded as Item Forensics,) this excerpt was coded deductively with the *a priori code* Remember. Interviewee A went on to explain,

“There was a tendency for me to be like, Oh, you're applying a periodic trend, but that's, I don't think that's what's going on. That's not, that's not really, what's meant by, by ‘apply.’ [The students are] remembering periodic trends and sort of…using them to figure out which of these four things are the right answer.”

Interviewee A refutes his instinct to sort item Q1 as requiring the cognitive process of apply (lines 19-21; coded as Unconfidence) by citing the reasoning he assumed students would use when approaching this item (lines 21-23; coded as Assumptions of Student Reasoning or Learning Context.)

When asked for clarification about what “applying a periodic trend” meant, Interviewee A expanded on the reasoning he assumed students would use when answering this item (lines 24-28; coded as Assumptions of Student Reasoning or Learning Context.)

> “I believe that they are remembering either through some just memorized thing, rote learning, or sort of a heuristic that they use with the periodic table…students tell me that they, they envision like pictures superimposed on the table to help them remember these things. But I, I do, yes, I do. I don't think they're applying here.”

From these excerpts Interviewee A was coded as sorting item Q1 as existing at the most upper-left BCT intersection of Remember Factual Knowledge, a visualization of which can be seen in Figure 4.2.

Interviewee C

Interviewee C also began her sorting of item Q1 along the knowledge dimension, saying,

> “So I'm going to try looking at the…vertical column first in terms of knowledge dimension. So it's not metacognitive, it's not…So between conceptual knowledge and factual.”

Interviewee C narrowed her potential sorting to between factual and conceptual knowledge (lines 2-3; coded as Unconfidence) along what she referred to as the “vertical column” of the BCT. As she worked her way up the knowledge dimension, she used the heading descriptions (lines 1-3; coded as Interaction with BCT) to remove metacognitive and procedural knowledge from consideration (lines 2-3; coded as Elimination.) Then Interviewee C considered her two remaining options,

> “And now those two…I know there's…those two components. I don't know. Do you have to pick one or the other? …it has to be one because I think it's…a little piece of both because, yes, there is a factual piece involved there because [the students] have to know…what is electronegativity and…how it goes on the periodic table, but I will lean maybe towards, conceptual knowledge in this case because, um, because [the BCT] says, ‘interrelationships among basic components.’ So it seemed…’theories, models, general systems.’ And so I think they have to know…the electronegativity increases from left to right, or from the bottom to top of the periodic table. And, you know, in order to…they could just memorize it. But…I would lean towards conceptual, but I think there is some factual in there as well.”

Because Interviewee C concluded that item Q1 required both factual and conceptual knowledge, but leaned towards the more abstract conceptual knowledge, this excerpt was assigned the *a priori* code Conceptual Knowledge using directed content analysis. Interviewee C sorted item Q1 as requiring conceptual knowledge based on the reasoning she assumed students would undergo when tasked with this item (lines 7-8 and 11-13; coded as Assumptions of Student Reasoning or Learning Context,) and on the knowledge dimension header language (lines 9-11; coded as Interaction with the BCT) but throughout the process she questioned both choices and whether she had to pick one (lines 6-9 and 13-14; coded as Uncertainty.)

Moving on to sort item Q1 along the Cognitive Process Dimension, Interviewee C said,

> “...In terms of the cognitive process, I'm just reading through. I would say somewhere in range between remember and understand, because again, [the students] can recall just maybe from a list or where things should be, but I think understand because they have to kind of generalize or explain…which one it is based on their knowledge of…how things should go.”

Because Interviewee C sorted item Q1 as requiring either remember or understand cognitive processes depending on the line of reasoning she assumed a student would use when presented with this item (lines 16-19; coded as Conditional,) this excerpt was deductively coded as both Remember and Understand. Like with the Knowledge Dimension, Interviewee C sorted item Q1 based on the reasoning she assumed students would undergo when tasked with this item (lines 16-19; coded as Assumptions of Student Reasoning or Learning Context,) and on the Cognitive Process Dimension header language (line 15; coded as Interaction with the BCT.)

From these excerpts Interviewee C was coded as double sorting item Q1 as existing at both the BCT intersection of Remember Conceptual Knowledge and of Understand Conceptual Knowledge, a visualization of which can be seen in Figure 4.2.

Interviewee E

Interviewee E began her sorting of item Q1 along the Cognitive Process Dimension, saying,

> "It's factual. I'm not sure that it gets too conceptual, 'cause I think if you can simply remember periodic trends…you're gonna be able to do that. So that puts it also squarely in remember, so I think this is factual and remember."

Because Interviewee E concluded that item Q1 should be sorted as existing at the BCT intersection of Remember Factual Knowledge, this excerpt was assigned the *a priori* codes of Remember and Factual Knowledge using directed content analysis. Interviewee C made this determination based on the line of reasoning she assumed students would undertake when tasked with this item (lines 1-2; coded as Assumptions of Student Reasoning or Learning Context.) Interestingly, while she was at first uncertain whether this item would elicit conceptual knowledge, Interviewee E used her determination of the necessary cognitive process to triangulate her sorting and eliminate conceptual knowledge (lines 1-3; coded as Elimination.)

Interviewee E then moved along both dimensions of the BCT (lines 4-14; coded as Interaction with BCT) to check her work, saying,

> "And if we're going down…I don't think it quite needs even a conceptual understanding of electronegativity because you can answer it simply with periodic trends. I really don't know that I think it's procedural cause it is a remembering thing, and it's definitely not metacognitive. Certainly remembering, I don't think that you even have to understand what electronegativity is to answer this. I think you just have to remember periodic trends. I don't think this is even applying periodic trends because I think it's, again, remembering. It's not analyzing anything that's…in front of them because…they're not presented with raw data and asking them to orient the electronegativity that way. And similarly not evaluate because…they're not discerning like someone else's ranking to this

> or…given something like that and it's definitely close-ended. It doesn't feel like a create. So I think it's a remember factual knowledge."

In this excerpt, Interviewee E used the line of reasoning which she assumed students would undertake when tasked with this item (lines 4-14; coded as Assumptions of Student Reasoning or Learning Context) as a touchstone to eliminate the remaining categories of both BCT Dimensions.

From these excerpts Interviewee E was coded as sorting item Q1 as existing at the BCT intersection of Remember Factual Knowledge, a visualization of which can be seen in Figure 4.2.

Interviewee K

Interviewee K began his sorting of item Q1 along the Cognitive Process Dimension, starting with an explanation of the reasoning he thought the item was trying to elicit, saying

> "And so here we have increasing electronegativity and so students would have to be familiar with the term 'electronegativity.' And so then they're comparing things and so they would have to have an idea of what electronegativity means. Right. And so to me, when a student reads the word electronegativity, then they should have like a little, I guess, a little set of ideas in their head. 'Okay. Electronegativity means dadadadada and something is going to be more electronegative because of this and something is going to be less electronegative because of that.' Right.
>
> "I'll start with the, I think with the X[-axis] I think so. Yes…So here, I guess I would say analyze 'students are asked in order to break material into its componen-, constituent parts and ascertain how the parts relate to one another…compare contrast'…so I would say analyze, because you're comparing electronegativities of different

> elements. So let's say analyze…because students are being asked to compare the electronegativities of those different elements.”

Because Interviewee K concluded that item Q1 required only the analyze cognitive process, this excerpt was assigned the *a priori* code Analyze using directed content analysis. Interviewee K sorted item Q1 as requiring the cognitive process of analyze by citing both the reasoning he assumed students would use when approaching this item (lines 1-8 and 13-14; coded as Assumptions of Student Reasoning or Learning Context) and the reasoning he thought the item was trying to elicit (lines 15-16; coded as Item Forensics) as well as the heading descriptions along what he referred to as the “X-axis” of the tool (lines 10-13; coded as Interaction with BCT.) This excerpt was also coded as Unconfidence to reflect Interviewee K’s uncertain language in sorting this item.

Moving on to sort item Q1 along the Knowledge Dimension, Interviewee K said,

> “Now, depending on how the students have learned it, I guess it could be factual knowledge. Y'know like if students kind of say… ‘electronegativity increases this way and that way,’ then that's what they know. Like that's the little bit of factual knowledge. And so then they can compare the elements in terms of where they are and then they can answer it like that. And so it could be analyze factual knowledge, but I think it could also be analyze conceptual knowledge…because we're looking at theories and models of electronegativity.”

Because Interviewee K sorted item Q1 as requiring either factual or conceptual knowledge depending on the model of knowledge he assumed a student would use when presented with this item (lines 17-24; coded as Conditional,) this excerpt was deductively coded as both Factual and Conceptual Knowledge. Like with the Cognitive Process Dimension, Interviewee K sorted item Q1 based on the reasoning he assumed students would undergo when tasked with this item (lines 17-21; coded as Assumptions of Student Reasoning or Learning Context,) and on the reasoning he thought the item was trying to elicit (lines 22-24; coded as Item Forensics.) Interviewee K goes on to draw parallels between the BCT and his preferred dynamic.

> "I think this is kind of like the exercises versus a problem, right, so I think initially when you're presenting it in class and explaining it, I think if you ask this question, like as the first question that students do on this topic, then I would say that it would be analyze conceptual knowledge because students are kind of still thinking about electronegativity and how it all works together and everything. But I think maybe after a few questions like this, then maybe they wouldn't think so much about the conceptual aspect of it, and maybe, will just look at the factual knowledge…like 'okay. It increases up and that way.' Yes."

While Interviewee K is deferring to his preferred method of sorting items by cognitive level (lines 25-33; coded as Internalized Prior Taxonomy Experience,) this mode of thinking also shows the contextual dependence of item sorting (lines 25-33; coded as both Conditional and Assumptions of Student Reasoning or Learning Context.)

From these excerpts Interviewee K was coded as double sorting item Q1 as existing at both the BCT intersection of Analyze Factual Knowledge and Analyze Conceptual Knowledge, a visualization of which can be seen in Figure 4.2.

Interviewee L

Interviewee L began his sorting of item Q1 along the Knowledge Dimension, saying,

> "...If I follow the same steps you did, and I'm not asking anything
> metacognitive, I'm not asking, or this question is not
> asking…something that would cause students to be aware of their own
> thinking…procedural…'skills, algorithms, techniques, methods
> specific to chemistry and their criteria for appropriate practice.' So
> 'which set of elements is arranged'...this would not be a procedural
> question. It's not about a…'skill, technique, method, specific to
> chemistry.' It is a ranking question and I don't see that as a skill, per
> se; I see that as a concept. So [the students are] looking at the
> interrelation between electronegativity and how that electronegativity
> concept is associated with the elements, assuming periodic trends.
> "That's my assumption. It doesn't say that there, but that would be
> what I would expect my students to be noticing in a question like this."

Interviewee L sorted item Q1 by following the same steps the interviewer previously took in sorting the example item, specifically looking at the reasoning he thought the item was trying to elicit (lines 1-8; coded as Item Forensics) and the reasoning he thought students would employ when approaching item Q1 (lines 8-12; coded as Assumptions of Student Reasoning or Learning Context.) As he worked his way up the Knowledge Dimension, he used the heading descriptions (lines 3-7; coded as Interaction with BCT) to remove metacognitive and procedural knowledge from consideration (lines 1-7; coded as Elimination.) Then Interviewee L considered his two remaining Knowledge Dimension options, saying

> "So 'which set of elements is arranged in order of increasing electronegativity,'...'in a relationship among the basic components, the fundamentals,' I mean, this is kind of like what you were saying, the fundamentals, it could be a factual question, um, based on just like, 'this is the trend, this is electronegativity and, uh, these are the facts.' However, I see electronegativity and its relationship to the elements and the periodic trend…as an interrelationship…I see that as more conceptual…I'm just reading these again…'factual: the fundamentals, terminology, technical definitions, and specific details students must know to be acquainted with chemistry,' actually, hold on, let me think about this. 'Conceptual: the interrelationships among the basic components, classifications, categories, principles, generalizations, theories, models, and systems within chemistry that enable them to function together'...I think I'm going to stick with it being a conceptual question. As much as I could see electronegativity…having a technical definition…I'm going to stick with conceptual and I think I will do that because I see it being an interrelationship between …electronegativity and the elements that's manifested in periodic trends as a concept. And, okay, so there's the knowledge dimension. I think it's conceptual knowledge that they'll be working with here."

Because Interviewee L concluded that item Q1 required only conceptual knowledge, this excerpt was assigned the *a priori* code Conceptual Knowledge using directed content analysis. Though Interviewee L eventually sorted item Q1 as requiring conceptual knowledge based largely on the heading descriptions (lines 14-15, 19-25, and 30-31; coded as Interaction with BCT) paired with the reasoning he thought the item was trying to elicit (lines 13-17; coded as Item Forensics), he did question whether the item should be coded for the concept of electronegativity or the fact of periodic trends.

Moving left along the Cognitive Process Dimension, Interviewee L considered the three categories of create, evaluate, and analyze for sorting item Q1, saying,

"...'Students are asked in order to form a coherent or functional whole to construct, design, hypothesize, plan, produce, reorganize, model, develop, devise, simulate, or draft, create'...'which set of elements is arranged in order of increasing electronegativity?' So [the students are] given the elements and they are supposed to, I would say evaluate–that's kind of where my mind jumps to right now is, they're given these elements, they are assigning essentially value to them based on their understanding of periodic trends and then choosing the appropriate order for that, so I would not say it's create, they're not asked 'to form a coherent or functional whole or construct, design,' et cetera. It appears as though they are being asked to use their conceptual understanding to apply some kind of value or assigned value to each of these, ranking them. 'Analyze: students are asked in order to break material into its constituent parts and ascertain how those parts relate to one another to attribute, to compare contrast, deconstruct, examine, parse, distinguish, reconstruct, outline, integrate, or structure.'…I could hear some arguments for analyze, since you're given thr- or four different like samples, one might say, and look at how 'breaking those apart into their constituent pieces to understand how they relate to one another' being their electronegativity.

"I see this more as a[n] evaluate, I think in that [the students are] assigning specific values…'based on criteria and standards.' I might be arguing for analyze. Now, let me, let me keep working backwards. 'Analyze: students are asked within a given context to employ calculate, carry out, change conduct, execute, implement, infer, modify, operate, predict, use, or substitute'... 'students are asked within a given context to employ. So I don't see 'which set of elements is arranged in order of increasing electronegativity'...they've got the concept electronegativity, I think they, that they're given a set of data and I think that puts it in analyze and evaluate."

Interviewee L narrowed his potential sorting to between analyze and evaluate (lines 37-38 and 47-60; coded as Unconfidence) based on both the reasoning he assumed students would use when approaching this item (lines 36-40 and 47-53; coded as Assumptions of Student Reasoning or Learning Context,) and the reasoning he thought the item was trying to elicit (lines 35-36, 41-44, and 57-60; coded as Item Forensics.) Interviewee L once again used the heading descriptions (lines 33-35 and 41-57; coded as Interaction with BCT) to remove Create from consideration (lines 33-42; coded as Elimination.)

Interviewee L continued to move left along the Cognitive Process Dimension of the BCT (lines 61-65; coded as Interaction with BCT) to check his work and eliminate all remaining cognitive processes (lines 61-65; coded as Elimination,) saying,

> “They're not asked to clarify anything or explain anything. They're not asked to recall anything in particular besides maybe what is electronegativity. They are applying the concept of electronegativity, so that gets into apply, but they're applying that concept to…a set of information and determining which one is correct. I think I'm going to go with either analyze or evaluate and kind of stick it in there. Uh, do I need, would you like me to choose one or?”

Because Interviewee L could see arguments for both analyze and evaluate and questioned if he was required to choose one (lines 65-66; coded as Unconfidence) these excerpts were deductively coded as both Analyze and Evaluate and Interviewee K was coded as double sorting item Q1 as existing at both the BCT intersection of Analyze Conceptual Knowledge and Evaluate Conceptual Knowledge, a visualization of which can be seen in Figure 4.2.

Interviewee M

Interviewee M began her sorting of item Q1 along the Knowledge Dimension, saying,

> 1 "…I'm going back to your use of the word 'ranking' because it's rank-y
> 2 kind of, but I still am going to stick with probably it's factual
> 3 knowledge being the knowledge dimension that I think this is
> 4 targeting."

Because Interviewee M concluded that item Q1 required only factual knowledge, this excerpt was assigned the *a priori* code Factual Knowledge using directed content analysis. Interviewee M sorted item Q1 as requiring factual knowledge by following the same steps the interviewer previously took in sorting the example item and keying into a verb listed in the tool as an example, without further consideration (lines 1-3; coded as Pattern-Matching.) Interestingly, like interviewee E, Interviewee M used her instinct about the necessary cognitive process to triangulate her sorting. This excerpt (lines 1-3) was also coded as Unconfidence to reflect Interviewee M's uncertain language in sorting this item.

Moving on to sort item Q1 along the Cognitive Process Dimension, Interviewee M said,

> 5 "…Then I would probably pop it in, remember too, yeah. So I think my
> 6 thinking is what do [the students] need to know to answer this
> 7 question? And they need to know the electronegativity trend, the
> 8 periodic trend. I suppose they have to apply the periodic trend. I don't
> 9 really think of that as something that's 'apply.' So I think I'm going to
> 10 stick with remember. That still feels right to me. And my reasoning is
> 11 that what this question requires is recalling the trend. Once a student

can recall the trend, selecting an answer is pretty straight forward from there. So I think if there was something in the context that sort of like in the framing of the question that maybe pushed it just a step further than like 'which one matches the trend' that…I would maybe consider that like 'applying' the trend, but as is…I think that what it requires of the student is to recall the trend."

Because she thought item Q1 was answerable through memorization, this excerpt was deductively assigned the *a priori* code Remember. Interviewee M refuted her instinct to sort item Q1 as requiring the cognitive process of apply (lines 4-15; coded as Unconfidence) by citing both the reasoning she assumed students would use when approaching this item (lines 4-7 and 10-15; coded as Assumptions of Student Reasoning or Learning Context) and the reasoning she thought the item was trying to elicit (lines 9-15; coded as Item Forensics) as well as her previous experience with Bloom's taxonomy (lines 7-8 and 11-14; coded as Internalized Prior Experience) though she does state that the item could be considered requiring an apply cognitive process if there were something more involved about the item's presentation (lines 11-15; coded as Conditional.)

From these excerpts Interviewee M was coded as sorting item Q1 as existing at the BCT intersection of Remember Factual Knowledge, a visualization of which can be seen in Figure 4.2.

Visualizing and Comparing Interviewee Sortings

Overlaying all interviewee sortings onto the grid from the BCT provides a way to visualize the full results for each item. In this section, heatmaps for all 9 items are provided and individual interviewees' sortings are compared to one another.

Item Q1

Item Q1 was sorted by interviewees A, C, E, K, L, and M. Figure 4.2 presents a heat map of the bimodal distribution of the six BCT intersections at which the six interviewees thought item Q1 could exist. In this heat map, the sortings of interviewees A, E, M, and the dual coding of interviewee C's sorting comprise one "mode" at the interfaces of remember/understand and factual knowledge/conceptual knowledge, while the dual coded sortings of interviewees K and L comprise a second mode along the analyze/evaluate interface.

Knowledge Dimension	**Cognitive Process Dimension**					
Factual Knowledge	Remember	Understand	Apply	Analyze	Evaluate	Create
Factual Knowledge	**A E M**			**K**		
Conceptual Knowledge	**C**	**C**		**K L**	**L**	
Procedural Knowledge						
Metacognitive Knowledge						

Figure 4.2. Heat map of the item Q1 sortings of six interviewees (A, C, E, K, L, and M) where yellow cells indicate an individual sorting, orange cells indicate an overlap of 2 sortings, and red cells indicate an overlap of 3 sortings.

Item Q1 had the most bimodal sorting distribution of all nine items, and was also the item with the highest number of completely overlapped sortings (i.e., Interviewees A, E, and M all completely agreed on the sorting of item Q1.) However, the defense that these three interviewees gave for their sorting was not identical: Interviewees E and M both sorted item Q1 based on how they assumed a student would answer it, which is different from interviewee A who said he chose the sorting because of features of the item itself (i.e., its multiple choice structure.) In the other mode, interviewee K and interviewee L both sorted item Q1 as requiring the cognitive process of analyze based on the Cognitive Process Dimension header, but they honed in on different parts of the definition–Interviewee K thought the item required students to compare the electronegativities of different elements, whereas Interviewee L thought that because the students were given a set of data, that item Q1 aligned with asking them to employ something within a given context.

While the interviewees sorted item Q1 in four of the six possible cognitive processes, there was relative consensus that item Q1 would require either factual or conceptual knowledge to answer. The discrete bimodal distribution of the sortings of this item indicate that there is some determining factor as to why each interviewee would sort item Q1 in one mode or the other. One of the inductive codes assigned to the transcripts of the interviewees' sortings provides insight into this distinction. First, when an interviewee's sorting was coded as Assumptions of Student Reasoning or Learning Context, the assumed lines of student reasoning upon which the interviewees based their sorting consisted of either an "ideal" or a "minimal" path. If the interviewee assumed a student tasked with item Q1 would take a more idealized path involving the consideration of the concept of electronegativity and determinations about which of the multiple-choice distractors is correct, then the interviewee was more likely to sort item Q1 in the further right mode, whereas if the interviewee assumed students answering item Q1 would use a more minimal path requiring only the memorization of periodic trends (or just test-taking strategies, as was the case with interviewee A) then the interviewee was more likely to sort item Q1 in the further left mode. This leads to very different rankings for the same item. However, the code Assumptions of Student Reasoning or Learning Context, was also applied to interviewee assumptions of learning context. Assumptions of learning context seemed to most influence into which Knowledge Dimension category an interviewee was more likely to sort item Q1, as with interviewee K drawing parallels between factual versus conceptual knowledge and exercises versus problems.

Item Q2

Item Q2 was sorted by two interviewees: G and J. Figure 4.3 presents a heat map of the two BCT intersections at which the interviewees thought item Q2 could exist. Interviewee G and Interviewee J both thought item Q2 could exist at the BCT intersection of Remember Factual Knowledge, while Interviewee J double coded item Q2 as also existing at the BCT intersection of Apply Factual Knowledge.

Knowledge Dimension	**Cognitive Process Dimension**					
	Remember	Understand	Apply	Analyze	Evaluate	Create
Factual Knowledge	**G J**		**J**			
Conceptual Knowledge						
Procedural Knowledge						
Metacogniti ve Knowledge						

Figure 4.3. Heat map of the item Q2 sortings of two interviewees (G and J) where yellow cells indicate an individual sorting and orange cells indicate an overlap of 2 sortings.

The interviewees sorted item Q2 into two of the six possible cognitive processes, but there was consensus that item Q2 would require only factual knowledge to answer. Item Q2 did not have any entirely overlapped sortings, though Interviewees G and J did partially overlap in their sortings–both said that based on the familiarity that students usually have with the ideal gas law, that item Q2 would often only require the remembering of factual knowledge. While both Interviewees G and J thought that in most cases item Q2 would exist at the BCT intersection of Remember Factual Knowledge, they both provided conditional scenarios in which they thought item Q2 might exist at the BCT intersection of Apply Factual Knowledge. Interviewee

J thought that if a student was unfamiliar with the ideal gas law and rearranging it to find units, that item Q2 might require the cognitive process of apply. Interviewee G thought that if the value of the gas constant, R, were not provided in the question stem and that the students were required to solve for it that this would require the cognitive process of apply. As the question is written, only Interviewee J's scenario is possible, which is why only interviewee J was double coded at both BCT intersections Remember Factual Knowledge and Apply Factual Knowledge.

Item Q3

Item Q3 was sorted by interviewees A F, J, K, and M. Figure 4.4 presents a heat map of the six BCT intersections at which the five interviewees thought item Q3 could exist. Item Q3 elicited interviewee sortings from the cognitive processes of understand to evaluate and from factual knowledge to procedural knowledge on the knowledge dimension. These sortings do not have a central point of consensus, instead interviewees F, K, and M sorted item Q3 at separate points in the range, and interviewees A and J chose three adjacent intersections at the cognitive process evaluate and the conceptual knowledge categories.

Knowledge Dimension	Cognitive Process Dimension					
	Remember	Understand	Apply	Analyze	Evaluate	Create
Factual Knowledge					**A**	
Conceptual Knowledge		**J M**	**J**	**J F**	**A**	
Procedural Knowledge					**A K**	
Metacogniti ve Knowledge						

Figure 4.4. Heat map of the item Q3 sortings of five interviewees (A, F, J, K, and M) where yellow cells indicate an individual sorting and orange cells indicate an overlap of 2 sortings.

The interviewees sorted item Q3 in four of the six possible cognitive processes, and three of the four possible knowledge categories. However, three of the six intersections proposed by the interviewees were chosen by more than one interviewee, and four of the five interviewees sorted item Q3 as requiring conceptual knowledge, suggesting some consensus on the sorting of item Q3. Interviewee J chose conceptual knowledge because the students are trying to relate pieces of knowledge to one another, whereas Interviewee M chose conceptual knowledge based on pattern-matching item Q3 with its Cognitive Process Dimension sorting, and interviewee F based her sorting on assumptions of unfamiliarity with the material. Interviewees A and K both chose different knowledge categories, Interviewee A choosing factual knowledge and procedural knowledge because he uses a knowledge-in-pieces model and didn't know which piece of knowledge students might surface first, and interviewee K choosing procedural knowledge because he believed students would have had to go through the empirical formula process in order to answer this item.

Item Q4

Item Q4 was only sorted by interviewee H. Figure 4.5 presents a heat map of the distribution of the four BCT intersections at which interviewee H thought item Q4 could exist between the apply/analyze and conceptual knowledge/procedural knowledge interfaces.

Knowledge Dimension	**Cognitive Process Dimension**					
	Remember	Understand	Apply	Analyze	Evaluate	Create
Factual Knowledge						
Conceptual Knowledge			**H**	**H**		
Procedural Knowledge			**H**	**H**		
Metacogniti ve Knowledge						

Figure 4.5. Heat map of the item Q4 sortings of one interviewee (H).

Item Q5

Item Q5 was sorted by two interviewees: A and B. Figure 4.6 presents a heat map of the two BCT intersections at which the two interviewees thought item Q5 could exist.

Knowledge Dimension	Cognitive Process Dimension					
	Remember	Understand	Apply	Analyze	Evaluate	Create
Factual Knowledge	**A**					
Conceptual Knowledge				**B**		
Procedural Knowledge						
Metacogniti ve Knowledge						

Figure 4.6. Heat map of the item Q5 sortings of two interviewees (A and B).

Item Q5 did not have any overlapped sortings–Interviewee A and Interviewee B chose very different sorting for this item. Interviewee A thought that the item “masquerades potentially as something, a lot more higher-order than maybe what students might do to give a response,” and that students would answer the item based on previous factual knowledge about oxygen that they brought into the learning environment, whereas Interviewee B thought that the item was giving students relative data and asking them to make conclusions based on their understanding of how oxygen interacts with water, which would be conceptual knowledge.

Item Q6

Item Q6 was sorted by 8 interviewees: interviewees B, C, D, E, G, H, I, and J. Figure 4.7 presents a heat map of the ten BCT intersections at which the eight interviewees thought item Q6 could exist. Item Q6 elicited interviewee sortings across all cognitive processes and most knowledge categories, with the exception of metacognitive knowledge.

Knowledge Dimension	**Cognitive Process Dimension**					
	Remember	Understand	Apply	Analyze	Evaluate	Create
Factual Knowledge	**G**		**J**			**J**
Conceptual Knowledge			**B C J**			**I J**
Procedural Knowledge		**E**	C E G H J	**D**	**G**	**D J**
Metacogniti ve Knowledge						

Figure 4.7. Heat map of the item Q6 sortings of eight interviewees (B, C, D, E, G, H, I, J) where yellow cells indicate an individual sorting, orange cells indicate an overlap of 2 sortings, red cells indicate an overlap of 3 sortings, and oxblood cells indicate an overlap of 5 sortings.

While the sortings for item Q6 were spread across much of the BCT, there was evidence of convergence among the sortings. The intersection with the most agreement amongst interviewees is Apply Procedural Knowledge; five out of eight interviewees sorted item Q6 into at least this intersection, and all eight interviewees thought the item required at least the cognitive processes of apply or create–based on whether they thought the students were applying a separation method they had learned or whether they were having to create a method based on what they knew about the properties of each mixture. Interviewees D, E, and G all gravitated toward other cognitive process dimension categories based on this same distinction–how familiar the students were with the materials of the item and how much they had to develop their own sorting techniques. Interestingly, interviewees D and E both reported being more comfortable with their more leftward cognitive process sorting.

Half of the BCT intersections at which the eight interviewees thought that item Q6 could exist were at the procedural knowledge level; while seventy percent of the BCT intersections were at either the conceptual knowledge or procedural knowledge levels. Interviewees C, B, I, and J all based their selection of Conceptual Knowledge on what students would need to choose between separation methods, while Interviewees C and J also included Procedural Knowledge in their sorting based on what students would need to know about each separation technique. Interviewees D, E and

H all honed in on the methodological nature of the item to choose procedural knowledge.

Nine of the ten BCT intersections at which the eight interviewees thought item Q6 could exist are adjacent to one another; only one sorting is physically removed from the group: Interviewee G thought that the separation of a solid iron and sulfur mixture was common enough an exercise in general chemistry that it would exist at the Remember Factual Knowledge intersection. Interviewee G was the only participant who explicitly sorted each of the three tasks of item Q6 at separate intersections, and Interviewees G and J were the only interviewees to sort item Q6 as requiring factual knowledge. Interviewee G suggested that knowledge dimension level is a function of the students' familiarity with the material, and that as knowledge of each method becomes less novel, it transitions from being procedural to being more factual in nature, while interviewee J said she could see arguments for Factual Knowledge, Conceptual Knowledge, or Procedural Knowledge, but based her placement on Factual Knowledge on students needing to know the components of each mixture.

Item Q7

Item Q7 was sorted by 6 interviewees: interviewees C, E, L, M, N, and O. Figure 4.8 presents a heat map of the five BCT intersections at which the six interviewees thought item Q7 could exist. Item Q7 elicited interviewee sortings across four of the six cognitive processes and in two of the knowledge dimension categories. Of all items presented to interviewees, this heat map has the most resemblance to a more classic bell curve; there is a clear point of centrality at the apply/analyze interface, with less agreement in the adjacent intersections.

Knowledge Dimension	**Cognitive Process Dimension**					
	Remember	Understand	Apply	Analyze	Evaluate	Create
Factual Knowledge						
Conceptual Knowledge		**C**	**M N O**	**C M N**	**L M**	
Procedural Knowledge			**E**			
Metacogniti ve Knowledge						

Figure 4.8. Heat map of the item Q7 sortings of six interviewees (C, E, L, M, N, and O) where yellow cells indicate an individual sorting, orange cells indicate an overlap of 2 sortings, and red cells indicate an overlap of 3 sortings.

The intersections with the most agreement amongst interviewees are Apply Conceptual Knowledge and Analyze Conceptual Knowledge; four out of the six interviewees sorted item Q7 into at least one of these two intersections. Interviewees C, M, and N all were unable to specify item Q7 to one intersection, instead each overlapping with one another in their double coding of the item. Interviewees M, N, and O all gravitated towards the cognitive process of apply based on familiarity with the material, but while interviewees M and N said that the item would require the cognitive process of analyze (or even evaluate in the case of interviewee M) if the material were novel, interviewee O thought that the cognitive process of apply was appropriate for an unfamiliar context.

As opposed to differentiating the cognitive process categories by level of familiarity, interviewees C and L debated between two non-adjacent cognitive processes based on the provided Cognitive Process Dimension headings–interviewee C could not narrow her sorting between the cognitive processes of understand and analyze, while interviewee L eventually removed the cognitive process of apply from consideration and settled on just the cognitive process of evaluate. Interviewee C thought that item Q7 required students to describe and explain, which corresponded to the cognitive process of understand, but it also required them to compare and contrast, which corresponded to the cognitive process of analyze, while not requiring them to calculate or execute anything, which precluded the cognitive process of apply. On the other hand, interviewee L thought that item Q7 did not require students to break information into parts, which corresponded to the cognitive process of analyze, it did ask them to predict in which container dissolution was the fastest, which could correspond to the cognitive process of apply, but because that discrimination was based on a set of criteria and standards, item Q7 ultimately required the cognitive process of evaluate.

While the interviewees sorted item Q7 in four of the six possible cognitive processes, there was relative consensus that item Q7 would require either conceptual or procedural knowledge to answer. Interviewee E thought that students would need to engage in a procedural process to answer this particular question, whereas the other five interviewees gravitated towards conceptual knowledge because students would need to think about the conceptual relationship between surface area and solubility.

Item Q8

Item Q8 was sorted by four interviewees: interviewees H, J, M, and N. Figure 4.9 presents a heat map of the distribution of the five BCT intersections at which the four interviewees thought item Q8 could exist. The shape of this heatmap is slightly bimodal; the sortings of interviewees H and N comprise one mode centered at the interfaces of remember/understand and factual knowledge, while the sortings of interviewees J and M comprise a second mode along the apply/analyze and conceptual knowledge interface.

Knowledge Dimension	Cognitive Process Dimension					
	Remember	Understand	Apply	Analyze	Evaluate	Create
Factual Knowledge	**H N**	**H N**	**N**			
Conceptual Knowledge			**J**	**J M**		
Procedural Knowledge						
Metacogniti ve Knowledge						

Figure 4.9. Heat map of the item Q8 sortings of four interviewees (H, J, M, and N) where yellow cells indicate an individual sorting and orange cells indicate an overlap of 2 sortings.

While the interviewees sorted item Q8 in four of the six possible cognitive processes, there was relative consensus that item Q1 would require either factual or conceptual knowledge to answer. In the first mode of the heat map, Interviewees J and M narrowed in on the classic difficulty of buffer problems for students, therefore promoting the item along both axes– Conceptual Knowledge because students have to relate pieces of knowledge to each other, and Analyze based on the language of the Cognitive Process Dimension header (though interviewee J notes that based on the familiarity the student has with the material might put it in the Apply category.) In the second mode, interviewee H and interviewee N thought that that item Q9 would not elicit more than factual knowledge: Interviewee H based this conclusion on the multiple choice nature of the question, saying this required less cognitive engagement, while interviewee N thought that familiarity with the subject matter would eventually lead to less cognitive engagement.

Item Q9

Item Q9 was sorted by interviewees A,F, K, and N. Figure 4.10 presents a heat map of the five BCT intersections at which these four interviewees thought item Q9 could exist. The heat map possesses one point of consensus for all four interviewees, and then secondary sortings of three of the interviewees spread out across four of the six possible cognitive processes and three of the four Knowledge Dimension categories. Of all the items presented to the interviewees, this is the only item in which all interviewees at least partially agreed on the same intersection.

Knowledge Dimension	Cognitive Process Dimension					
	Remember	Understand	Apply	Analyze	Evaluate	Create
Factual Knowledge	**A**					
Conceptual Knowledge		**A**				**K**
Procedural Knowledge			**F**			**A F N**
Metacogniti ve Knowledge						

Figure 4.10. Heat map of the item Q9 sortings of four interviewees (A, F, K, and N) where yellow cells indicate an individual sorting and red cells indicate an overlap of 3 sortings.

All four interviewees thought the item could require the cognitive process of Create, while three of the four interviewees thought item Q9 could exist at the BCT intersection of Create Procedural Knowledge, though Interviewee F thought that the intersection of Create Procedural Knowledge was the less likely of her dual coded intersections and that it would most likely exist at the intersection of Apply Procedural Knowledge. Interviewee A had the widest spread of sortings; he sorted item Q9 as also existing at the BCT intersections of Remember Factual Knowledge, Understand Conceptual Knowledge, and Apply Procedural Knowledge. Interviewee A was the only participant to say that item Q9 could explicitly elicit Factual Knowledge depending on how familiar the student was with the material, however, both he and Interviewee K used their cognitive process dimension sorting of Q9 to triangulate a Knowledge Dimension sorting of Conceptual Knowledge. Interviewee K said item Q9 would only require procedural knowledge if the students had to perform some calculation with their collected data, whereas the other three interviewees agreed on the Knowledge Dimension category of Procedural Knowledge based on the experimental nature of the item.

Themes from the rest of the interviews

As noted in the Methodology chapter, the tool-testing interviews consisted of three distinct-but-related sections. The rest of this chapter will discuss the findings from the first and third of those sections.

Interviewee profiles

In the first section, interviewees were asked about their experience with and opinion on learning taxonomies. For many interviewees, Bloom's Taxonomy was the first learning taxonomy with which they became familiar, but the way they integrated it into their teaching or research practice varied. While analyzing this section, it became apparent that some interviewees seemed to occupy a few archetypes of how they interacted with learning taxonomies. Refining of these archetypes resulted in four Interviewee Profiles, summarized in Table 4.4. In the following sections, characteristics of each profile will be summarized, and example quotes from two interviewees that exemplified that profile will be provided.

Table 4.4. The four identified interviewee profiles.

Profile	**Definition**	**Example Interviewees**
Bloom's Intensivist	Interviewee tends to use a bottom-up approach to taxonomy praxis. Bloom's taxonomy tends to drive the design of their teaching (through informing how they write assessment items and learning goals.) Often uses Bloom's as a tool to train TAs and other teachers.	Interviewee B Interviewee J
Taxonomical After-Thoughtist	Interviewee tends to use a top-down approach to taxonomy praxis. They tend to design their teaching first and then apply a learning taxonomy to it, typically as a tool to communicate something about the teaching to others.	Interviewee C Interviewee H
Prefers-Another-Taxonomy	Interviewee knows Bloom's taxonomy, but likes a different formal taxonomy or informal dynamic better.	Interviewee I Interviewee K
Taxonomy Agnostic	Interviewee has no attachment to a particular taxonomy, instead drifting between frameworks as needs arise.	Interviewee E Interviewee G

Profile 1: Bloom's Intensivist

The first interviewee profile is that of a Bloom's Intensivist. This is someone who uses Bloom's Taxonomy to drive the design of their teaching, research, or both. Interviewees B and J were both good examples of Bloom's Intensivists. Both interviewees discussed the ubiquity of Bloom's taxonomy in their early teaching career, with interviewee B saying,

> "[Bloom's Taxonomy] was definitely, always kind of there with me, as I was thinking about assessment items and learning goals. Mostly I say it dictates, like, do I really want this to be a remember or understand learning goal? Or do I really mean something different? Like, should, should I be writing this at a different level? That's kind of how I use Bloom's taxonomy as a way to reflect on what is it that I really want students to know and be able to do. And then, okay, now I can go and say, I really think this should be an application question. So what, what should this question look like? And Bloom's taxonomy kind of helps you frame both the learning objective and the assessment item itself kind of in parallel. Cause I, I often develop them together. Like I write the, I write the objective and I'm like, okay, can I visualize, what is the assessment going to look like? I may not know the context or the system or anything, but I have a general sense of what the stem of the question is going to look like based on that."

Interviewee J also relied heavily on Bloom's taxonomy as a new faculty, which resulted in science of teaching and learning data she wished to share in publication.

> "I used Bloom's taxonomy a lot to think about my test questions in particular and sort of what level, what, what kinds of skills I'm asking from students…So I, um, you know, where I, where I decided to kind of start with the exam blueprinting process was in that I was, I was trying to do two things. Um, I was trying to better communicate with my students, what the expectations were, um, in order to reconcile that question of like 'the exam was nothing like what I expected' kind of thing. So managing student expectations, uh, which was very important to me as a pre-tenure faculty member at the time. And then, um, the second piece was that I was trying to, uh, systematize my own work. Writing exams takes a long time for me as I'm sort of writing questions and thinking about kind of making a complete exam or how the pieces fit together. And so for me, the idea of making an exam blueprint, I thought about my learning goals and then correlated that with the level using Bloom's taxonomy that I was expecting my students to be able to perform. Um, that was, that was an organizing principle for me. Um, that was also, for me, that was also kind of streamlining my workflow because I, part of that was, I had just had a baby. And so I was learning and at a personal level, I was trying to

> figure out how to, how to balance, um, and integrate my work and my life."

Though they may have been introduced to Bloom's by a mentor or supervisor, Bloom's intensivists often took the time to pursue more readings on the versions of Bloom's Taxonomy. Interviewee B notes that she didn't have much choice of what taxonomy she was going to use for her research: "I think, and I guess I never really spent much time debating, like, should we use Bloom's?...I just, because my postdoc-advisor used Bloom's, that's what we went with," but even so, on her own years later she still read volumes one and two of the original taxonomy as well as the revised taxonomy. Interviewee J first learned about the original taxonomy in elementary school, but was reintroduced to it via an Ionic Viper Community workshop in her first year of teaching, and then read about the Blooming Biology Tool.

Many times, Bloom's Intensivists resonate with the structure or rationale of Bloom's taxonomy enough that they train colleagues or mentees in the use of it. Interviewee B uses it for mentoring new instructors, noting,

> "It seems for faculty, once I moved into that stage of working more with faculty, it seemed very approachable and logical for faculty to think about Bloom's taxonomy. And so I just kind of stuck with that and didn't move into using any other taxonomy…When I bring in new…graduate students who teach alongside of me. So what we have is we don't have a formal name for this program, but we have multiple sections of intro bio and a collaborating instructor, whether they're a graduate student, a post-bac, or faculty member will teach their own section in parallel with me and I teach them about Bloom's taxonomy, just so that they can think and get this exposure to, Oh yeah, I'm asking this question. And this is a, this is a, you know, an understand

> or a factual recall sort of a thing, but I don't really mean it to be that, you know, so to kind of introduce them. So that's kind of like a mentoring thing that I do with them in terms of research."

Interviewee J agrees that, she personally is interested in "if we teach students about learning taxonomies in general and Bloom's in particular, um, can that help students better understand what's expected of them?" Interviewee B noted that she hasn't used Bloom's with students directly, but that, "as I'm moving more into thinking about the role of feedback and self-evaluation, I could see that coming back up…to practice and prepare for the next exam."

Over their career, Bloom's intensivists may reference Bloom's Taxonomy less often, but they note that they still use it internally. Interviewee B says, "now I think it's probably still there, but I think it's a little bit more, I'm more fluent with it, right. So it's just part of, that's just how I go about doing uh doing course development and assessment, item development, sorts of things." She also mentioned that her department is using it for holistic programmatic assessment.

Profile 2: Taxonomical After-Thoughtist

The second interviewee profile is that of the Taxonomical After-thoughtist. This is someone who uses a learning taxonomy in a top-down fashion to fill a gap in their teaching, research, or both. Interviewees C and H were both examples of this profile. Members of this profile often did not specifically seek out a learning taxonomy, but rather some way to communicate some aspect of their teaching/research (e.g., to communicate to students what they are learning, or to communicate something of the nature of the learning environment to manuscript readers.) For example, interviewee C uses Bloom's with her undergraduate mentees as a way to discuss metacognition and study skills, but when she needed some assistance in data analysis in one of her research projects, Marzano's taxonomy was suggested by a colleague, so she used it in conjunction with the Interactive-Constructive-Active-Passive framework to perform discourse analysis on group quizzes. Before consulting her colleague, she did not have intentions of sorting the quiz items by cognitive level, saying, "I had no idea at that point, uh, we had the data and then I said, 'okay, this is what we have. How can we use this to, to make, how can we make sense of this?' And she, and when, while we was speaking about it and going through it, this was at a Gordon conference, and she was like, 'Hey, you know, have you considered, you know, Manzano?'" (Interestingly, Interviewee C didn't know it at the time that she selected the data analysis method, but the group quizzes she was analyzing were written using Bloom's taxonomy.)

Similarly, when interviewee H engaged in course reform for an analytical program using problem-based learning, he only engaged Bloom's taxonomy in the writing process to communicate the different types of potential learning environments. Interviewee H learned about Bloom’s taxonomy early in his teaching career, and returned to it to communicate with other educators about his teaching experiences. Since that project, interviewee H has used Bloom's taxonomy to explain his teaching rationale to his general chemistry students, saying,

> “I've been looking at it more in terms of my lower-division, my uh general chemistry course and introducing the students to that and sort of to explain why I approach things the way I do um using Bloom's. And I don't know how much, that it's early in the semester, and I don't know how much they absorb. It makes me feel better, if nothing else that I'm explaining why I'm doing what I'm doing and it has some literature base for it.”

Interviewee H has also developed integrated problem activities for the general chemistry course, where students answer questions across the cognitive process dimension of Bloom’s taxonomy throughout both semesters. Interviewee H explains, saying “So in the second semester they might also be reviewing some thermodynamic information based on that single topic, all…to sort of get them to build a, synthesize that material by brute force, I would say.”

In this way, interviewee H started using Bloom’s taxonomy as a communication tool, but eventually it became useful in his development of class material, while interviewee C does not use her taxonomies of choice to develop her teaching or research, but rather solely as a way to scaffold her discussion of it to students or colleagues.

Profile 3: Prefers-Another-Taxonomy

The third interviewee profile is someone who, based on their interest in measurement or teaching, Prefers-Another-Taxonomy besides Bloom's. Interviewees I and K were both good examples of this profile. Each interviewee had their own preferred dynamic–Interviewee I was attached to the 3D-Learning Assessment Protocol while Interviewee K had experience using the Expanded Framework for Assessing General Chemistry Exams (EF).

Interviewee I focused item sorting and writing on what "intellectual work is being emphasized," saying,

> "I think on our organic course right now, every single exam we're going to give this coming semester, cause we we've written all of them, which I'm pretty proud of as a, as a team. And they're all at least 50% of the points dedicated to 3D tasks. So where you've got to use a big idea of a scientific practices frame by a crosscutting lab. That's awesome. I'm pretty happy with that. I think that says, we care about that sort of work. If you look at assessments and another class that gives the same intro lecture and they're all, you know, calculate, draw, whatever that's going to be pretty much no points on 3D learning. And that tells you something from a research perspective, I get sort of irritated when people just talk about learning outcomes in very vague terms, especially those outcomes are either exam performance or course grade, because I have no idea what that means and I'm pretty sure it means nothing. So it would be useful if we had some sense what sort of intellectual work was emphasized on the assessments and something like the 3D-LAP can do that."

While Interviewee K initially used the EF to attempt to predict students' performance on exam items, in light of the COVID-19 pandemic, he now uses it to align his learning outcomes with his assessments, saying,

> “Specifically for the general chemistry course, then I went back to [the Expanded Framework.] And so this past summer I went back to that and aligned everything and use them as the objectives and, and it worked out very well…It's funny what, what I did think at the time, you know, like when I was doing it this past summer is I, that I should have been doing this all along. That was my main thought. And I thought, cause I couldn't figure out I was like, “why, why was I not doing this?” You know? And so, so I was really happy to actually go back to it and have it as an anchor, you know, and have it as something else that that would be beneficial for everybody, you know, for the students, for me, for everybody involved. And so my main thought was “I should have been doing this” and I, I just couldn't figure out why I hadn't.

Members of this profile may or may not have been very familiar with Bloom’s taxonomy. Of his familiarity, interviewee I said, “Somewhat familiar. I mean, I must often use sort of frameworks aligned with 3D learning, which supposes less of a hierarchy than Bloom's taxonomy.” Interviewee K said,

> “I haven't looked at it too much recently, but of course I did, you know, like a good amount of work with it. And so I would say that I'm pretty familiar. I think I have a good familiarity with it. And, um, and as I start to get more into it, you know, things are going to start to come flooding back. But back then when I was doing the work, I was totally knee deep in it. And so I was super familiar with it at the time. Um, and so I'm still pretty familiar with it.”

When defending their preferred dynamic, members of this profile often echoed the criticisms of the original taxonomy–particularly about the “pyramid” structure of and the lack of a theoretical basis behind the categories. For example, Interviewee I elucidated on his preference for 3D-LAP saying, “There's no presumption that one [3D-LAP Practice] is more or less sophisticated than any other. And that I believe differs from how Bloom’s is typically conceptualized, because you often see the sort of pyramid." He also discussed the ambiguity of education research community terms, saying,

> "I think we need to really be careful about what we think matters and how it's parameterized. And I expect this is something that you've run into, there is a long and distinguished history in chemistry education and in Ed in general of saying really fun-sounding buzzwords that mean jack-squat. Nobody knows what 'critical thinking' means. Nobody knows what 'inquiry' means…And in like all cases, there is no discussion of what the heck could the outcomes mean. It's just like course assessments or course grades. And nobody has looked at, I mean, frankly, I always think just frickin' 3D-LAP, your exam. I mean, it's, it's not that hard. You could just do it. And it would tell me something right now. I know nothing. It could just be an art class. I mean, not that there's anything wrong with art classes, but I don't think you can say much about the importance of this helping you be a doctor, if all you're doing is getting really good at drawing hexagons."

Interviewee K wanted learning taxonomies to be useful for predicting student performance on exams, saying,

> "I was doing a lot of work looking at some of the exams, you know, like looking at exam results, like the ACS exams. And then I was thinking about, well, how can these be sorted? Can you predict students' performances? You know, and can you, can you actually try to develop a framework that says, okay, this will let you know how students are going to perform in this area. This might predict how students are gonna perform in that area. And so we did a lot of looking through the exams, trying to see what common items were with the themes or in everything in the exams. Um, then we tried to play the framework to sort of predicting how students would do on exams, but that was a little bit tougher to do. And so coming up with the framework and sort of using it as a tool to say, okay, well, you might want to have these types of items represented on your exams. And these are common types on other exams that work pretty well, but predicting students' performance was a little bit harder."

Profile 4: Taxonomy Agnostic

The last interviewee profile is the Taxonomy Agnostic–someone who does not have particular attachment to any taxonomy in particular, and will switch between them as needs arise. Interviewees in this profile often encountered Bloom's taxonomy first, and therefore used it in research because it was the only framework they knew. The choice in learning taxonomy was not particularly intentional and might be based on convenience. Interviewees E and G were both good examples of this profile–both interviewees came across Bloom's taxonomy at academic conferences, and were attracted to it because of its accessibility. From there, they both elected to use it for research projects early in their teaching careers, but have since drifted away from it in their own research, teaching, or both.

Interviewee E learned about Bloom's taxonomy at a conference early in her post-doc and first considered using Bloom's taxonomy for her teaching, based on its tangibility, saying,

> "I was like really excited cause I can grab on to it. You know, it made intuitive sense to me. Um, and, and all of that. Um, so I'm pretty sure I heard it at a conference and you know, probably made some notes about it as, 'Oh, this might be useful later.'…I'm almost completely sure I heard about it at a conference first…at that time especially, I would hear something at a conference and then I'd go immediately back and like, think about 'how does this apply to my teaching?' or something like that."

Later on, when Interviewee E found herself needing a learning taxonomy for a project in her early teaching career, she already had one in reserve and was able to apply it to her research, saying.

> "In some sense it was, it's because it's so widely known and it was like what I knew about, right? So I don't know if at that time, I was like, I wasn't necessarily saying, um, you know, 'here are other frameworks that we could think about helping to guide this.' I was like, if, you know, that's the one thing that I knew, one thing that I knew of, and I felt comfortable with, um, was Bloom's taxonomy and that, that I felt like I could interpret it sufficiently so as to guide the construction of our questions. Um, but, but I admittedly like, um, because I was a postdoc who was trying to guide a graduate student from another group and doing a project, um, I don't think either one of us was spending a lot of time, you know, in the literature going, 'what other alternatives, why?' You know, so even if I think if I went back to the paper, I doubt there's much of a rationale in terms of 'here's why Bloom's was useful for this' versus something else that you might consider. I'm not sure we even got that deep with it. We were just like, 'this feels, this is something I know about. Let's look at it and it feels like a good fit.'"

Likewise, Interviewee G learned about Bloom's taxonomy from a colleague at a conference while discussing how to communicate his research findings, and was attracted to it because of what he considered to be overlap with his background in physical chemistry, saying,

> "So, uh, having read, eh, Uri Zoller's, papers about HOCS or LOCS, which being myself, a quantum chemist at that time, uh, I very like very much uh this HOMO/LUMO distinctions in quantum chemistry. So associated with that, that, that was one way of attracting me. So that's the, the interaction between, uh, my science and the scientific part and the educational part. So it was a conference in Rome when I discussed, uh, with Uri Zoller about my data and we decided to try to explain them in terms of his HOCS and LOCS approach."

Both interviewees have used several frameworks or taxonomies in their research careers, though interviewee E is clear that none of the other dynamics she used were called taxonomies, saying,

> "Uh, I mean, not things that are called taxonomies. There are frameworks that we have used, um, like Russ's framework on mechanistic reasoning, um, other frameworks that I've used in the, in the, in like research. And I think about that as being, not unlike that in a sense it's just more specific to a different kind of thinking and, and you can also, it also kind of has a sense of difficulty in it. There's a framework for, um, kind of scientific thinking, what is it called…that we used in, uh, in the writing? So, so I guess these come out more of our writing-to-learn projects. Um, whereas our, we really just stuck with PCK theory rather than a taxonomy for, um, the PCK research? So even, even in our, you know, um, we've stopped constructing the same kind of question, so it hasn't come up there again, but the, but we definitely tend to use more frameworks in an analytical way in, um, in the writing-to-learn research that we do. Um, one of the more popular ones has been, or at least starting off, we were looking at at Russ's mechanistic reasoning framework. There's a scientific reasoning, um, framework that, um, by Greenberg and Hand that we had used, um, for one study in particular. And then we're kind of like seeing where else we go with that. Um, so I think we really like frameworks, but, but they've kind of moved cl--drifted closer to like what I think of science practices and NDSS, and, and thinking that might kind of coincide with that or, or, and, or be chemistry-specific."

Likewise, Interviewee G discussed basing research throughout his career on Piagetian as well as Balian theory and using the SOLO taxonomy, but focuses on the similarities across the various frameworks, saying,

> "There are many things which are related to these taxonomies, but I find that more or less, they they're similar. They are different in a number of things, but they are are similar in many other things…I mean, my own approach is that uh one has to use all, all theories, which are available. Like we are doing the science in physics. I mean, uh, quantum mechanics is very important today for physics, but we don't forget classical physics. We don't forget Newton. So everybody's there. And the same applies in chemistry. I mean, when we don't forget, uh, Arrhenius and, uh, uh, all these guys."

Both interviewees attribute this approach to having varied research interests. Interviewee G said,

> "An issue here is that, uh, uh, in my research in educational research, I have quite a spread of topics. So I have not, uh, concentrated on a one or two things, but I'm doing a a range of things. And this was a result of, uh, of the field being a new one. So you want to explore it even for yourself as widely as, uh, as you can. It also, it's a field where you, don't only interested in theory, but you also want to be useful in the applications…Uh, but, uh, uh, this is the issue that, uh, I was interested in many things while if I had gone on with doing quantum chemistry, my field could be, you know, more narrow."

Interviewee E echoes this, saying,

> "I mean, I just say that, like, if you asked me, I'm not sure that I used it in anything other than this paper or related or the related papers, right. There are a couple that came out around the same time or within a year or two all on PCK. Those are the only ones that I could think of using, because we don't really try to, I think we've moved away from trying to design like items or measures for things we tend to do more interview based research in which case we're not really doing that, not completely. Um, but in that as we've we I tried to move away from that. And then as we're kind of returning to some more instrument measure type things, we tend to, um, we've been doing it in more specific ways. Like we've been thinking about mechanisms in organic chemistry and that those have very specialized frameworks. So this to me, um, it made sense for that project. There just hasn't been another project or movement in a project where it made sense to that for that."

While Interviewee E didn't find much more use for Bloom's taxonomy in her research, early in her teaching, Interviewee E relied more on Bloom's, saying,

> "At some point early on, when I began teaching this course, when I had very little teaching experience, [Bloom's taxonomy] was like this anchor that I could hold on to, um, to the point that…I even ordered a, a Bloom's book. Probably after that conference, I think it was probably a Gordon conference that I heard it at, or something, I don't remember where, but I came back and ordered a book. So I have this like Bloom's taxonomy book in my office."

However, as she has progressed more in her teaching, she reflected that she doesn’t use Bloom’s taxonomy as much anymore, saying,

> “And so like, if I were in my office right now, I'd probably would open it up and, um, and think, yeah, when's the last time, like, it's probably been a while since I've actually sat down and like developed activities or questions or things for students, like specifically going, how does line up with Bloom's? Like, I, I'm probably much farther detached from it than when I first saw it and I would be writing a course. I'd be like opening it up and thinking about it. And now I don't really do that anymore. And I, and in part maybe, maybe in some sense, my teaching has also gone more specific-y in organic as opposed to, um, this was, was just this really general thing that, that I could hold onto. Um, and, and now I think about it a little differently.”

And Interviewee E feels her current teaching load and philosophies have contributed to her drift away from Bloom’s taxonomy, saying,

> “So I teach organic two lab course the most. That's the introductory course that I teach, um, where I'm really thinking about this. Um, I, I teach a graduate course in chem ed and, um, we've, you know, people have brought up Bloom's. So I know some of my students know about Bloom's, but I haven't like used it in that class much. I really use it more in this introductory and not deliberately like opening up Bloom's, as I said…I loosely just go, um, you know, what am I asking them to do here?...Yeah, I think, I don't even think, like, it's not like I have Bloom's open and I'm going as I write this question, you know, ‘which Bloom's category does it go into?’ I think once, once I had especially applied it to my research and thought about it now, when I write practice questions or test questions, um, it's more of a sense of, um, it's, it's more of a way of thinking than it is like specifically applying, you know, this, this recognition that, um, there are different kinds of knowledge and that there are different levels of difficulty that might be associated with that, or different kinds of thinking patterns that, or problem solving type things that you want students to engage in…And, and I guess, because I teach lab, one of the things that I've drifted towards, which to some sense, I think actually maps onto Bloom's a little bit is, um, the science practices and NGSS. So, so like this has analyze like it's lab, they analyze a lot of data. Of course that's

> analyzing data, something very specific, but, um, you know, I can see some overlaps there and thinking about thinking about it more in, in that sense, I guess. Um, so I guess I just, you know, I'm kind of an experimenter in my teaching and so I'll use something for awhile. And then I know it has this like, lasting impact on my t- thinking even if I don't like explicitly go back and use that and go, 'now I'm going to unplug, you know, open up the Bloom's taxonomy in my brain and apply it to this question.' You know, like I know it, it influences the sort of sense that I have of 'what kind of question am I writing?' Um, so, so I guess as I'm saying that I kind of re-turned around and think like, I really was using it mostly for quiz questions and thinking about that. And, and I'm kind of at the point where in this class, like, we're going to just not even have quizzes this semester too, so there's just been so much drift away from it."

Likewise, Interviewee G says that he uses Bloom's taxonomy in his teaching of science education courses, but that his research has focused on problem solving and while "there was some visiting of the HOCS/LOCS matter" he has not continued to do systematic work with Bloom's Taxonomy.

Debrief of BCT

In the third section of the tool-testing interviews, interviewees were asked if they had any difficulties using the BCT and if they had any suggestions for its improvement. For analysis, these questions were combined with the final part of the first section of the interviews, where the participants were asked what value they saw for a tool like the BCT and how they saw it fitting in the greater CER landscape. The responses to these two lines of inquiry have been separated into the two following sections, Proposed Use of the BCT in the CER Community, and Proposed Improvements to the BCT.

Proposed Use of the BCT in the CER Community

In the tool testing interviews, participants were asked if they saw a need for a tool like the BCT in the chemistry education research community, and what uses they saw in that landscape for such a tool. As interviewee G pointed out, arguing for discipline-specific tools, chemistry occupies a unique position in STEM because of the necessity of learning both calculative and conceptual skills,

> “Chemistry is very unique in that if you take physics, for instance, any area of physics–be that mechanics, be electricity, magnetism, quantum mechanics, anything, it has uh mathematics involved in it, but chemistry can be quite different, you know, cause both the quantitative and qualitative aspects. And so to be able to categorize the difficulty, the complexity of problems and questions, it's very important. And so chemistry, because it's unique in that, it needs its own, uh, instruments.”

Interviewee suggestions for uses of a discipline-specific item sorting tool such as the BCT spanned all components of the CER community–there were proposed uses for researchers, practitioners, and students, though there was the most consensus that due to its accessibility and ubiquity, that a discipline-specific approach to Bloom’s taxonomy had the most potential as a tool for instructor training. In the following sections, the proposed uses of the BCT within each realm of the CER landscape will be summarized, and example quotes from the interviewees will be provided.

For Researchers

Following the original intention of Benjamin Bloom, in the realm of research, interviewees found the most use for the BCT as a way to communicate and standardize their findings for other community members. Interviewee J discussed how consistent use of such a tool could remove the burden for each CER researcher to build the language to communicate their findings similarly to how the Blooming Biology Tool has affected the Biology Education Research community, saying,

> "I could see that having a fairly consistent tool might help us to compare results…I do remember when I learned about the Blooming Biology [Tool,] it was, I did have the thought of like, 'Oh, it'd be kind of cool if there was one for chemistry' and I looked and there wasn't, and that must you know, um, and to have something that we could use to compare, so we don't have basically, so that you don't have to do what I did, which is reinvent, reinvent the wheel, right…I think it would help the research community because you could then you could study something else instead of making a new instrument every time."

Interviewee E and H both discussed communicating their findings with their practitioner colleagues, and how common language could assist in that, with interviewee E saying,

> "I grapple with that all the time, um, in different ways or with colleagues, when I'm trying to explain to them what we're doing in various projects and things like that, um, you know, like 'what does modeling mean,' or synthesis of course is, uh, we, you know, writers, people in the writing world, they talk about synthesis all the time: there's synthesis of ideas in writing. And I know that has a very specific meaning in chemistry. And so, um, it's like some code-

> switching that you have to do in your mind when you're going between the two sets of language."

Similarly, Interviewee H said,

> "If I'm talking to colleagues who don't have that mindset…I don't know if all of my colleagues think about it in the same way I do. And it would be, it would be nice to have a set of terminology that would be consistent and specific that we could, I could use to be more persuasive for my arguments. So I guess that I could see a real benefit for those people who aren't necessarily as into, that's probably a bad term too, but into education based research as I am to find it more useful because of the, the dual meaning of some of those terms."

Other interviewees proposed the kinds of research questions they thought a tool such as the BCT could help answer. Interviewees C and E proposed more traditional uses of Bloom's taxonomy in research. Interviewee E discussed conducting research to evaluate the kinds of activities that are emphasized in different learning environments, and pointed out the accessibility of those findings for the community if they were conducted on chemistry-specific assessments with a chemistry-specific tool, saying,

> "Oh, yeah. I mean, you know, I mean, even, even just sitting here there are like a couple of ideas are coming into my head of things that you can do with that…Like, because my impression is that, that they, we do a lot of things that promote memorization and not higher up the taxonomy. Um, so you could really answer that question. Um, and, and if it's a chemistry-specific one, you know, then, then, you know, there's less of a leap to be made in terms of how it applies."

Interviewees A and L agreed with this kind of use of Bloom's taxonomy with interviewee A saying that using Bloom's taxonomy "can be just categorizing questions to put some boundaries on your research study, like [we] did way back in the day. And yes, I think this can be useful for, for that type of classification when it's needed.” Similarly, interviewee F thought the BCT could be used to retrospectively look at how assessment has changed in chemistry over time, which would also be applicable to the rapid changes that had to occur in chemistry teaching in light of the COVID-19 pandemic, saying, “It could be…one way to look at changes, you know, say changes through the pandem–it doesn't have to be pandemic-focused, but changes through the years of, of assessments. Are we moving anywhere at all in it? Certainly I could see it as an instrument to be used in research.” Looking forward in CER, Interviewee C discussed using tools like Bloom’s taxonomy to examine the outcomes of active learning modalities, saying,

> “I think we, at the point now where ‘let me think about active learning,’ um, whatever method it is. We had a point where people know it works and we see the outcome of it in terms of, um, grades, you know, DFW. And we see, you see a lot of that, but I no- now, I think people are thinking about how, what are the components of it that make it work well, you know, as an organic chemist, I say, what's the mechanism? Like, why, why, why does this work? You know, how does it work? How does it work best? Who does it work for? I think now we, at that point in the, in that, in the research endeavor where we try to figure out what are the things that make this work, what are the underlying factors? And I think this is one of the ways, because most of these active learning always include some kind of group activities. Students are working maybe on clicker questions in class, they are doing, you know, so whatever it is, how does this match up with Bloom's taxonomy? And then if, if we can match it up, then how does it relate to student outcomes?”

On the other hand, some interviewees felt there wasn’t a lot of novel research that could be done with Bloom’s Taxonomy. Interviewee B discussed her drift away from using Bloom's taxonomy in her research by saying, "I think for research purposes…other things have kind of caught my attention and and maybe I just kind of got stuck and lost creativity in terms of like, what can you do with blooming other than, you know, categorizing what things are."

For Practitioners

Several interviewees discussed the way that either they already use Bloom's taxonomy in their teaching, or proposed ways they could see practitioners using Bloom's taxonomy in their classrooms. Both Interviewee C and J use Bloom's taxonomy to align their course content. Interviewee J said that she thinks something like the BCT could be "kind of one tool in the grand toolbox of ways that we approach the classroom" and described how she uses Bloom's taxonomy to align her formative and summative assessments, saying,

> "For me, the primary utility of sorting by cognitive level is to try to match my kind of, to scaffold my class and match my summative assessments with my formative assessments, with my goals. Right. And really thinking about like, what is it that I expect students in this class to be able to do. And so then having words for that, for the differences between those things is a, is a useful concept."

Likewise, Interviewee C uses Bloom's taxonomy to write her learning objectives for her teaching.

> "How do you write, um, how do you know, what do you want your students to be able to do after the class is over? And what does that mean in a chemistry context? I don't think, I didn't know. I don't think

> a lot of faculty even think about things this way. I mean, I do it because when I w- I don't, I don't even think about it...I do learning outcomes for every chapter I go through. And I think about, okay, what do I want the students to have at the end of this chapter? And I write out all that, put it there for them."

While interviewee K does not use Bloom's for this purpose in his own classroom, he did suggest that instructors might find use in the BCT for aligning their exams with their learning objectives, saying,

> "The way that I would sort of view it as to be used in the classroom is I would kind of use it as a check as somebody's constructing an exam, you know, like just to kind of see, 'okay, do I have a representation of these different types of, of, um, of dimensions,' you know, like along this dimension, 'do I have different representations along this dimension? Do I have different representations?' Because sometimes again, like people can be really algorithmic heavy, you know? And so you might have a professor and an instructor and when you look at the exam, like it's complete straight calculate, calculate, calculate, calculate. And so I think this sort of tool could be useful to say, 'Oh, but wait a minute, there are other things that are important as well. You know, like there's a lot of conceptual things that are important.' Um, it's important to have students analyze representations to analyze diagrams and things like, and so maybe consider, you know, like taking out some of the calculations and putting it in some other things. Yes. And so for me, that's one of the primary things that I could, you know, think of and that I would suggest, uh, to possibly use it for…More in teaching than in research and…so then in terms of that alignment, I definitely think it would be useful."

Interviewees G and H also discussed how they see the BCT being used by practitioners in their classroom. Interviewee G suggested using it to more intentionally select homework problems for students, saying,

> "There is a distance between, uh, the results of research and the application of these results in the classroom. So, uh, quite often, you know, uh, take, take, for instance, the problem solving of question in general, in textbooks, you have, uh, a number of questions in the end of

> each chapter or within each chapter. And, uh, sometimes students have to practice these, uh, these questions, uh, quite often, uh, the books contain too many questions. The students do not have time. They start to, cannot allocate all the questions to students. So the point is for selecting some questions and what criteria teachers should be using to to to select questions. So, uh, to have some, uh, tools which can be variable tools like this HOCS/LOCS, uh, division, or whether, uh, the questions, they involve abstract or concrete thinking if the concepts are complex, or they're simple. If there are mathematics involved or it's a more descriptive the approach, uh, you have so many things in which you can distinguish and the tool, uh, can be useful if it can, uh, serve many, many purposes."

Interviewee H supplied that the BCT could make larger CER studies more accessible for practitioners, saying, "I think, yeah, it's pr- I could see it serving as a bridge between those sort of high-number standardized studies and a small somebody who's just trying to take the research that they're reading about and apply it."

Across interviewees, there was the most consensus that the BCT would be useful for instructor-training purposes. Interviewees A, B, C, E, F, and I all discussed how they saw Bloom's taxonomy fitting into instructor-trainings. Interviewee E specifically discussed the accessibility of Bloom's taxonomy as a strength for practitioners, saying,

> "I know that for, for some people, Bloom's has gone out of favor, but I think, you know, it's kind of like, I think about it, like the food pyramid kind of thing. Like, yeah, you might want to shuffle some things around and maybe, you know, it's just like any model. It's not the perfect model, but what's nice about it is I think it makes a lot of intuitive sense. It's very accessible. Um, especially, um, even though I don't, I haven't used it in any of my recent research. I do feel like as a practitioner, it almost helped me more as a practitioner than it has as a researcher."

She went on to say that she sees Bloom's taxonomy as an entry point for new faculty, saying,

> "Um, so I guess I think about it when I think about it, I think about it as this nice, like entry point into that as a way of mapping thinking for people, especially that could be useful for like me at the time, you know, I was a new post-doc, I hadn't had any classes or anything on education, and I didn't really know, I didn't know what I was doing teaching-wise. I didn't fully know what I was doing research-wise. And, and so this was like, I just could grab onto it, you know?"

However, interviewee E was clear that tools like the BCT can be useful for faculty at any stage in their career, saying,

> "People who've been teaching for a long time in chemistry still need this. I think that, you know, we don't, we, the way that we're trained to teach there's there's, we, we just write questions the way the questions were written for our quizzes and tests and, or we designed problem sets that are, you know, that were the same kinds that we had. And, and so we're relying completely on a textbook or a problem bank or our own brains, you know, which are just modeling after the way that we were taught to do this kind of thing…but what do we write is like a lot of, um, you know, kind of factual questions that are, remember, remember, you know, like w- we, we trigger memory a lot, or we maybe even like, trigger. Like, I think when we write those sorts of questions, then we emphasize memorization and, and that leads to problematic downstream effects. And I think about that, particularly with organic chemistry, where students are trying to memorize their way through organic chemistry, which is really problematic. Um, and I think that if you know, that, that for the same reason, this was useful to me as an entry point to thinking about, you know, okay, there, there is actually some thought that can be put into the design of things. Um, and, and to really think about what I want students to know and be able to do, I think, um, it could be really useful. And I do sort of agree that, that with the utility of it, I could see it as being, you know, for the, the reason like chemistry instructors, it, you know, at least that I've encountered, they don't have a lot of time. Some will invest more time in teaching because they really want to. Um, but I know that even though we have the center for research on teaching and learning, you

> know, our, our, our learning center, not very many people in my department walk across the road and go to that learning center and get help or learn things like that might be helpful to them. And i- Bloom's tax-, you know, like Bloom's taxonomy, it doesn't happen. Um, so I always feel, you know, I've come to the point where I feel like if we can meet people, um, part of the way, meet them where they're at, right? Like, you're, we want, we design things that are very general and then ask someone to do the work of figuring out how it works for them, you know? And, and that's fine, except that it means that, that I think that keeps probably a large number of people from using it, because they're not going to do that work. Right. So, so meeting them where they're are, where they are. And I think making a chemistry-specific or chemistry version or whatever you're calling it of, of, um, Bloom's, I think could help that need, right. It becomes more accessible. Um, there's still always the problem of dissemination of things. You know, even if it is chemistry specific, it doesn't matter general or specific or what it is, however useful it is. There's always the challenge of dissemination. But I think if you take that step closer, making it more discipline-Bloom's specific, I think could be really helpful."

Interviewee F agreed that Bloom's taxonomy is a useful entry point for new faculty and those who need to make changes to their teaching, for example in light of the COVID-19 pandemic, saying,

> "We just need to look at the pandemic and it was a rapid pivot to online for a, for a large majority of, of institutions and people having to think about how they're going to assess in, in this new way, or trying to make whatever they're doing now online as 'normal,' as, as possible, as normal as they can make it for, what they feel is, is normal, and appropriate. And, and, you know, in a way it's a great opportunity to think about assessment in different ways and have a bit of a wedge there to say, 'Hey, this is an opportunity for these conversations about, you know, what, well, what are we, you know, going to assess what, what really matters to us? What do we value? How are we going to know if, if, if students get there?'...Um and it means that people who are trained as chemists and doing assessment work need to normally learn different ways to assess, but also be able to have conversations about it and it needs to be accessible. So I think I, you know, I think Bloom's taxonomy is one really excellent entry point into that…I, I do

> think it has a really important value…So I can easily see the value for educators."

Interviewee A likewise discussed using Bloom's taxonomy to train faculty on how to develop items that are aligned with the type of evidence they are trying to elicit, saying,

> "I come down more on the side of, of instructor training than I do research for this. And the reason, the reason I believe that I think is that…is, it is really fundamental that you are in one way confident in another way are, you know, have something to point to that says that this item is going to elicit evidence of what you want. Right. And the extent to which mapping that construct onto a level of cognitive engagement seems superfluous when you're looking at it from that, from that perspective. Right. if I want my, if I want my students to remember the symbol for lead, all right, that that is what I will test on. Not, and not be sway– dissuaded by the fact that that is a very low cognitive engagement. Now I, I'm probably not going to ask questions like that, but it is the instructor's purview to sort of assess what they, what they want to know. And so that, that connection is more fundamental for me than than knowing what level of a cognitive engagement this is."

Interviewees B and C both discussed how tools like Bloom's taxonomy makes faculty think differently about their assessments. Interviewee B discussed how it helps faculty write multiple-choice assessments that are not just at the recall level, saying,

> "I think as long as we have multiple choice tests, I think there is a utility here. It's a tool that helps us break free from simple recall questions. It helps us think more deeply about the words we use on our assessment. What is it we're actually asking students to do? So I do see a value for these tools. I think they become very, very much like a second language for faculty who really study and work with them to be like, 'Oh, I really want this question to be, you know, an analysis. I want them to analyze stuff. So here's the language I'm going to use. Here's how I'm going to set it up, you know, those sorts of things.' So I

> do see that there is, there's value in these, these taxonomies to help us move beyond simple rote memorization questions, the reason why these become important. Right. It's just what Bloom said is like we've got bunches of large enrollment courses at all sorts of different universities. And we want to understand, you know, equivalence basically, right? Like what are we doing? And, and, and are we all asking them factual questions? Are we asking them these range of questions? Right. So it does, there is utility here in terms of helping us develop better assessments, helping us think broadly about what it is we're asking students to do, to know, be able to do those sorts of things. So I think it really works well for that kind of jumping off point into better assessment design, better classroom and curriculum design…I could imagine that like using Bloom's taxonomy as, you know, faculty professional development to help faculty develop different assessments or to think about their assessments in different ways."

Interviewee C was much more interested in how tools like Bloom's can help faculty learn how to align learning objectives, saying,

> "I think it can. I mean, because I think it forces you to, um, to think more deeply about what is the learning objective and what do you want students to achieve by the end, because bas- then, and then structuring your teaching. Like you said, people say you should be, you know, backward design. Right. And I think some, most of our instructors, they just go with, 'okay, this is what I, this is the book of how to teach chapters one through 12,' you know, but do you need to teach everything in every chapter? You know, and at what level, because there's some things I pick and choose every time it's like, you know, 'this is not that important. I can just mention it and they just need to recall it' or something like that, but they don't need to know, you know, but 'here's the something that they really need to understand really well. And this is what I expect them to understand about it.'"

Interviewees E and I both had suggestions for how instructor training with the BCT should be conducted. Interviewee I wanted to consider what training would look like for people with no learning taxonomy experience, saying,

> “So what sort of training do you anticipate going with this? Cause I think if you give this to somebody like me, for example, with no training at all, you're going to get my lens on what an explanation means and what an argument is and so forth. And it's, we have the same problem with with the LAP. I mean, we're in a workshop I think all the 3d learning stuff has the same problem because colloquial ideas of argument for example, are pretty different from what we mean by scientific argument.”

Likewise, interviewee E discussed how important worked examples of sorted items and modeling how to use the tool would be for training people new to the BCT, but that she saw no issue training people to use the tool for research, saying,

> “What was really helpful to me is that you w-walked through that first example. Um, I mean, I, I think I could have done it, but it helped me to, it helped me to reorient to the taxonomy in, and think about, you know, a way to approach that. Um, I think partially it probably also felt easier because, um, maybe because it felt like you went the reverse in terms of the difficulty hierarchy. Um, so that was kind of helpful because I definitely, I had a harder time. I felt, I felt like I had a harder time with my first independent question, which is [item Q6,] um, in that one, uh, just because of the form of it, but like the last one [item Q1] to me that was super clear. So I could still argue, I could still kind of argue with myself a little bit on the first one, but the last one I don't, I feel pretty confident in…because I know, um, I'm familiar with Bloom's taxonomy or at least I had used it quite a bit prior, um, you know, in years prior, uh, you know, it's very, it's very accessible to me already. I guess if I were someone who, um, if, if I were someone that, uh, was not as familiar with like, like someone who's a new instructor, who had never heard of Bloom's taxonomy, um, then having a worked example would be good. Right. And even a chemistry-specific example…But I think a worked example, um, would help them to use those to an extent…from a practitioner side, I'm just aware of, of people just, you know, really having subject matter knowledge and chemistry and not understanding things related to, um, you know, educational theory or practice. Um, so for those folks, it, you know, it's not that I don't think that they could do this, even from that. I think it would just make it a step easier if there was a worked example. But,

> um, it, I don't feel like I needed that though. You did really give me a worked example in the beginning. Um, but yeah, research-wise, I don't feel like this would be a problem at all. And I think it could be really useful."

For Students

Interviewees B, C, and E discussed the way that the BCT could be used as a metacognitive tool for students. Interviewee C already uses Bloom's taxonomy to communicate her learning objectives to her students, saying,

> "I said, okay, this is some way for you to look and say, 'okay, look at each learning outcome. Am I able to do this? Can I, you know, perform this? Do I understand this?' Um, so I asked them to use it as a guide for when they study, like, you know, go through each ones, 'Okay. I'm going to try to look at learning outcome one today. Can I do that one, can I go through these problems? And I know that I understand it,' then, you know, move on. It's a way of helping them to kind of, um, have a goal for when they go study things like that."

Interviewee E said that the interview made her consider why she isn't already using Bloom's taxonomy in this way, saying,

> "I mean this, even having this conversation makes me think, why haven't I discussed in my class, you know, and it may be because it's chemistry education, I'm trying to, to show them chemistry-specific things, but that's me not remembering what it's like to just go, 'I don't know anything about education,' you know, and like, um, what, it's, what it's like to be starting off and thinking about this is actually really accessible and useful. Um, so, uh, yeah, I could see a lot of potential for that."

Finally, Interviewee B considered how in her field of Biology, the COVID-19 pandemic has only exacerbated the rapidly switch in emphasis of skills that students will need going forward, saying

> "I'm coming at this from definitely a biologist standpoint where everything in our field is changing so rapidly. I mean, COVID-19 has taught us that hook, line, and sinker, right? And our students, we can teach them a lot of facts, but at the end of the day, they're actually going to need to be able to do those other things. They're going to need to think about applying and, and those, you know, procedural things

> like, 'How does this work? What do I already know? What systems am I familiar with that I can apply here? What do I know? What don't I know,' that metacognitive piece of things. So I think for our biology students, as the field changes so rapidly, some of those other things are going to be much more important for them than being able to tell me the steps of photosynthesis, for example, right? That is probably going to be something that is a lot less important for our students moving forward."

Proposed Improvements to the BCT

In the last line of inquiry from the tool-testing interviews, participants were asked if they foresaw any issues implementing the BCT into their classroom or their research. Suggestions for improvement to the BCT followed several themes, all of which correspond either to the BCT's construction or its use. For clarity, these themes have been combined into the two following sections: Structure and Function. In these sections, the proposed improvements to the BCT will be summarized, and example quotes from the interviewees will be provided.

Structure

The first set of proposed improvements to the BCT related to the construction of the tool. Interviewees had widespread opinions on whether Bloom's taxonomy, and by extension the BCT, should be considered to be hierarchical, (which points to some of the challenges of addressing misinformation about Bloom's taxonomy when attempting to disseminate the BCT,) and how that affects the included cognitive processes and the Knowledge Dimension.

The Problem of Hierarchy

Several interviewees had different opinions on the assumed hierarchical structure of Bloom's taxonomy. One issue that came up repeatedly in interviews was that of an assumption of hierarchy in the Cognitive Process Dimension. Some interviewees thought that the Revised Taxonomy, and by extension the BCT, was still assumed to have the same cumulative hierarchy as its predecessor and objected to it, while others assumed there was a cumulative hierarchy to all iterations of Bloom's taxonomy and appreciated it. Some interviewees knew that there was no longer an assumption of cumulative hierarchy in the revised taxonomy, but objected to any assumption of hierarchy in sorting cognitive items. Interviewees B and E were among those that spoke highly of the assumption of hierarchy in the Cognitive Process Dimension, while interviewees I and M were less convinced. Interviewee E thought the structure of a hierarchy was useful for beginning instructors, saying,

> "You know, I think, I mean, I think the hierarchy is useful. Um, I think one thing though that, or I think if you're a beginning instructor as it was, for me, it's helpful in going 'wait,' you know, because I think that you need to have a balance of the, because there is somewhat a difficulty it's not as difficult to remember something as it is to analyze something. Um, so I think that hierarchy is useful, whether or not there's necessarily like is, is analyzing and evaluate, you know, those feel pretty close to me in, whereas remember versus analyze, feel quite different in their difficulty. Um, so maybe, you know, there it's harder, there, there isn't as much of a gap like between individual things. Um, but, but it, I think the, the overall hierarchy makes sense.

Interestingly, Interviewee B agreed that she prefers a hierarchical taxonomy, but that she is comfortable with items not fitting cleanly into those boundaries, saying,

> "I also liked that Bloom's is hierarchical. I mean, most [learning taxonomies] tend to be, but I liked that hierarchy that if you're asking somebody a question at the level of synthesizing or evaluating, they necessarily need some of these lower level cognitive skills. They need to have some you know, facts under their belt that they understand and can apply and use, sorts of things. So I like that it's nested, but I'm, I'm an ecologist by training. And so hierarchies are kind of where we live. We'd like when things are nested, because that just intuitively makes sense to us that, that, that it would build, you know, upon those lower levels…I see that as, I don't think it's neither a pro or con, but I think other folks feel it's a con, like a negative that, that there is a lot of, for all that scientists embrace, tend to to, the tentative nature of science in the classroom, I think they really want very clear articulated sharp boundaries for things like 'this is lower-level, this is higher-level' sorts of things. And the idea that it could be fluid, right, that, that an application question could be both, could be lower-level or a different application question could be higher-level just depending on that, I think is uncomfortable for many faculty to, to sit in that space, to know that 'I could write application questions that can go kind of either way.' So I kind of fall on the side of 'it is what it is.' But I think for faculty, it feels like a negative.

As an example of Interviewee B's previous point, interviewee I expressed his discomfort with a flexible hierarchy, saying,

> "I don't think that it's reasonable to suppose a hierarchy, the practices, and that's my, I believe there's not a lot of theoretical basis for that hierarchy. That was my understanding…But that's going to make it really important to justify the hierarchy. And I don't know how to do that…because you're probably not just going to explain stuff without models and chemistry, cause it won't make sense. So I see that goal is pretty sophisticated. But I don't, I mean, you might be able to, you, you can come up with some really sophisticated explanations. I mean, you talk to some phys-org people and they'll blow your mind, with the crazy stuff orbitals do. I mean, and I, the status quo is at least in my

> experience, so focused on skills and recall that even the crude like percent three-dimensional thing is a huge improvement. And so I, I guess I can see the value hypothetically in a hierarchy, but I, I have a hard time imagining what that would look like and what sort of theoretical justification it would need. Cause you'd have to be really sound on that because otherwise values are just going to bleed through in a hierarchy, as I think happened with, with Bloom."

On the other hand, Interviewee M convinced herself of the hierarchy of the first three levels of the Knowledge Dimension during the interview by discussing how conceptual knowledge nests within procedural knowledge, saying, "So I think that maybe once you get to the higher level for procedural, it seems to me that it would inevitably intersect with conceptual knowledge, that those are a lot harder to disentangle when you get up there. Just because you can't design an experiment with, I mean, you have to draw on your conceptual knowledge."

A common objection to the idea of Bloom's taxonomy having a hierarchy came from interviewees conflating cognitive skill with difficulty. Interviewees A, E, and J each had thoughts on how difficulty relates to the assumption of hierarchy in Bloom's taxonomy. After discussing the lack of assumption of cumulative hierarchy in the Revised Taxonomy, interviewee J mentioned that when she has coded using Bloom's taxonomy in the past, she hesitated to code more difficult questions as lower-order items, saying,

> "I do feel like when I was coding, I would sort of struggle with, especially understand if it's things like, 'well, this can't be an understand question because it's kind of hard, um, and understand is a low level skill,' but sha-, I like this, um, sort of alternative way of

> framing [the hierarchy,] that these, these are overlapping in terms of, of, um, I guess difficulty, cognitive requirement. Yeah. And so then is there a different, I guess your, part of your question then is, is there a different, or a better way to distinguish between those three things so that, um, we get back to more discreet, easy, medium, hard…I get, in like my colleagues and I, as we're discussing like how to educate students, um, I see increased or like, I guess, decreased emphasis on anything that has to be memorized, um, that we are really against asking students to, or like as a group, there's this push to move away from remembering anything. And the problem is that it, if you fee- the one, the one I'm thinking about is something like solubility rules, but if you don't remember your solubility rules and like you don't re- sort of remember that that is a thing, then it makes it harder for you to understand solubility equilibrium, because you don't have, you don't have this concept of like, 'things are soluble and things are insoluble.' Um, and so then 'let's add some complexity to that' instead just comes out of left field, like, 'oh, and apparently you can do this for solids.' Um, so I think that's really, I think there's also something about our educational moment in like in the generation, you know, now that, that, that, um, you know, I'm, I'm a millennial, right? Like I grew up with Wikipedia and I can look anything up on Wikipedia or on my smartphone. And so why do I need to remember anything? And my colleagues that are newly retired or close to retirement age, they can tell you what all the colorimetric tests are for ions. Right. And that is useful knowledge. And maybe, I shouldn't remember those things too, instead of, that elevates, the remember domain."

Interviewee E agreed that the order of Bloom's taxonomy is linked to difficulty, but that feels comfortable to her, though she does note that she feels like Conceptual Knowledge and Procedural Knowledge could be switched and it would not affect the order of difficulty for her, saying,

> "If I kind of look at conceptual and procedural, again, those are really hard to parse in terms of their difficulty. Some things conceptually are incredibly difficult and would be more difficult. Sometimes I feel like procedural is closer to factual. Right. Um, although factual feels like a step towards conceptual, you know? Um, so, so it's easy. Like the factual feels like the anchor to me. I don't know, like you could have re- you could, um, reverse conceptual and procedural and I wouldn't argue

> with it. I think those are, those are problems-specific. Those are content-specific. So remember, understand, apply that order is intuitive to me because, and I, you know, I don't know why that doe- the ones that could potentially be reversed are analyze and evaluate, um, create, definitely feels like it's the upper anchor. Um, but again, I wouldn't argue with any order of analyze or evaluate and, and I probably am more comfortable or feel like there's less argument for switching, analyze and evaluate than there is for switching conceptual and procedural, because I can probably easily come up with conceptual problems that are quite difficult relative to procedural problems, um, analyze and evaluate in terms of difficulty, they feel somewhat comparable. But I think in terms of a natural flow, you probably would analyze things on your own because evaluate feels more like you're, you're evaluating something that someone else has produced, whereas analyze is you're given something and it's more raw, um, what you're starting with. So it, it feels like a more natural progression. So I think it probably makes sense as it is, but in term, but, but more because of the progression of things. Um, and I do think that you have to be able to analyze something first before you can evaluate someone else's analysis, if that's the way that you want to think about it. Um, and only because these are so partially because they're so general, does it, does it feel like evaluate should come next? Um, I think that's, so I guess I'm talking myself in a circle, which really is just like, I think probably that one is fine though. Those, those in terms of like strict difficulty, I don't know that they're that much more difficult. Um, but it's the conceptual and procedural are probably the two that still like of any of the, anything on the taxonomy are the two that I feel like are completely interchangeable in terms of their difficulty. Um, that makes a lot of sense to me."

Interviewee A thought that the two constructs should not be related, but that they were inextricably linked in Bloom's taxonomy which reduced the usability of Bloom's taxonomy for research purposes for him, saying,

> "What I'm trying to say is that level of cognitive engagement should not, and it's probably a duh statement for you, be acquainted with difficulty…And so, so yeah, I, I suppose in closing, here I, I do struggle a little bit to see how categorization on Bloom's taxonomy as it's existing right now would be, would be useful, but also maybe I've just

> not had the experience or thought of an experimental protocol where it might be useful."

Ultimately, Interviewee A's objection to the idea of a hierarchy is a matter of evidence, saying,

> "I wish we had more evidence that these are not only independent of one another, but also hierarchical. Um why, why is create more cognitively engaging than analyzing?...if Bloom's taxonomy in the past was ever presented as a 'you're just classifying and there are some grand distinctions among these six things, but you shouldn't treat them as a pyramid,' I probably, you know, I would have a lot less problem with it."

The Problem of Explain

A related issue to that of assumptions of hierarchy and conflation with difficulty is that interviewees seem to interpret the verb "explain" in multiple ways. Interviewees A, G, I, and J expressed discomfort either with the verb "explain" being under the heading of the understand cognitive process, or with the idea of "understand" being perceived as lower on the Cognitive Process Dimension, while Interviewee L expressed a similar sentiment about the term "conclude" while sorting item Q7 in part two of his interview. Interviewee G discussed how understanding thermodynamics is not a simple process, saying,

> "Another thing I wanted to, to question was about this, uh, cognitive processes. Remember this is categorized as LOCS, understanding as LOCS, apply LOCS and HOCS, and then analyze, evaluate, and create as HOCS. I could think that, uh, this understanding level can be more demanding in some cases…So the thing is, what do you mean by understanding, for instance, you know, if you read the chemical thermodynamics, uh, now, this is a microscopic subset. Uh, it deals

> with the variables which are familiar, uh, objective, you know, like mass, volume, uh, pressure. And, uh, one might think in the first place that is, should be a simple subset, but to understand thermodynamics, it's very demanding. You understand it. Of course it's a different thing. Quite often, we might understand or think that we understand, uh, various fields of, uh, of, a discipline, like in physical chemistry. I can say, I can understand thermodynamics. I can understand statistical thermodynamics or quantum chemistry, but, uh, understanding of course, is when I read a book about it, I'm happy about it. I, I, I'm not have queries in my mind. Sometimes I might have. Quite often, you are not happy with something. You don't understand it. And, uh, you, you want to read more, you try another book, another reference to to to know more about it, but it's a different thing when you move into trying to solve problems. So if I go to problems with thermodynamics or quantum chemistry, uh, some of them you find hard to do, unless you have some experience with them, some practice with them, if you, if you deal with them for the first time, even if you understand the concepts, certainly the problem might be difficult. So that of course takes you further than, uh, knowledge and understanding. But still, I think that understanding, uh, can be, uh, demanding, especially, we are dealing with conceptual understanding. And now of course we have concepts which are abstra--abstract and concrete, and we might be thinking that concrete concepts are simpler, but they may not be, as I said earlier, thermodynamic concepts and variables are concrete, but that's mean that understanding thermodynamics is, it's more difficult now."

Interviewee I agreed that the process of explanation in science can be very sophisticated, explaining,

> "Yeah, I mean, my, my first thought when I was looking at how describe, explain, generalize, and illustrate are all up together is that doesn't gel with how I think about explanation…If you think about the scientific practice of explanation, it's much more than illustrating and describing, it's fairly sophisticated…the other element that sometimes is missing in three-dimensional learning, I don't know the extent to which it's missing in the Bloom's work, is the central importance of phenomenon. So in fact, for someone to engage in three-dimensional learning, you would want to explain or model or analyze or whatever a thing that happened…You're not going to just sit on a mountaintop and just come up with some sort of explanation for tertiary

carbocations stability because the sun hit you in such a way, and you just feel moved to, that doesn't make any sense…Yet, that's how we conceive a lot of assessments! I mean, heck it's like, 'why is this carbocation more stable than this one' is a chestnut organic question, despite the fact it's not grounded in anything observable as, as I've just framed that. And so when you think about explaining, I mean, so we do some high school work. So there's an anchoring phenomenon: we use evaporative cooling. I mean, why, why do you feel colder when you're wet and when you're dry? So explaining that is really non-trivial thinking about phase changes in terms of disruption of interactions between populations of molecules energy transfer from your body to this liquid on the surface, disrupting those. And then some of them leave as gas molecules, whisking kinetic energy away from your, from your body. That's hard. And you could illustrate or describe that in a very vague sense and not have any idea how and why it happened. So I didn't like that being in that box….it isn't clear to me why some things [in the BCT] are grouped together. And so the thing I brought up was 'describe, explain, generalize, illustrate.' So I would argue explanation is the heart of much of science. I think model-based explanation is the heart of chemistry…So this idea of relating actors at a causal level below a phenomenon to how that phenomenon happens, and also thinking about the cause sort of the, how and the why, which you've also talked about in terms of the knowledge dimension, that's all under the heading of explanation to me. And I think that's consistent with a lot of the scientific practices, literature around that practice that is really different from just describing a thing that happened. Like I can just take a look at a beaker and you know, you put vinegar and baking soda together in that beaker now, 'oh, it bubbles,' that's a description…The explanation, well at first you have a proton transfer to bicarbonate and then you have a solvent assisted decarboxylation, that's driven both by the departure of the carbon dioxide, the stability, the carbon oxygen, double bond, and the fact that you just made a gas. And so you've got more microstates and it's entropically favorable. That's not the same. Those are way different. And they're next to each other and it doesn't make sense to me. And then I, and this is, I have, anyone who has worked with me any length of time will tell you, I just fricking hate stoic because that's become the way we like, make gen chem discriminate against people for no reason….I see I go on up [on the BCT] and I see calculate. And I'm like, 'Oh, no, calculate is not more sophisticated than a causal mechanistic explanation for a phenomenon!' This is not true!...Th-the reason I like the science practices, is there's a lot of literature behind…argumentation

> modeling, explanation analyze, interpret data, all those have a ton of literature. I know what those mean, like really precisely. And I worry that if people don't have a precise idea of what all the terms under these bins mean, you're going to really introduce some major bias depending on the community using the tool…I mean, of course the question I had at the outset was 'why don't you just use the 3D-LAP?'"

Interviewees A also compared the act of explanation with the relative complexity of the term "modeling," saying,

> "Well, I mean, there was that one instance where I went, 'this question has the word explain in it, but I do think it's another, I think it's a very much higher-order thing than understanding.' Right….And I do think in a, in nowhere, do I see for example, Oh, well it under create there is model, which I think is a very, very science, chemistry practice. Right. but that, I think that's being used under create as the authentic practice of modeling. Right. Not using an established model to explain something, right? Or, or, or bringing. So I think there is, there is a lot of gray area between explaining with a model and modeling that this tool doesn't capture sort of the cognitive engagement of, of, of that idea."

Interviewee A also discussed the homonymity and discipline-specific nature of the term analyze, saying, "I think analysis is, is is another one of these tricky things because we as scientists have a very particular idea of data analysis and depending on what that is, you know, that is that, is that a procedure or is that cognitive engagement?" Interviewee J continued, discussing the overlap of the cognitive processes of apply and analyze, and how that relates to students' perceptions of difficulty saying,

> "From my previous experience, I have found that most of the questions that I ask students are in the apply and analyze categories. And so then, um, the, I, I found in my work that students couldn't distinguish between what was needed for those two different modes. Um, and you know, yeah, probably I have difficulty distinguishing sometime, you

> know, in some cases. Um, but so like what, what is the, what, what is the difference? The way I distinguish them for myself is mostly…So like, if I give you information and ask you to explain it, that is probably more of an analyze thing, because you figure out it's important. But, um, so yeah, like most of what I want students to do in my classes, both at the lower-level and at the higher-level fall into that apply and analyze category. And so then this limits the usefulness for telling my students what I'm expecting from them, because it's like…for every learning goal, we have to figure out an apply or analyze version of this and then we're done. Um, so I, I wish that there were maybe a little, are there, is there a different way to slice it that, uh, for a, for a chemist sort of separates those activities. But some of that is also kind of dependent on like, my context, you know, in my institution. Um, I'm also really thinking about the like, summative testing environment. I remember my students were really terrified of create-level anything, um, especially in a timed context because like, 'Oh, we have to like come up with something.' Um, but in, in my study, I remember we did a practice example where students had to devise a mechanism for a, uh, organometallic reaction that they had never seen before. And, you know, I said something like, 'so if I gave you this on a test, an organometallic reaction you've never seen before, just draw a mechanism, in this class, that would be a create level of question.' And they were like, 'oh yeah, that is hard. And, um, also like it takes time. And so I see why that's not on the exam.' Um, so in that way, like, that distinguishing between create and some of the other things is useful."

The Problem of Bidimensionality

An aspect of the BCT that was new to many interviewees was the Knowledge Dimension. Several interviewees discussed their frustration with the pyramid presentation of the Original Taxonomy. Interviewees A, G, and J all brought up their issues with the popular presentation of the Original taxonomy. For example, interviewee G did not understand the purpose of the pyramid narrowing as it went up, saying,

> "I think I have not understood about this triangle here is that, you know, as you go higher, you have a shorter base to the triangle. Does

> that mean anything? Well. One way of seeing that could be that know in education, we deal more with knowledge than with comprehension and, uh, as we go up with deal less in terms of, uh, uh, the time we dedicate our teaching, during our effort, otherwise I cannot see, uh, a reason for having these triangles there."

Likewise, Interviewee J had questions about the pyramid presentation on the matter of unidimensionality and how it related to the BCT, saying,

> "The other thing that I feel like I hear from other people is, well, 'students have to, they have to use knowledge in order to do any of these other things. So you're always testing knowledge.' I don't think that that's necessarily always true, but, um, yeah, what's the, I guess, as an instructor, I would kind of want to know, like, what is the relationship among these things, they're often a pyramid? Is that the best representation?"

Some interviewees felt that the bidimensional nature of the BCT was an improvement on the Original Taxonomy's unidimensionality, whereas other interviewees argued that the appeal of Bloom's taxonomy is its ubiquity and recognizability, and that adding elements to it reduces its usefulness. Interviewees A and G both appreciated the addition of the Knowledge Dimension, while interviewee B discussed how in her previous research, the use of the bidimensional structure over complicated the tool for their faculty, saying,

> "When we started, we started using the language of the revised taxonomy and faculty were like, 'what is this? Like?' Cause they had already had some familiarity with Bloom's original and to them, it, the the dimension that changed, you know, things from, you know, comprehend and understand those sorts of things just seemed like semantics and faculty just didn't have time for that. And when we were trying to introduce and integrate with this second dimension of that, right, which you have on your tool, which is really um interesting

> to see, faculty, I don't know, it was overwhelming, I think for folks to think about that…And, and it was just, it was a struggle because faculty, weren't ready to think about these two dimensions and how to prioritize. They were okay with one dimension and thinking about, 'okay, I want most of my items to be, let's say I'm an intro bio teacher. I want most of my items to kind of be in the middle, maybe middle lower end of things.' You know, maybe, I don't know, that's just an example, but to then think about this second dimension, that was just more than, I think a lot of faculty could handle, so we abandoned it."

Interviewee B also continued that for the purposes of the research they were trying to accomplish, a second dimension resulted in higher granularity than they needed, saying, "We did start to try to create a tool that looked at both of those [dimensions] and when we could fill it out and it became onerous for us and not that useful. And so we abandoned it, but I mean, part of it was, we were just looking for, we were looking for tools that could um, be quickly implemented by faculty and be interesting. And we found that faculty were less interested in that second dimension."

On the other hand, some interviewees wondered if the BCT could accommodate additional dimensions. Interviewee G suggested adding levels of cognitive development, saying,

> "Could you think of a third dimension to it? Uh, it could be, you know, psychological part, you know, you know, if that involves field dependence, field independence, if it involves, uh, uh, working memory, if it involves Piagetian-level of cognitive development. Things like being, uh, creative thinking, divergent thinking, convergent thinking, just other things, because you have all these things and quite often they are scattered around and, uh, it's good to to to try to combine them, so but your instrument I think is, is very good in combining in this, these two dimensions."

Interviewee J had two suggestions for additions to BCT. In her teaching, she considers whether an her assessments are balanced in terms of algorithmic, conceptual and descriptive problems, saying,

> "I do think about like, uh, sort of questions that require, especially at the gen-chem level, questions that require a numerical answer, um, versus questions that require a conceptual answer versus questions that require sort of, uh, I guess, um, more of a like identification kind of thing. And so, yeah, so something that helps- would help me think about the relationship between concepts, visualizations, and pictures. I teach a lot of gen chem, if you can't tell, um, and so the difference between thinking about what the molecules are doing versus the calculating calculation, um, problem solving piece, um, yeah, that those are sort of things I think about differently. And I, and when I'm writing an assessment, I am thinking about the balance between, 'okay, are you explaining those things with molecules versus explaining them with math?'"

She also thinks about the complexity of items in regard to how many steps it might require, saying,

> "Another way that I think about like, the sort of 'hardness' of a question is in terms of how many concepts get connected, and also like how many kind of, how many steps in the sense of, do you have to solve for one thing in order to solve for another thing? So like, do you have to, um, use the half-life equation to get the rate constant, and then you can figure out the concentration at a particular time. Um, because that, like, there's a complexity there that I see my students struggling with, which is the, and I guess this is related to that conceptual knowledge of how do these, how many different pieces can we string together? So I could see that being useful."

The Problem of Metacognitive Knowledge

Despite being mostly pleased with the bidimensional structure of the BCT, some interviewees were perplexed by the inclusion of the metacognitive knowledge level in the Knowledge Dimension, either because they weren't sure how it could be used in assessment, or because they thought it didn't quite mesh with the other levels of the Knowledge Dimension. Interviewees E, F, and H felt that metacognitive knowledge was slightly removed from the rest of the Knowledge Dimension. Interviewee F thought that maybe metacognitive knowledge affects the rest of the BCT intersections, saying, "So I do love the Knowledge Dimension. It seems interesting to me that the metacognitive piece is almost mixed in, on the same dimension as the knowledge piece. To me, it almost, almost seems to be like a third, a third or integrated dimension across all of it," while interviewee E thought that metacognitive knowledge shouldn't be included in the structure because it didn't follow the rest of the dimension's hierarchy, saying,

> "One thing that might be challenging is like metacognitive knowledge, um, you know, thinking about thinking it, it should be in some sense that almost feels like it should, it should be used for everything, like once you have it. Um, but I can also see how that is, that is like, context-dependent. Like if you have metacognitive knowledge about general chemistry problems and then you get to organic chemistry, those are, it's a very different kind of, um, thinking and problem types that are in organic chemistry. So, um, you, it's not necessarily going to transfer over and, and, but I guess it's not something I don't think that you have to have, like factual knowledge of something before you can have metacognitive knowledge or use metacognitive knowledge in that scenario. So that might be the one that kind of feels like an outlier to me, or, you know, and maybe that just has to do with the way that I, that I like to teach or even the way that I was, or tried to be as a student."

Interviewee H did not necessarily feel that metacognitive knowledge should be removed from the BCT but liked that it was at the bottom of the

Knowledge Dimension, which he felt de-emphasized it, because he did worry about practitioners not knowing how to employ it, saying,

> "I think a lot of people don't understand what [metacognition] means and I'm not sure they're aware of their own cognition. For lack of a better term, would be 'dumb' enough for some of my colleagues to understand it…Yeah. I I I don't know if this necessarily needs to be removed, but I like the fact that it's down at the bottom, which means it's the last one you come to. So you're not, cause I think you're going to be operating more up in that top corner generally, depending on the questions you're asking and the context. So I like the fact that it's down at the bottom. It may be that idea of summative versus formative assessments may, it may be a good idea to put that, that this would be, you know, more of a formative assessment just in that one. But again, that's like, I struggle, because a lot of those terms are education-specific terms that I'm not sure…That we're all familiar with, but nobody else necessarily understands or distinguishes between formative and summative."

Interviewees C, L, and M each had questions about what metacognitive items might look like in college chemistry. Interviewee C wanted to know if metacognitive knowledge applied to college chemistry, saying, "I wonder what meta- metacognitive, um, knowledge will look like in a chemistry context and whether it, you know, applies that's one thing. Um, but I, you know, in terms of, um, if you were creating learning objectives or you were looking at it in a research context, um, where would that fall for chemists? Um, so I didn't use that one at all in at least in these questions." Interviewee L also asked about what metacognitive items might look like, but then worried about faculty wanting to only focus on what they perceived to be the highest skills on the BCT, saying,

> "There's a part of me that's like, I'd really like to know what a metacognitive question would be, um, for like a general chemistry course. Um, just off the top of my head, I'm having difficulty thinking of one…And that's, that's another problem is, we tend to look at this and, and I would say that most faculty who want to use this tool would say, 'okay, so I need to be building questions that point students towards metacognitive-create'. And it's like, 'no, no, no, no, no, no'...so obviously I just fell into that trap: 'well, what would this look like? A metacognitive question?' Cause, cause we want students to be metacognitive. That's a goal, but we also want them to think conceptually and critically, um, and understand like, procedural skills and, and in all these different areas of, of cognitive processes."

Interviewee M similarly asked about metacognitive questions in college chemistry but also asked about other intersections on the BCT such as creating factual knowledge and creating procedural knowledge.

Interviewee B had addressed questions of whether all intersections of the revised taxonomy resulted in useful skills for students to practice in her previous research, saying,

> "I remember sitting in the box that would be where remember coincides with uh, metacognitive knowledge. Right. And I remember thinking 'how useful is that box for a faculty member or for a student,' and then faculty looking at this and saying, 'well, do I need a question in every one of these boxes? That's insane. And, and how do, how do I prioritize?'...We started trying to put items into, you know, like 'where would items fit?' And we had a lot of holes and we tried to figure out what would fill those holes. And that's where we kind of stumbled. And we wondered, we started wondering, 'do we need to fill that hole? What's the, you know, what's the importance of that hole?' And then, you know, then you go existential and you're just like, 'well, what's the point of assessment in general?' And you're just like, 'oh, wow this is, this is not a good place to be.' It turns out. Um and you have a bad semester where you assess nothing. No, I'm just kidding. But it just became like trying to fill that out. We, we found we couldn't fill it out and we just didn't, we weren't sure of the utility and being able to fill it out, at least at that point in time for our, for our faculty and our

> students…some of these boxes are not going to be filled in. That's the one thing that I always I always think about is like, 'I don't think all of the boxes are, are necessarily going to be filled in.'...No, I think you've done a really thoughtful job here…I think it's also interes, might be interesting to look at the things that are not on the intersection, that are not maybe as important as 'why are they less important? Could they be helpful? Could they be important?'

Function

The second set of proposed improvements to the BCT related to the use of the tool. Interviewees had suggestions for how to simplify operation of the BCT and how to further specify it to college chemistry.

The Problem of Granularity

A major theme in interviewee suggestions for using a variant of Bloom's taxonomy was to consolidate cognitive processes for ease of use and to align it more with other learning taxonomies such as Marzano's taxonomy, the SOLO taxonomy, and the ICAP framework. Interviewees A, C, J and N all felt that the level of granularity in Bloom's was excessive and compared Bloom's taxonomy to Marzano's taxonomy, with interviewee A saying, "I'll just go back to, like, why, why six cat–? Like, why are there these six categories?...I think we, we talked about this before. I think the language of, of the six checkboxes is really deleterious in a lot of ways to sort of what we practice as chemistry as a scientist." Interviewees C likewise said,

> "[Marzano's taxonomy] doesn't have as many levels. So I think it was, it was just the four or five levels and, um, it kind of summarize and I probably have to look at it to remember all the, um, I think it was just a little bit more precise and I didn't have to go through, you know,

> Bloom's, is this a seven, how many different levels? Six. Yeah. Um, let me see. I'm trying to see if I can. Yeah. So it has four levels. And I think what, one of the things I noticed, like in some of the Mazano tax-taxonomy, um, they, they have the level, but they also have like a mental process that goes along with that, with that level. So maybe like a level four would be things like investigating, experimenting, problem solving, decision making. So those are mental processes they think go along with this knowledge utilization level four. Um, whereas the level one is the kind of recalling, executing, recognizing that kind of thing. Um, and I know in the, the verb is in verbs and phrases that go along with it, similar to, um, Bloom's, but it's just those four levels. And instead of having to deal with, you know, the six levels, I think it was, it was just simpler for us to use it. Um, and just categorize a question into one of these four levels."

Similarly, interviewee J explained that she thought the Knowledge Dimension would be easier to sort along for undergraduate students because there were less possibilities, saying,

> "One thing that I've been thinking about is, I think that six levels is too many for an average undergraduate student. And so anyway, I've been thinking about, are there like, would it use a simple, like Marzano's, I think it has four levels and …that's a better way of thinking for chemists, for example, because, um, there are fewer differences…I'm less familiar with the knowledge dimensions, those were sort of easier for me to, to sort. The cognitive processes were harder for me to sort, because there sort of are so many of them."

Another issue of granularity in Bloom's taxonomy is that of higher- and lower-order separations of the Cognitive Process Dimension. Referring back to her previous research using the Revised Taxonomy, Interviewee B addressed this dynamic, saying,

> "So Bloom's, at least when we initially were using it, we were just using the cognitive dimension. And so there's kind of six not kind of, there are six areas, not areas, but levels to it. And they can be broken down even more like you can just focus on lower-order versus higher-

order. Although that brings up a con, which is where does the lower-, higher-, where is that bar is that, you know, some people are like, 'well only levels one and two knowledge and comprehension…are lower. Everything else is higher-order.' Some people say, 'no, it's right in the middle,' and there's really no evidence. There's no justification, that it's all very arbitrary and subjective there's no, and I don't even know that you could objectively, you know, come up with a lower- versus higher-order because even that, that idea of cognitive skill level, that's also subjective, right?

Not all interviewees were sold on consolidating the Cognitive Process Dimension. For example, Interviewee H said, "I liked the description. I didn't pay that much attention to the labels until I was actually assigning the box. I think that the descriptions do a good job of separating those levels. The, yeah, so I don't, I don't see any issue where there, there's, you know, you can argue one way or another." while interviewee L suggested keeping all six processes, but focusing on better stratifying the levels, saying,

> "A recommendation might be to, to emphasize certain pieces that help to differentiate different definitions. Yeah. So, so I mean, I know that they're basically a list of adjectives, not adjectives, a list of verbs for most of these. Um, but like what makes it different? What makes analyze different from 'evaluate: judge, assess, assign,' versus 'ascertain how those parts relate to one another,' 'attribute, compare, contrast?'"

The Problem of Pattern-Matching

Interviewees had differing opinions on the lists of verbs provided in the headings for each of the Cognitive Process Dimension levels. While interviewee M, who relied on pattern-matching in her sortings said, "I will say the verbs are really helpful. Just like a lot of verbs helped me think about items and because when I write questions, I have to think about all these words. So it made it a little bit easier to connect items to categories," Interviewee A cautioned against listing verbs to prevent users from engaging in pattern-matching, saying,

> "There is a, there's a huge misconception that just, when you use one of these words [in the headings,] it automatically keys to this cognitive engagement level. Right. So as soon as I say 'discriminate,' I'm evaluating. Right. And so I wish the tool wasn't as, as deceptively matchy, as it often is presented…And I don't know necessarily how to do that, but I guess that's what you're after. But there, there often is that, 'oh, I use that word. It must be this thing' rather than understanding what cognitive engagement actually even is."

Interviewee J agreed that verbs used in items do not correlate with just one cognitive process, saying, "I have definitely used those big, um, like tables of verbs, but what I noticed about those tables of verbs is that they do overlap, right? Like the 'distinguish' ends up being in all sorts of places because, um, we don't necessarily have words for these different things."

A related issue with the cognitive process headings was that of readability. Interviewees B, J, and L each had issues reading different headings. In his sorting of item Q7 in part two of his interview, interviewee L discussed his difficulty reading the heading for the cognitive process of understand, while interviewee J mentioned difficulty with the phrasing of the analyze cognitive process heading, saying,

> "The one place I had a little snag when I was reading this is the, in the analyze, um, description, it was just like kind of, I tripped over the sentence a little bit. "Students are asked comma in order to break

> material into its constituent parts and ascertain how those partss relate to each other" to do these verbs. And I guess I thought of analyzing, as "students are asked to break material into its constituent parts and ascertain how those parts relate to one another." Um, and so, yeah, there, there's something about that description that I had a hard time with the sentence there's, I mean, certainly like that's parallel to the way that the evaluate and create versions are written and they are shorter and I didn't have as much problem with that, but yeah, just the way that that analyze was described a little bit, I tripped over it."

Likewise, interviewee B expressed difficulty parsing the heading for the create cognitive process, saying,

> "Yeah. And you've got a lot of the verbs in here. Uh it was the last one. Students are asked in order to form a coherent or functional whole. And I wanted like a noun there, like 'a whole of what?' Like, what are we, yeah. What are we putting together? And my brain just put in system or molecule or, you know, I was like, 'okay, let's, let's make those, the words here.' And then I could use the rest of the, the rest of the, the, the, the part there. Right. Which is 'to construct design, hypothesize, et cetera.' But I needed that little noun there. And so I don't know if more clarifying words there might, might help. Like 'what is that functional coherent whole,' but again, that might be something that chemists are like, 'no, we already understand that. That's cool. We don't need anything else here.' You know, I'm not sure."

In a similar vein, interviewee I expressed that he didn't know what "carry out" or "operate" means in the context of the BCT.

The Problem of Chemistry-Specificity

While the BCT is a chemistry-specific translation of Bloom's taxonomy, interviewees C and J expressed a desire to have more chemistry-specific language in the BCT headings or in worked examples. Interviewee J compared the items sorted in the interview with the provided examples in the publication of the Blooming Biology Tool, saying,

> "Is there something more chemistry-specific? And so I think that, is there an interim, maybe this was just in the, sort of in the accompanying materials is a little bit of training about, um, what, what makes these, what makes these categories distinct? And examples really help. I remember reading about that with the Blooming Biology thing that they could, they, in whatever I was reading, it was like, 'here's an example of each of these levels.' Maybe it was applied at the same topic even, and that was really helpful for kind of understanding how those go together. So are there, I guess I would say, are there examples that are relevant to chemistry teaching? Um, you know, so that, that make, that would make this, um, sort of more accessible for, for practitioners. So something like, 'where does analyzing a spectrum go and how is that different from calculating a rate constant or, um, proposing a mechanism or proposing a method or designing an experiment or,' um, yeah, like some sort of those connections might be, might be helpful."

Interviewee C felt the headings of the Cognitive Process Dimension could be made more chemistry-specific, citing the generalist nature of most of the terms and the homonymity of words like "synthesize," saying,

> "Do you think if the wording was more chemistry specific, would it make a difference?..Like in the, in the, um, knowledge dimension, there is things, it says chemistry, right. Um, 'theories of chemistry.' Um, but on the, um, cognitive processes, those are just what people do. Right. But [these verbs] could apply to any, to any subject. Do you think? I, it was, I don't know if it was changed or, or it was, was more…Because 'synthesize' can mean different things, depending on your discipline, 'synthesizing literature,' you know, that kind of thing. Um, and people could think that's evaluating higher up on, uh, you know, that kind of thing looking through. Um, Hmm. But the words were helpful for me to try to figure out what, you know, what goes, where."

On the other hand, interviewees B and E pointed out terms used throughout the tool that they felt were appropriately specific. Interviewee B complimented the robustness of the definitions of the Knowledge Dimension, saying,

> "I'd looked at the knowledge dimension first and I thought that seemed um pretty reasonable. I like it. You include, like in the factual knowledge, you include examples, you know, fundamentals and 'what do we mean by fundamentals? The terms, technical definitions.' So the more you can robustly put in those um clarifying examples or general ideas, I think the more helpful that is for folks. And I like, you know, I think there's, it's a good job that, that has happened on that knowledge dimension."

Interviewee E noted that she liked some of the terms used in the Cognitive Process Dimension like "symbolize," but like interviewee C, discussed the homonymity of certain chemistry terms like "model," saying,

> "I don't have like a perfect memory of Bloom's taxonomy, like the original version, but like, I know what I, when I looked at glanced at [the BCT] briefly, I was like, I, I like to see 'symbolize' for example, because that's such an important part of chemistry. Like that's really important. But then in my mind, I was like, is that part, you know, I was like, I can't even remember what's on the original Bloom's taxonomy. So that's part of the problem, um, 'exemplify in order to draw meaning from,' um, that feels general, but there, there are some specific words to me that felt like, um, let's see what else there was, um, like 'model' I, model feels like it could be general or specific, cause I think that's one of those terms that can be, has multiple meanings in different settings. In fact, I have a collaboration with an analytical chemist and 'modeling' means something completely different to her in analytical chemistry than it does in education. Right. So it can have a very chemistry-specific feel to it. But it, you could also, like if I was a chemist reading this, I might also, like if I had an education background, I might think about that in a different way."

The Problem of Context

A fundamental consideration when employing Bloom's taxonomy is that of the context in which an item was presented to students. Most of the interviewees referenced this issue at some point, with interviewee G succinctly stating, "It depends on the problem and on the solver. So you have two factors, two two two protagonists here…one is the problem, and the other is the student."

Interviewees C, F, and K discussed the context of the item, with interviewee F remarking on how the sorting of an item changes with its placement in a term, saying, "I guess what what really strikes me and strikes me anytime that I've, I've worked with that cognitive aspect of Bloom's taxonomy is the, it's just how variable things can seem depending on what, what the context is. Even early to late in one of my, you know, one of my own classes or when I've tried to explain to someone else, but yeah, that's, I guess that's the part that I that I land on. Similarly, interviewee K came back to the conversation from his item sorting about the importance of knowing where in a unit an item was presented to determine its BCT intersection, saying,

> "When you're teaching, and initially, you know, you can get the concept across and you can try to get students to understand the concept. And that part for sure is conceptual knowledge. But then to me, it kind of, I don't want to say 'degrades,' you know, but, but then after that, I think it kind of essentially degrades and, and students don't think conceptually about it so much after that, like it just becomes, you

> know, like, ‘okay, gzzgzz’ kind of like that. Yes. But I don't know how to incorporate that into something like this, you know, like, I don't know if you'd have to sort of identify the context, you know, like when students are initially learning about this, dadada that I- I- I'm not sure.”

In a similar vein, Interviewee C discussed calibrating her sorting to consider the most complex cognitive process that a student might use, saying,

> “I didn't know that at the start, but thinking of what's the highest level that, that I think when you go to the, to the knowledge dimensions, you could have more than one. Right. And I think that's okay, but when it comes to the cognitive processes, I think, um, knowing that, ‘okay, what is the highest,’ if you want to pick the highest level that this question, um, could, um, get to, and then focusing on that, I think that would be real helpful. What would have been really helpful to me that they, okay, yeah that I'm just focused on ‘what is the highest level of, um, cognitive process that this question can I think that will be,’ you know, maybe make it easier to pick one because then I'm kind of going through like, ‘well, is it this? Or is it, but if, but if it reaches to analyze,’ then that will be the one that we can say, ‘this question is an analyze question.’”

Interviewees A and B both discussed the importance of having context of a learning environment when performing item sorting in research. Of his own early-career research, Interviewee A reflected that his coders did not have that context, saying,

> “My panel of [coders] had no idea what was being taught in the class that we had our research in, they were just given items and we did not take them through any, any extended interview protocols or, to get their ideas about things. They just, you know, labeled, right? And so how they did that? Mmm, don't know, but I did trust them. I did know that they knew what Bloom's taxonomy was, but they had no idea.

Interviewee B also discussed how not having the context of how an item was presented in class could affect research, saying, “One of our papers, we were looking at chem, chemistry or biochemistry and physics classes and, and it was not, originally we thought that our main team of researchers could be the bloomers and we quickly realized we couldn't be, like, I, I couldn't, I could do some of the biochemistry stuff, but not all of it because I don't have that background. Like you're saying, like, you know, some of these are, they're heuristics that are really important for students to be able to understand and apply. Right. And so, whereas I might see that as just straight up memorization, the chemist is like, ‘no, they have to know this and then do this,’ right. Or vice versa, whereas I'm thinking of it as this really big and intense thing, and the chemist is like, ‘hmm, no, not so much,’ you know, sort of thing. So there is that content knowledge that really comes into play when using, when using these types of tools…My upbringing with Bloom's is much more conservative, we tend to gr grade, categorize lower than upper. We, at least that's been my impression in talking with other people has been, they're like, ‘oh, this is an analysis question.’ And I'm like, ‘is that really like, think about this?’ Cause we're, we definitely try to think about context. ‘Where was this, was this question asked,’ right. And sometimes the question, you know, if it were asked in an upper-division class, I could totally see it being an analysis question because you can expect the students to be able to do more, to have more content knowledge. But sometimes the questions you know they're being asked in a first-year biology class, the student doesn't have enough background information to actually answer that question, which means they must have had something in the class, right, that they're, that they're, that they're pinging. Right. So it changes the question from being this higher level question to something much more rote and recallish…There was a couple of examples where I read it and I said, ‘I can't answer this question. There's no way I can answer this question right now. I needed to have had something that happened in lecture. That means this is a recall question. This is not a synthesis question at all.’ And once we started realizing that, like, that was like, a game-changer. And so that's part of why we didn't use the [Blooming Biology Tool] is because we were coming to this realization, that context really mattered and really changed things. And we weren't sure how that was reflected in the Biology in Bloom tool. That said to this day, I tell my colleagues go use that tool. It's really, user-friendly, it's really helpful. Right. It's, it's the thing that if it had been developed, you know, just one or two years earlier, we absolutely would have used. Right. But we were kind of developing our own way through using Bloom's taxonomy um a little bit separately. So yeah. So that's why we

> didn't use it in, in part is that we'd already started. And then we found this context effect that we thought was really important. And, and, and we were, I think a little bit more conservative in our blooming of things."

When it comes to the context of the student, interviewees B, J and M discussed the various assumptions researchers and practitioners make about how students are answering items. Interviewee M was most explicit in her distrust of sorting items along Bloom's taxonomy due to this disparity, saying,

> "So I would say that I still felt the same apprehension applying [the BCT] that I've felt with Bloom's generally, which is that, um, we are still, at best, making a prediction about what we think students are going to do to solve this problem. Granted, sometimes I feel more confident making that prediction, like in the [item Q8] case, I felt pretty confident thinking about what it was requiring the students, cause I've seen students engage in that kind of reasoning. Um, but I'm still left wishing I had, and this is going to be like the chem-edness inside of me, wishing that I had student, um, evidence that, w- w- well, one, there's quite a few things I would kind of want. One, that student remembering looks different than a student understanding versus applying. Those are actually corresponding to different cognitive processes, that students engage in. Um, and I would then like to see some evidence of matching, right, that what an instructor characterized as an understand question, evoked that cognitive process in students. Those were like the two it's- I'm just like desperate for student perspective data here, right. Cause there's just a- it's limited when it's only considering our perspective, even if that perspective is informed by what we've observed in our students, it's still limited without some evidence of- from students."

Interviewee J agreed that when building assessments for her own students she is mostly aware of what her students know, saying, "A question that, you know, basically could be a create question can be knocked all the way down to a knowledge question, if the, if the instructor has sort of already told you what the answer is. And so then that, um, that kind of changes that a little bit, but if I'm thinking about using my own teaching, then I know I know what I've told my students for the most part.” While an instructor might know what they told a student, Interviewee B discussed that each student has their own knowledge that they bring to an assessment, saying

> “Students all the time circumvent what our intentions are. Right. We intend this to be a synthesis question, but the students are super savvy and they know how to turn it, turn it. I mean, they're not, but they're just like, what have I memorized that I can then put plug into this equation or this, this problem or whatnot sort of a thing. And so they're, they're constantly circumventing our intentions and, and that's important as well.”

Interviewees C and H discussed how assumptions of student reasoning affect research outcomes, with interviewee C outlining a study on acid equilibrium concept inventorying, saying,

> “One of my students did a study on how students understand acid equilibrium across the curriculum from gen chem through biochem. And even when students answer the questions, we can tell students who are answering them because they just memorized thing, but they couldn't really explain the underlying factors related to the acid equilibrium question, or they can calculate, or they will say, ‘well, if you give me a calculation, I can do it,’ but they didn't have the, they didn't know what all of these things meant, but they know how to plug it into an equation. Um, so what level are they operating in to just plug it into the equation?...So what do you expect as an instructor versus what students actually do and how do you use that in assessment? Cause i f I, if I make a learning objective for my students based on, on this, then that's my expectations. When I assess them, I'm going to assess them at that level and they should be able to do this. I expect them to use a certain way in process of going through it.”

On the other hand, interviewee H cautions against researchers sorting items based on how the instructor intends for students to answer them, saying,

> "Depending on what story you're trying to tell, you could, it fits in a variety of places, which is good, but it also makes some, like, it almost ends up being to the researcher's advantage because they can describe something that, that across that range to fit wherever they want. Where students, again, I'm not sure it's clear from the student side that they would have that same understanding. Does that make sense?...How are they actually answering the question? And I think that a lot of times that's where I see colleagues struggle is that they're looking at it only from their side which would put it in a higher-order. Like for example, the last one [item Q8,] I could see somebody trying to classify that as conceptual understanding, right, because that's what they want, but uh students are gonna look at it and say, 'well, it's a buffer. So it's an acid and a base.' They're just remembering it. So I think that, that is one of the issues that I see, and I don't know how to, to approach that, but that would be one thing that in research I would, you would have to, or at least I would be looking at, 'are we looking at this problem from the student's perspective or my intent and is there a difference between those two,' which would be an interesting discussion, I think."

Of course, after participating in item sorting, an overarching concern for interviewees was how to address items having the possibility of existing at multiple BCT intersections based on context. Interviewees C, I and J approached this concern in different ways, with Interviewee C wanting the BCT to include guidance on whether learning objectives could cover multiple Knowledge Dimension categories, saying,

> "Because you can have stuff that's factual knowledge and, but also has conceptual and procedural, if you think being in the lab, all three might be at work, right. Um, so yeah, I think maybe having some guide to put people on how to, how to really deal with the, you know, maybe this ambiguity of like where these things fall. Like, is it okay if, if I have, um, if I make a learning objectives, and I say this is factual,

> conceptual and procedural, you know, um, those are the kinds of things that I can kind of get stuck on.

Similarly, interviewee J assumed based on her item sorting examples that the interconnectedness of an item was a positive quality, but wanted to know what to do with that information as an instructor, saying,

> “As an instructor, I guess the kind of, maybe last thing that I would want to sort of work through is, so given that, a lot of times, these, all of these things are really interconnected, right. And hard to just like in all of, in your examples, when I kind of coded it, um, often I was kind of left with, well, we're, we're kind of using a lot of things here. Um, how, yeah. So how, as an instructor, what am I supposed to do with the idea that my question is using multiple codes? Like how, how is that like, so like, assuming that that's a good thing that tells me more information about the sort of, uh, you know, assessments I'm giving my students, um, what am I supposed to make of that?”

Interviewee I was less convinced that multiple possible BCT intersections was sustainable, saying, “Everybody's going to want to go with ‘which intersection does it exist at?’…I'd bet money.”

CHAPTER FIVE: DISCUSSION

Conclusions

In this project, a taxonomy of college chemistry problems based on previous work in chemistry-specific cognitive skills was developed. The Blooming Chemistry Tool (BCT) was developed by combining the Expanded Framework for Analyzing General Chemistry Exams with Anderson and Krathwohl's Revised Taxonomy (RT), translating it to chemistry-specific cognitive processes and knowledge touchstones and modeling it after the accessible structure of the Blooming Biology Tool.

This project addressed the following two research questions:

1. What is the structure of, and what are the categories of skill necessary to construct a taxonomy of college chemistry problems?
2. How do members of the CER community with prior experience sorting items by cognitive level engage with such a taxonomy? What value do they anticipate this tool having in the CER community?

The next two sections will discuss the findings of these two research questions in light of the feedback the BCT received from interviewees.

Research Question 1

The BCT retained the blank, bidimensional matrix structure, as well as all levels of RT knowledge and cognitive skills with level headings written to prioritize chemistry-specific skills and types of knowledge. Interviewees found benefits and drawbacks to each of these decisions.

BCT Construction

Interviewees opposed an assumption of hierarchy in the BCT

As discussed in the literature review chapter, Zoller and colleagues' investigative work into the nature of Higher-Order Cognitive Skills (HOCS) in college chemistry (Zoller, et all., 1995, Zoller & Tsaparlis, 1997; Zoller, 2001; Zoller et al., 2002, Zoller & Tsaparlis, 2003; Zoller & Scholz, 2004; Zoller, 2011) has been highly influential in the CER community. When Zoller and Tsaparlis rejected the Original Taxonomy based on the lack of cumulative hierarchy they found, much of the CER community followed. As evidenced by several interviewees, many members of the CER community are unfamiliar with Anderson and Krathwohl's revision of the OT that addressed most of its major criticisms. namely that of an assumed cumulative hierarchy. Several interviewees thought that the RT, and by extension the BCT, was still assumed to have the same cumulative hierarchy as its predecessor. Interviewees I and M explicitly objected to the idea of a cumulative hierarchy, while Interviewees B and E appreciated the assumption of cumulative hierarchy in the Cognitive Process Dimension, though Interviewee B qualified her assumption by expressing that she is comfortable with items not fitting cleanly into those bounds, while

Interviewees E expressed some discomfort in the idea of a cumulative hierarchy in the Knowledge Dimension of the BCT.

Other interviewees knew that there was no longer an assumption of cumulative hierarchy in the revised taxonomy, but objected to any assumption of hierarchy in sorting cognitive items. This objection seemed to come from interviewees conflating cognitive skill with difficulty. Both interviewees A and E mention having used Bloom's taxonomy as a tool to articulate item difficulty, and interviewee J mentioned hesitance in coding more difficult questions as "lower-order." However, this is problematic because while some higher-Blooms skills (e.g., create) might reliably be more "difficult" than some lower-Blooms skills (e.g., remember), there are many instances where this is not necessarily the case. For example, it would be quite simple to write an "analyze" item that is harder or easier than a particular "apply" item. This highlights a need in the literature for a tool that more directly targets difficulty, or even potentially a need to triangulate difficulty through multiple measures. Certainly it suggests that any instructions or training provided with the BCT include explicit discussion of how cognitive process levels are not equivalent to difficulty levels.

Interviewees were cautious about the high granularity of the BCT

An aspect of the BCT that was new to many interviewees was the Knowledge Dimension. Interviewees A, J, and G compared the BCT favorably to the OT, expressing their frustration with the unidimensional pyramid presentation of the OT. Interviewee G and J even suggested adding additional dimensions or dynamics to the BCT, such as aligning it with stages of cognitive development, a discussion of algorithmic versus conceptual items, or an index of complexity. Other interviewees argued that the appeal of Bloom's taxonomy is its ubiquity and recognizability, and that adding elements to it reduces its usefulness, such as Interviewee B, who discussed that in her previous research, the use of the RT over-complicated the tool for their faculty.

Related to the matter of bidimensionality, a major theme in interviewee suggestions for using a variant of Bloom's taxonomy was to reduce granularity of the Cognitive Process Dimension for ease of use and to align it more with other learning taxonomies such as Marzano's taxonomy, or the SOLO taxonomy. Interviewees A, C, J and N all compared Bloom's taxonomy unfavorably to Marzano's taxonomy regarding the number of levels in the Cognitive Process Dimension. Similarly, interviewee J thought the Knowledge Dimension would be easier to sort along for undergraduate students because there were less possibilities. Some interviewees aligned more with Zoller and colleagues in their move towards simply separating the Cognitive Process Dimension into HOCS and LOCS, particularly with regards to practitioner use. Interviewees A and E agreed with Interviewee B's caution about bidimensionality by discussing how a higher level of granularity in the number of levels of the Cognitive Process Dimension were a drawback of Bloom's especially for practitioners less familiar with CER or the Science of Teaching and Learning, and new instructors who want to teach well, but have very little time. However, not all interviewees were sold on consolidating the Cognitive Process Dimension. Interviewee H and L both liked the six levels, though interviewee L advocated for better stratification between levels, which will be discussed more in the next section.

BCT Content

Interviewees advocated for more chemistry-specificity in the BCT

On the other side of the issue of granularity is that many interviewees have chemistry-specific definitions of several cognitive processes, and expressed a desire for a learning taxonomy to encompass that discipline-specific homonymity. The most frequent example of this was interviewees interpreting "explain" in multiple ways. In particular, there seems to be a larger "Explain" where interviewees say that the explanation must be according to the scientific method in some way (e.g., how does one prove a fact is true, using the scientific method where there is data being employed to build an argument), as opposed to a smaller "explain" where students can justify their answer using established facts (e.g., "This is true because the molecule is polar.") Because of this, Interviewees A, G, I, and J expressed discomfort either with the verb “explain” being under the heading of the understand cognitive process, or with the idea of “understand” being perceived as lower on the Cognitive Process Dimension. Interviewee A, C, and L expressed a similar reluctance about including the term “conclude” in the understand cognitive process heading. Interviewees A and E both discussed the relative complexity of the term “modeling” and how it tied the cognitive process of understand to that of the cognitive process of create. However, this overlap between the idea of true understanding requiring synthesizing information in order to create models from which one can draw conclusions might explain why many interviewees seemed to discount the cognitive process of create automatically in their sortings. Of course, this may be a result of the nature of the questions chosen for sorting, but there may also be an assumption that “create” must mean to create brand-new knowledge (e.g., research) whereas the knowledge a student creates doesn’t have to be new to the world, simply new to the *student*, to activate this category.

Another example of this confluence came in the form of the interface between the cognitive processes of apply and analyze and how they overlap with procedural knowledge. Interviewees A and J specifically spoke to this overlap, while several more interviewees inquired about the difference in terms during their sortings, but most interviewees reflexively connected procedural knowledge with laboratory and experimental methods, procedures, and techniques. While many laboratory methods may activate procedural knowledge, it's possible not all laboratory methods do, and many problems that are not lab-based in general can activate procedural knowledge (e.g., using a memorized procedure for a calculation). Picking up on this, several interviewees expressed uncertainty whether "apply" is meant to extend to applying a memorized heuristic. For example, interviewee A discussed how faculty might question whether item Q1 was a memorization question (because the students memorized a periodic trend) versus an "apply," because the students used that memorized trend to rank the elements? Because of this discipline specific-homonymity interviewees C, E and J expressed a desire to have less generalist language in the BCT headings or in worked examples in favor of more chemistry-specific language, with Interviewee J explicitly suggesting the types of discipline-specific examples provided in the Blooming Biology Tool.

However, the push for more chemistry-specific language in order to include the discipline-specific homonymity of some terms and to better differentiate the difference between adjacent categories such as the apply/analyze interface runs directly counter to the discussion of whether six Cognitive Process Dimension categories provide too much granularity. In college chemistry, especially because of the tension of providing balance in conceptual and algorithmic items, many assigned questions fall in the range of understand/apply/analyze, and being able to better explicate those differences is a motivation to retain the six cognitive process dimension categories of the RT.

Interviewees emphasized the need for accessibility in the BCT

Since many interviewees discussed the potential of the BCT to serve as an instructor-training tool, a high priority of interviewees was that the tool be as accessible as possible. There were two ways in which interviewees addressed accessibility. The first was the matter of readability. Interviewees B, J, and L each had issues reading different headings based on either wordiness or cadence. Interviewees I and J both expressed difficulty with the number of verbs used in each cognitive process dimension heading–Interviewee J because the verbs are not mutually exclusive to one category, and Interviewee I because of the vagueness of certain verbs. Interviewee A specifically cautioned against making long lists of verbs to prevent users from engaging in pattern-matching in the way Interviewee M did across her items. Based on this feedback and the matter of chemistry-specificity, future iterations of the BCT would be recommended to use less, but more explicitly discipline-specific verbs in the Cognitive Process Dimension headings.

The second way in which interviewees addressed accessibility was through the efficiency of use of the BCT–whether any of the included elements were superfluous. For example, some interviewees were perplexed by the inclusion of the metacognitive knowledge level in the Knowledge Dimension, either because they weren't sure how it could be used in assessment, or because they thought it didn't quite mesh with the other levels of the Knowledge Dimension. Interviewees E and F felt that metacognitive knowledge was slightly removed from the rest of the Knowledge Dimension, while interviewee H was glad of the physical separation of the metacognitive knowledge level, being at the bottom of the knowledge dimension, since he worried practitioners would not know how to use it. As an example of interviewee H's worry, interviewees C, L, and M each had questions about what metacognitive items might look like in college chemistry, while interviewees B and M discussed whether all BCT intersections have practical use in the college chemistry classroom. In the construction of the BCT, it was decided to retain all levels of the RT knowledge and cognitive skills, even though college chemistry classrooms may not regularly conduct assessments at all those levels so that members of the CER community could determine the boundaries of their own assessments, a topic which will be discussed in more detail in the next section.

Research Question 2

In the tool-testing interviews, participants were observed in how they engaged with the BCT as they sorted a selection of items and then were asked how they saw the BCT being used by various members of the CER community. In this section, reasons for sorting variance among participants is discussed, as well as how interviewees anticipated the BCT fitting into the CER landscape.

BCT Engagement

In the previous section, some reasons for rater inconsistency among the interviewees while sorting items along the BCT have already been addressed, such as interviewees having different interpretations of the terms within the BCT (e.g., ‘explain) and interviewees defaulting back to prior experiences with the OT or RT (rather than using the BCT as it is designed.) This falling-back on prior experience seems to be related to the interviewees' profiles, discussed in further detail later in the results chapter. While these interviewee profiles are not mutually exclusive, they were useful for contextualizing interviewee sortings. For example, interviewees who fit into the Bloom's Intesivist profile or the Prefers-Another-Taxonomy profile were more likely to have preconceived notions of how to employ Bloom's taxonomy, and therefore had to integrate the new aspects of the BCT into their understanding of sorting items by cognitive level, while interviewees in the Bloom's After-Thoughtist and Taxonomy Agnostic profiles were much more flexible in applying Bloom's Taxonomy and therefore were more open to changes in the tool's construction and content. That being said, the most common reason for rater inconsistency among the interviewees while sorting items along the BCT was a matter of context. The more assumptions that an interviewee had to make about the context of the item (how it might've been presented in class, how the student would be expected to answer it, what level of mastery the instructor might be looking for,) the more variance present in the way that item was sorted along the BCT.

Interviewee Assumptions of Learning Context

A fundamental consideration when employing Bloom's taxonomy is that of the context in which an item was presented to students. This is a recognized issue in research using learning taxonomies, and both interviewees A and B discussed the impracticality of coding items without knowing how they were presented in class. However, the learning context in which an item was presented to students encompasses more than the particular textbook or if an instructor explicitly told students a particular thing; it also includes at what point in time the item is being given to students (e.g., at the start of a unit vs the end of a unit). Interviewees F, J, K, and M all remarked on how the sorting of an item changes with its placement in a term. This fact was the crux of interviewee K's preferred item-sorting dynamic of exercises versus problems, which he used to contextualize his sorting on multiple items–whether the item was calling for more rote skills or crystalized knowledge or whether students were still learning how to accomplish a given task. While interviewee K used this dynamic across multiple cognitive processes, the matter of novelty was discussed by several interviewees as a way to specifically distinguish between the cognitive processes of apply and analyze in a college chemistry classroom. Apply and analyze were definitely considered by interviewees to have more overlap than many of the other cognitive process dimensions, but having a community-defined way of separating them gives resolution power to the BCT, especially in chemistry-specific contexts such as whether an item was asked in lecture or in lab.

Interviewee Assumptions about Student Reasoning

While an instructor might know what they told their class, interviewees recognize that each student has their own knowledge that they bring to an assessment. Interviewees tended to sort items along the BCT according to two assumptions of student reasoning, irrespective of interviewee profile: some interviewees used the BCT to categorize the "ideal path" of things a student should know to get the correct answer for the "right" reason (e.g., in a ranking electronegativity question, that students would understand the concept of electronegativity, therefore it's "understand,") whereas others categorized based on the "minimum set" of things a student needs to know to get the answer right (e.g., a rank electronegativity question could be correctly answered without without any conceptual understanding necessary, as long as students have memorized the periodic trend, therefore it's "remember.") This leads to very different rankings for the same item.

However, it is worth considering whether the "ideal path" to which some interviewees default actually refers to the skills they as a subject-matter expert might use to answer an item, versus the skills they might expect students to use, and where those two expectations might diverge. Interviewees B, J, and M each discussed the various assumptions researchers and practitioners make about how students are answering items though Interviewee M was most explicit in her distrust of sorting items along Bloom's taxonomy due to this disparity. The BCT was designed with a "highest takes all" approach in mind, and CER members desiring a more fine-grained examination of an individual question might instead pursue a different tool, such as the 3D-Leaning Assessment Protocol.

Of course, after participating in item sorting, an overarching concern of interviewees was how to address items having the possibility of existing at multiple BCT intersections based on context. Some interviewees, such as interviewees C and J wanted guidance on what to do with that information as an instructor, either as part of the BCT or its training, while Interviewee I suggested the BCT would not be usable if items did not have clear-cut intersections at which they existed, and interviewee L worried that if items existed at multiple intersections that practitioners would ignore lower, more fundamental intersections in favor of promoting “higher order” skills. However, based on the interviews from this project, assessment items do not have an inherent sorting–it really is all about context–which runs counter to previous forays into sorting items by cognitive level in the research literature that have presented rankings for items along Bloom’s taxonomy as if they were inherent to the item itself and not based on the context in which it was asked. This is a substantial research finding that has some implications for how the BCT might fit into the greater CER landscape.

BCT Value

Interviewees were asked if they saw a need for a tool like the BCT in the chemistry education research community, and what uses they saw in that landscape for such a tool. Interviewee suggestions for uses of a discipline-specific item sorting tool such as the BCT spanned all components of the CER community–there were proposed uses for researchers, practitioners, and students, though there was the most consensus that due to its accessibility and ubiquity, a discipline-specific approach to Bloom’s taxonomy had the most potential as a tool for instructor training.

Implications for Practitioners

Fundamentally, the power of a discipline-specific Bloom’s taxonomy translation as an instructor-training tool is that it provides a structure for people to think about their teaching. While any tool that provides that structure is beneficial, Bloom's taxonomy in particular seems to be something that is more accessible because interviewees said that new teachers find it intuitive, and because many people are often at least a little familiar with it already. Several interviewees discussed the way that either they already use Bloom's taxonomy in their teaching–both Interviewee C and J use Bloom's taxonomy to align their course content–or proposed ways they could see practitioners using Bloom's taxonomy in their classrooms, such as the classical use of aligning exams with learning objectives, or for more intentionally selecting homework problems for students, or for making larger CER studies more accessible for practitioners.

Other interviewees talked about the advantage of Bloom's taxonomy as a teaching training tool. Interviewees B and C both discussed how tools like Bloom's taxonomy makes faculty think differently about their assessments, and Bloom's taxonomy can help new instructors *develop things* such as assessment items, learning goals, and assessments aligned with learning goals, as well as to think about what one assesses, how to assess a variety of skills, and what the point of the course/chapter/lesson is. Interviewee C noted that Bloom's taxonomy also provides a way to help new instructors *communicate things to their students*, such as learning expectations and goals, and Interviewees B, C, and E discussed the way that the BCT could help new instructors perform metacognitive exercises with their students to guide the students self-assessment of what types of things they are learning and studying. Neither of these skills involve training new instructors the "right" sorting for particular types of chemistry items. Rather, the point is for new instructors to consciously engage with their teaching and assessment and develop an internal understanding of the sorting cognitive items looks like in their classroom.

Virtually all interviewees saw value in a Bloom's-like resource as a tool for *people who are starting to teach*. Across interviewees, there was the most consensus that the BCT would be useful for instructor-training purposes. Interviewees A, B, C, E, F, and I all discussed how they saw Bloom's taxonomy fitting into instructor-trainings. Some interviewees, such as Interviewees A, B, and C, already use some version of it for this purpose, contributing to the fact that regardless of their interviewee profile, the OT was usually the first learning taxonomy with which interviewees became familiar. Interviewees E and F agreed that Bloom's taxonomy is a useful entry point for new faculty and those who find that they need to make changes to their teaching, (for example, in light of the COVID-19 pandemic.) In either case, the advantage of a discipline-specific Blooming tool is immediately evident: because of how generalist Bloom's taxonomy is, it takes time and work to think of how to apply it to a particular disciplinary context especially when considering language that can be semantically confusing to chemists, (e.g., synthesis, explain) and, as interviewee A said, “new faculty are incredibly overwhelmed,” to say nothing of the practitioners faced with a roughly overnight transition to virtual learning in light of the COVID-19 pandemic. In addition to the BCT reducing the time and effort needed to engage with learning taxonomies as a beginning teacher, a tool that supplants the unidimensional OT in the CER community canon is especially valuable, as all interviewees discussed being familiar primarily or exclusively with the unidimensional, “pyramid” presentation of Bloom’s taxonomy.

Implications for Researchers

The OT and the RT were intended to be used to assess assessments and student responses to assessments and to develop learning outcomes, and while an individual practitioner can be trained to use the BCT to develop or improve a course by iteratively engaging in a cycle of (1) developing learning objectives, (2) using those LO to develop their assessments, and then (3) evaluating what level their students are engaging at on those assessments, having evidence that items do not have an “inherent” sorting also means the BCT could better establish another use for learning taxonomies: to structure the context necessary to compare classroom environments. While a few interviewees (A, C, E, F, and L) did propose the kinds of research questions they thought a tool such as the BCT could help answer, the best research use interviewees saw for the BCT was as a way to communicate and standardize their findings for other community members. For example, a hypothetical instructor that uses the BCT to structure her learning outcomes and assessments, could use the BCT to report the intersection at which an assessment item was aimed, as well as the intersection at which she taught the material, and the intersection at which her students responded. In this example, the BCT serves as both a framework upon which to hang her claims, as well as a lens through which other CER community members can view her reported course outcomes. As evidence-based teaching practices and active learning strategies become more prevalent in college chemistry classrooms, the ability to communicate course outcomes at a level of resolution higher than final course grade becomes more imperative. Consistent use of a tool such as the BCT could remove the burden for each CER researcher to build the language to communicate their findings for each project, similar to how the Blooming Biology Tool has affected the Biology Education Research community.

Indeed, the BCT was intended to be useful to the CER community in much the same way that the BBT has been useful to the BER community, and while it was clear that all interviewees saw value for more consistent use of *a* learning taxonomy in CER, it's possible that a discipline-specific translation of Bloom's taxonomy may not meet that need for the CER community. As stated in the last section, Bloom's taxonomy is widely known–

regardless of their interviewee profile, the OT was usually the first learning taxonomy with which interviewees became familiar. However for researchers, that early exposure was not necessarily considered by all interviewees to be a strength, as often interviewees "outgrew" the OT as they progressed in their careers. While Bloom's taxonomy's quality of "brand recognition" was considered to be a positive attribute when selecting a base learning taxonomy for this project, that also means there would be "bad press" surrounding the still-widely-held criticisms of the OT (regarding its unidimensionality, assumption of cumulative hierarchy, and lack of theoretical basis/validity as a taxonomy,) with which to engage when convincing CER community members that the BCT is both a useful and an acceptable tool. This was evidenced by resistance to accept changes to Bloom's taxonomy from interviewees who were either invested in an older version of the taxonomy or else invested in rejecting Bloom's taxonomy in favor of a more preferred dynamic. Beyond that, there are those in the CER community for whom a less-generalizable, more context-dependent approach to item sorting (such as the one outlined in this project) may never sit well. Of course, a heavy reliance on learning context may be an intrinsic, unavoidable issue for any taxonomy, and not just the BCT. As the community decides which kind of taxonomy around which they want to base such a needed common tool, Interviewee G cautions that,

> “What happens in these social science, sciences field is that, uh, many researchers want to invent their own, uh, models, their own theories, uh, even, uh, theories with an acronym. And quite often, uh, they want to override previous theories and replace them with the, with the new ones. But, uh, I think the better, the better, the best approach would be to try to combine and to explore it, most of them. It's good to go from Bloom's to revise[d] Bloom's. This is good and standard, in the pure science. But what I find is bad sometimes with guys is that they say, ‘well, this is useless. This is uh rubbish,’ you know…I think we must try to, to explore it, most of the contributions and in that way, uh, everyone of us is going to be useful with his work for future. Because if you have people who try to reject or find your work, which is wrong, of course, science proceeds and develops and progresses, this is normal and standard. Uh, what his bad is trying, you know, to reject all the things and try to do new, uh, new things.”

Limitations

The sample size for this project was small, consisting of only 15 interviewees. Furthermore, the sample of interviewees for this project consisted of a narrow subset of interviewees–each had to be (a) faculty or staff at an institution of higher learning, (b) in possession of a PhD in their field, and (c) have experience sorting items by cognitive level. In fact, two of the interviewees were members of the BER community with experience using the Blooming Biology Tool. The sortings and perspectives that this sample of interviewees gave will by no means be representative of the CER (or BER) communities.

Future Research

Depending how the CER community wishes to employ the BCT, there are a couple of paths that future researchers could pursue. If future researchers are interested in refining the BCT to improve accessibility for more members of the CER community, then they may engage in interview studies or literature reviews to determine more chemistry-specific language (in headings, examples, or overlaid layers) for the tool. If future researchers are interested in establishing evidence of validity and reliability for the BCT in order to use it across multiple course environments, then they may acquire a more representative sample of the CER community (and therefore better heatmaps of a given items sortings,) through a larger quantitative study where many community members, regardless of experience sorting items by cognitive level sort chemistry items along the BCT. If future researchers are interested in creating more in-depth profiles of courses, then they might sort the content from a given course and supplement that with interviews with the instructor or student to contextualize the levels at which the course was taught and understood.

REFERENCES

Allen, D., & Tanner, K. (2002). Approaches to Cell Biology Teaching: Questions about Questions. *Cell Biology Education*, *1*(3), 63–67. https://doi.org/10.1187/cbe.02-07-0021

Anderson, L. W., Krathwohl, D. R. D. R., & Bloom, B. S. (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. *New York Longman*, 302. https://doi.org/10.1207/s15430421tip4104_2

Arneson, J. B., & Offerdahl, E. G. (2018). Visual Literacy in Bloom: Using Bloom's Taxonomy to Support Visual Learning Skills. *Cell Biology Education*, *17*(1), ar7. https://doi.org/10.1187/cbe.17-08-0178

Bailin, S. (2002). Critical Thinking and Science Education. *Science & Education*, *11*(4), 361–375. https://doi.org/10.1023/A:1016042608621

Bergendahl, C., & Tibell, L. (2005). Boosting Complex Learning by Strategic Assessment and Course Design. *Journal of Chemical Education*, *82*(4), 645.

Bloom, B. S. (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals, Book 1: Cognitive Domain*. *Taxonomy of Educational Objectives*.

Booker, M. J. (2007). A Roof without Walls: Benjamin Bloom's Taxonomy and the Misdirection of American Education. *Academic Questions*, *20*(4), 347–355. https://doi.org/10.1007/s12129-007-9031-9

Bruck, A. D., & Towns, M. H. (2009). Analysis of classroom response system questions via four lenses in a General Chemistry course. *Chem. Educ. Res. Pract.*, *10*(4), 291–295. https://doi.org/10.1039/B920834H

Crowe, A., Dirks, C., & Wenderoth, M. P. (2008). Biology in Bloom: Implementing Bloom's Taxonomy to Enhance Student Learning in Biology. *Cell Biology Education*, *7*(4), 368–381. https://doi.org/10.1187/cbe.08-05-0024

Crowe, A., Dirks, C., & Wenderoth, M. P. (2009). Biology in Bloom: Implementing Bloom's Taxonomy to Enhance Student Learning in Biology Alison. *CBE Life Sciences Education*, *8*(1), 1–6. https://doi.org/10.1187/cbe.08

Dávila, K., & Talanquer, V. (2010). Classifying end-of-chapter questions and problems for selected general chemistry textbooks used in the United States. *Journal of Chemical Education*, *87*(1), 97–101. https://doi.org/10.1021/ed8000232

DeRuisseau, L. R. (2016). The flipped classroom allows for more class time devoted to critical thinking. *Advances in Physiology Education*, *40*(4), 522–528. https://doi.org/10.1152/advan.00033.2016

Edwards, N. (2010). An analysis of the alignment of the Grade 12 Physical Sciences examination and the core curriculum in South Africa. *South African Journal of Education*, *30*(4), 571–590. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-79951570483&partnerID=40&md5=4c89b4c8f80a76af90c8591ff1ea61b6

Elliott, S. L. (2010). Efficacy of Role Play in Concert with Lecture to Enhance Student Learning of Immunology. *Journal of Microbiology & Biology Education*, *11*(2), 113–118. https://doi.org/10.1128/jmbe.v11i2.211

Flinders, D. J., Anderson, L. W., & Sosniak, L. A. (1996). Bloom's Taxonomy: A Forty-Year Retrospective. *History of Education Quarterly*, *36*(1), 76. https://doi.org/10.2307/369314

Gillette, G., & Sanger, M. J. (2014). Analysing the distribution of questions in the gas law chapters of secondary and introductory college chemistry textbooks from the United States. *Chem. Educ. Res. Pract.*, *15*(4), 787–799. https://doi.org/10.1039/C4RP00115J

Harrow, A. (1972). *A Taxonomy of The Psychomotor Domain: A Guide for Developing Behavior. New york*. https://doi.org/10.2307/1161665

Holme, T. A., & Murphy, K. L. (2011). Assessing conceptual and algorithmic knowledge in general chemistry with ACS exams. *Journal of Chemical Education*, *88*(9), 1217–1222. https://doi.org/10.1021/ed100106k

Hsieh, H.-F., & Shannon, S. E. (2005). Three Approaches to Qualitative Content Analysis. Qualitative Health Research, 15(9), 1277–1288. https://doi.org/10.1177/1049732305276687

Kalas, P., O'Neill, A., Pollock, C., & Birol, G. (2013). Development of a meiosis concept inventory. *CBE Life Sciences Education*, *12*(4), 655–664. https://doi.org/10.1187/cbe.12-10-0174

Karamustafaoğlu, S., Sevim, S., Karamustafaoğlu, O., & Çepni, S. (2003). Analysis of Turkish High-School Chemistry-Examination Questions According To Bloom'S Taxonomy. *Chem. Educ. Res. Pract.*, *4*(1), 25–30. https://doi.org/10.1039/B2RP90034C

Kayima, F. (2018). *Exploring chemistry teachers ' perspectives on questioning and providing a new way of analyzing teacher questions in science classrooms*.

Krathwohl, D., Bloom, B., & Masia, B. (1964). *Taxonomy of educational objectives: The classification of educational goals - Handbook II: Affective Domain. David McKay Company*. Retrieved from http://scholar.google.co.uk/scholar?start=0&q=bloom+krathwohl&hl=en&lr=lang_en&as_sdt=0,5#8

Kreitzer, A. E., & Madaus, G. F. (1994). Empirical Investigations of the Hierarchical Structure of the Taxonomy. In *Bloom's Taxonomy: A Forty-year Retrospective* (pp. 64–81).

Morshead, R. W. (1965). Taxonomy of Educational Objectives Handbook II: Affective Domain. *Studies in Philosophy and Education*. https://doi.org/10.1007/BF00373956

Newton, G., & Martin, E. (2013). Blooming, SOLO Taxonomy, and

Phenomenography as Assessment Strategies in Undergraduate Science Education. *Journal of College Science Teaching*, *43*(2), 78–90. Retrieved from www.jstor.org/stable/43631075

Niaz, M., & Robinson, W. R. (1991). No Title. In *Teaching algorithmic problem solving or conceptual understanding: Role of developmental level, mental capacity, and cognitive style* (pp. 1–32). Lake Geneva, Wisconsin.

Nurrenbern, S. C., & Pickering, M. (1987). Concept learning versus problem solving: Is there a difference? *Journal of Chemical Education*, *64*(6), 508. https://doi.org/10.1021/ed064p508

O'Neill, A., Birol, G., & Pollock, C. (2010). A Report on the Implementation of the Blooming Biology Tool: Aligning Course Learning Outcomes with Assessments and Promoting Consistency in a Large Multi-Section First-Year Biology Course. *The Canadian Journal for the Scholarship of Teaching and Learning*, *1*(1), 8. https://doi.org/10.5206/cjsotl-rcacea.2010.1.8

Pappa, E. T., & Tsaparlis, G. (2011). Evaluation of questions in general chemistry textbooks according to the form of the questions and the Question-Answer Relationship (QAR): the case of intra- and intermolecular chemical bonding. *Chem. Educ. Res. Pract.*, *12*(2), 262–270. https://doi.org/10.1039/C1RP90031E

Ramirez, R. P. B., & Ganaden, M. S. (2008). Creative activities and students' higher order thinking skills. *Education Quarterly*, *66*(December), 22–33.

Rodrigues, S., Taylor, N., Cameron, M., Syme-Smith, L., & Fortuna, C. (2010). Questioning Chemistry: The Role of Level, Familiarity, Language and Taxonomy. *Science Education International*, *21*(1), 31–46. Retrieved from http://www.icaseonline.net/sei/march2010/p3.pdf

Saldana, J. (2015). The Coding Manual for Qualitative Researchers. upon Tyne: Sage.

Sanabria-Ríos, D., & Bretz, S. L. (2010). Investigating the relationship between faculty cognitive expectations about learning chemistry and the construction of exam questions. *Chemistry Education Research and Practice*, *11*(3), 212. https://doi.org/10.1039/c005470b

Simpson, E. J. (1972). The Classification of Educational Objectives in the Psychomotor Domain. *Education*, *3*(3), 43–56. Retrieved from http://eric.ed.gov/ERICWebPortal/recordDetail?accno=ED010368

Smith, K. C., Nakhleh, M. B., & Bretz, S. L. (2010). An expanded framework for analyzing general chemistry exams. *Chem. Educ. Res. Pract.*, *11*(3), 147–153. https://doi.org/10.1039/C005463C

Soozandehfar, S. M. A., & Adeli, M. R. (2016). A Critical Appraisal of Bloom's Taxonomy. *American Research Journal of English and Literature*, *2*, 1–9.

Stamovlasis, D., Tsaparlis, G., Kamilatos, C., Papaoikonomou, D., &

Zarotiadou, E. (2005). Conceptual understanding versus algorithmic problem solving: Further evidence from a national chemistry examination. *Chemistry Education Research and Practice*, *6*(2), 104–118. https://doi.org/10.1039/B2RP90001G

Terry, B. D. R., Lemons, P., Armstrong, N., Brickman, P., Ribbens, E., & Herreid, C. F. (2012). Eight Is Not Enough : The Level of, 2001.

Thompson, A. R., & O'Loughlin, V. D. (2015). The Blooming Anatomy Tool (BAT): A discipline-specific rubric for utilizing Bloom's taxonomy in the design and evaluation of assessments in the anatomical sciences. *Anatomical Sciences Education*, *8*(6), 493–501. https://doi.org/10.1002/ase.1507

Tikkanen, G., & Aksela, M. (2012). Analysis of Finnish chemistry Matriculation Examination questions according to Cognitive Complexity, *8*(3), 258–268.

Tsaparlis, G., & Zoller, U. (2003). Evaluation of higher vs. lower-order cognitive skills-type examinations in chemistry: implications for university in-class assessment and examinations. *University Chemistry Education*, *7*(2), 50–57.

Upahi, J. E., Issa, G. B., & Oyelekan, O. S. (2015). Analysis of senior school certificate examination chemistry questions for higher-order cognitive skills. *Cypriot Journal of Educational Sciences*, *10*(3), 218–227.

Upahi, J. E., & Jimoh, M. (2016). Classification of end-of-chapter questions in senior school chemistry textbooks used in Nigeria. *European Journal of Science and Mathematics Education*, *4*(1), 90–102.

Zaidi, N. B., Hwang, C., Scott, S., Stallard, S., Purkiss, J., & Hortsch, M. (2017). Climbing Bloom's taxonomy pyramid: Lessons from a graduate histology course. *Anatomical Sciences Education*, *10*(5), 456–464. https://doi.org/10.1002/ase.1685

Zheng, A. (2008). Application of Bloom's taxonomy debunks the "MCAT myth"'. *Science- …*, *319*(January), 2–3. https://doi.org/10.1126/science.1147852

Zoller, U. (1993). Are lecture and learning compatible? Maybe for LOCS: Unlikely for HOCS. *Journal of Chemical Education*, *70*(3), 195. https://doi.org/10.1021/ed070p195

Zoller, U. (2001). Alternative Assessment As (Critical) Means of Facilitating HOCS-Promoting Teaching and Learning in Chemistry Education. *Chemistry Education: Research and Practice in Europe*, *2*(1), 9–17.

Zoller, U. (2011). Science and Technology Education in the STES Context in Primary Schools: What Should It Science and Technology Education in the STES Context in Primary Schools: What Should It Take? *Journal of Science Education and Technology*, *20*(5), 444–453. https://doi.org/10.1007/S10956-01

Zoller, U., Dori, Y. J., & Lubezky, A. (2002). Algorithmic, LOCS and HOCS (chemistry) exam questions: Performance and attitudes of college students. *International Journal of Science Education*, *24*(2), 185–203. https://doi.org/10.1080/09500690110049060

Zoller, U., Lubezky, A., Nakhleh, M. B., Tessier, B., & Dori, Y. J. (1995). Success on Algorithmic and LOCS vs. Conceptual Chemistry Exam Questions. *Journal of Chemical Education*, *72*(11), 987. https://doi.org/10.1021/ed072p987

Zoller, U., & Scholz, R. W. (2004). The HOCS paradigm shift from disciplinary knowledge (LOCS) - To interdisciplinary evaluate, system thinking (HOCS): What should it take science-technology-environment-society oriented courses, curricula and assessment? *Water Science and Technology*, *49*(8), 27–36. https://doi.org/WOS:000222193400005

Zoller, U., & Tsaparlis, G. (1997). Higher and lower-order cognitive skills: The case of chemistry. *Research in Science Education*, *27*(1), 117–130. https://doi.org/10.1007/BF02463036

APPENDIX: OTHER ITEM SORTINGS

This appendix consists of a description of the coding process for the other eight items sorted by interviewees as part of the pilot testing interviews.

Item Q2

Item Q2 was sorted by the two interviewees G and J. Item Q2 was the last question sorted by both interviewees.

Interviewee G

Interviewee G began his sorting of item Q2 along the Cognitive Process Dimension, saying,

> "Well it surely…the cognitive process is remembering. Okay. And the knowledge dimension, again, it's factual knowledge, but, uh, of course students usually have done a lot of, uh, practice with the ideal gas equation. So, uh, they must be quite familiar with it."

Because Interviewee G concluded that item Q2 should be sorted as existing at the BCT intersection of Remember Factual Knowledge, this excerpt was assigned the *a priori* codes of Remember and Factual Knowledge using directed content analysis. Interviewee G made this determination based on the familiarity he assumed students would have with this type of item (lines 2-4; coded as Assumptions of Student Reasoning or Learning Context.)

Interviewee G then considered the phrasing of the question (lines 5-14; coded as both Item Forensics and Complexity,) saying, “Uh, an issue which is of interest here is dimensions of the, of the R constant, which depend on the dimensions used for the other variables, you know, for pressure, for volume. And, uh, uh, another item which can cause some complexity make, make, may make this question more difficult, if, if you don't give the value of the R constant because one can calculate the value himself or herself, uh, by using the ideal, uh, their standard conditions, you know, one mole under standard pressure and temperature. It occupies a volume 22.4 liters. So if you use the equation with these variables, you can deduce, you can calculate the value for R.”

When asked how students having to solve for the value of R would change his sorting, Interviewee G said,

“Well, that's, uh, uh, again, depends on, uh, if students have encountered that before, if they haven't, uh, it can be to the apply. Okay. Process application.

“Uh, it has factual knowledge in it. But if you have not done before, for instance, if you don't give the students, uh, the value of the R constant, okay, students may ask, what is the value of R, I need that to use in, in the calculation and you tell them, uh, I don't give the value to you because you can calculate that on your own. So, uh, if the student has not done that before, he has to think, so it's an application and it's, uh, uh, even analysis can be involved in that. So I could, I could stop at analysis.

“Uh, well it's procedural knowledge is there. Okay. So apart from, uh, factual knowledge, uh, conceptual knowledge depends on how practice you have. Okay. If you have, uh, enough practice with that you may not need to understand the concepts because the concepts are already there, but still you have to apply some procedure. So I could go as far as procedural knowledge and, uh, application, and even it might need some analysis I should think.”

In these two excerpts, interviewee G provided a scenario in which he thought item Q2 might require the cognitive processes of apply or analyze, or even the knowledge Dimension category of procedural knowledge (line 7-34; coded as both Conditional and Assumptions of Student Reasoning or Learning Context:) if the value of the gas constant R were not provided in the question stem and the students were required to solve for it (lines 5-26; coded as Complexity.) Interviewee G uses his prior experience with Bloom's taxonomy to determine what he considered to be procedural knowledge (lines 28-34; coded as Internalized Prior Experience.) As the question is written, students are given the value of R and would not have to calculate it, so interviewee G was not coded as double sorting item Q2.

Then Interviewee G expanded on the reasoning he assumed students would use when answering item Q2 (lines 40-49; coded as Assumptions of Student Reasoning or Learning Context.)

> "Understand, you know, if it is what we call algorithmic type of problem, the students have, uh, uh, solved many, many, uh, questions, many exercises with the ideal gas equation, and you don't have to understand, you just have to proceed with the application. So in this case, you jump here from remembering you go to application, without having to, to, to understand the thing. Of course there are situations where students can apply the ideal gas equation, all the standard conditions, uh, not for the proper systems. So you may have something which is in their liquid state. And you can have students apply the ideal gas equation."

From these excerpts Interviewee G was coded as sorting item Q2 as existing at the BCT intersection of Remember Factual Knowledge, a visualization of which can be seen in Figure 4.3.

Interviewee J

Interviewee J began her sorting of item Q2 along the Knowledge Dimension, saying,

> "I think that this is mostly factual knowledge, um, yeah, that, that's the strongest one for me."

Because Interviewee J concluded that item Q2 required only factual knowledge, this excerpt was assigned the *a priori* code Factual Knowledge using directed content analysis, but her indecision was coded as Unconfidence (lines 1-2.)

Moving on to sort item Q2 along the Cognitive Process Dimension, Interviewee J said,

> "And then um assuming that again the way that it's worded is 'write the ideal gas equation' to me that's a remember activity: retrieve, um, or, or find, look up somewhere what that equation is.
>
> "There's two pieces to this, right. And so then give the units used for each term in the equation. Um, give the units, depending on the context that might be, um, that might be a higher level. If the student hasn't really been asked to do that before, that could be a remember, right? Like we've sort of learned all these things before. Um, but that could be where you have to kind of like rearrange the equation and solve for all of those things. And so then that would be a high level, um, maybe an apply. Yeah."

Because Interviewee J sorted item Q2 as requiring either remember or apply cognitive processes depending on the requirements of the question (lines 8-14; coded as Conditional,) this excerpt was deductively coded as both Remember and Apply. Interviewee J started her sorting of item Q2 by examining the how the question was phrased (lines 3-8; coded as Item Forensics) and then based her decision on the potential dual nature (lines 7-14; coded as Complexity) of the actions students would need to take to answer the item (lines 5 and 9-13; coded as Assumptions of Student Reasoning or Learning Context.) Interviewee J thought that if a student was unfamiliar with the ideal gas law and rearranging it to find units, that item Q2 might require the cognitive process of apply, but that otherwise, it would be a remember process.

From these excerpts interviewee J was double coded as sorting item Q2 at both the BCT intersection of Remember Factual Knowledge and of Apply Factual Knowledge, a visualization of which can be seen in Figure 4.3.

Item Q3

Item Q3 was sorted by the 5 interviewees A, F, J, K, and M. Item Q3 was the first question sorted by interviewee J, the last question sorted by interviewee F, and a middle question for the remaining interviewees.

Interviewee A

Interviewee A began his sorting of item Q3 along the knowledge dimension, saying,

> “Ooh, so right off the bat, I see you know, something different about this question that students have to explain sort of a phenomenon rather than calculate something or, or do what you normally do with these, with these blasted, with these blasted ratios of elements, right? And so along the, along the, the left axis, I would, I would first say

6 that this is conceptual knowledge. They do need to have some sort of
7 concept about what these ratios mean, how they are experimentally
8 determined and things like that. Oh, well, I mean, that might be a
9 combination of factual, conceptual and procedural knowledge, right.
10 Because you know, depending on what you need to sort of attend to to
11 craft this explanation but let's for the sake of, of having to choose
12 something, I'll say that this is, this is getting at some conceptual
13 knowledge about the ratios of elements in, in, in some type of sample."

Interviewee A narrowed his potential sorting to conceptual knowledge by first examining the nature of the question (lines 1-4; coded as Item Forensics.) When trying to align that along what he referred to as the "left axis" of the BCT (line 5; coded as Interaction with BCT) he found he was having difficulty determining (lines 5-6 and 8-13; coded as Unconfidence) which line of reasoning he assumed students would use when tasked with this item (lines 6-8; coded as Assumptions of Student Reasoning or Learning Context) and could see arguments for multiple knowledge categories (lines 8-11; coded as Conditional.) However, when told that he did not have to narrow it to one intersection if he does not feel that choosing one is reasonable, he continued,

14 "Gotcha. thanks for giving me that, that prompt. Yeah. Okay. Then I
15 won't decide on a category here um because it's it's very hard to know
16 what a student would, how a student, we know what would be the first
17 attendance in sort of, I use like a knowledge-in-pieces type of picture
18 every time I think about students and their brains and that's what I'm
19 going with that here. And so you know, what is immediately
20 highlighted by this? I don't really know, but I think it could be
21 anything but perhaps metacognitive knowledge."

Because Interviewee A sorted item Q3 as requiring any type of knowledge except potentially metacognitive (lines 19-21; coded as Unconfidence, Interaction with BCT, and Elimination) depending on knowledge he assumed a student would have access to when presented with this item (lines 14-19; coded as both Conditional and Assumptions of Student Reasoning or Learning Context,) this excerpt was deductively coded as containing Factual Knowledge, Conceptual Knowledge, and Procedural Knowledge.

Moving on to sort item Q3 along the Cognitive Process Dimension, Interviewee A said,

> “Now along the other upside here, it says the word ‘explain’ which, you know, by just matching words, would key into understand, but I think this is well beyond understanding. I think this is more, I think this is more evaluate because students are asked to, students are given perhaps a, you know, a real, a real result and asked why that is different than what maybe a theoretical value would be, right? And so that's evaluating information, that is, to my mind anyway. And sure you can write, ‘explain’ there and they would, you know, they would have to do that. I don't know how you can do evaluation without explanation about what your evaluation is, but I, I think that it's pretty high on the, on the upper scale. I don't see them creating anything, but I do see them evaluating sort of a result against a theoretical standard.”

Because Interviewee A concluded that item Q3 required only the evaluate cognitive process, this excerpt was assigned the *a priori* code Evaluate using directed content analysis. Interviewee A discussed how one could pattern-match the wording of the item with the BCT Cognitive Process Dimension header of Understand (lines 22-23; coded as Pattern-Matching,) but dismissed this (lines 23-24; coded as Elimination) due to re-examining the complexity (lines 24-27) of what the question is trying to elicit (lines 25-27;

coded as Item Forensics.) Interviewee A relies on his prior experience with Bloom's taxonomy to determine (lines 27-34; coded as Internalized Prior Experience) what counts as a task further along the Cognitive Process Dimension (lines 30-34; coded as Interaction with BCT) insofar as students might approach it (lines 27-34; coded as Assumptions of Student Reasoning or Learning Context,) but he does express some uncertainty of how his experiences align with the BCT (lines 24, 28, and 30-34; coded as Unconfidence.)

From these excerpts Interviewee A was coded as sorting item Q3 as existing at the three BCT intersections of Evaluate Factual Knowledge, Evaluate Conceptual Knowledge, and Evaluate Procedural Knowledge, a visualization of which can be seen in Figure 4.4.

Interviewee F

Interviewee F began her sorting of item Q3 along the Knowledge Dimension, saying,

> "Ooh, fun. Okay. So I think I'd have to go conce--ah I think I have to go conceptual. I mean again, I might go factual if they had that this was verbatim, like, 'this is how this works.' And so they just need to repeat, 'but this is how this works.' Uh but my, my instinct is to say that they'd probably have to think about this. And then I would think they would need to analyze, I guess, is where I would land. So, so thinking about what those ratios are, what those ratios actually mean, how they relate to each other. I don't think I would put it on evaluate. I don't think I'd put it apply or below. I mean, so unless they really, you know, they already knew this and it fell into like factual knowledge and remember/understand. So maybe I could go there, but otherwise I would go conceptual knowledge and analyze. Yeah."

Because Interviewee F concluded that item Q3 should be sorted as existing at the BCT intersection of Analyze Conceptual Knowledge, this excerpt was assigned the *a priori* codes of Analyze and Conceptual Knowledge using directed content analysis. Interviewee F made this determination based on the context in which she assumed this item would be presented to students (lines 2-12; coded as Assumptions of Student Reasoning or Learning Context) whether that was as a verbatim repeat, or as a stand-alone item (lines 2-5 and 9-12; coded as Conditional) which helped her to remove the other cognitive processes from consideration (lines 8-9; coded as Elimination.) If it was a stand-alone item, she would expect the complexity of the item to affect its sorting. Interviewee F did show some hesitance in her sorting (lines 1-2, 4-6, and 8-11; sorted as Unconfidence.)

From these excerpts Interviewee F was coded as sorting item Q3 as existing at the BCT intersection of Analyze Conceptual Knowledge, a visualization of which can be seen in Figure 4.4.

Interviewee J

Interviewee J also began her sorting of item Q3 along the Knowledge Dimension (line 1; coded as Interaction with BCT,) saying,

> “So, um, again thinking about the knowledge dimension, I, it is not metacognitive and I would probably throw out factual knowledge as well. Um, without, again, without knowing what the instructor had previously told the students, um, I'd say this is probably conceptual

knowledge. The students might use a procedure in order to, or they might be familiar, be thinking about a procedure, but they're not really doing it right here. They're, they're thinking about the, um, sort of what the numbers mean and how they're, how they relate to a molecule. So I'd say conceptual knowledge."

Because Interviewee J concluded that item Q3 required only conceptual knowledge, this excerpt was assigned the *a priori* code Conceptual Knowledge using directed content analysis. Interviewee J mainly uses her assumptions of how this item would be presented to students and what lines of reasoning students may use to interact with the item (lines 3-9; coded as Assumptions of Student Reasoning or Learning Context) to remove metacognitive knowledge and factual knowledge from consideration (lines 1-3 and 6-7; coded as Elimination.) Though she admits some uncertainty in her sorting (lines 2-9; coded as Unconfidence.)

Moving on to sort item Q3 along the Cognitive Process Dimension, Interviewee J said,

"And then on the cognitive processes, um, again, I'm gonna, I'm going to assume that the instructor hasn't already told students, so this isn't a remember question, but it might depend, um, I would argue this also isn't a create question, um, not an evaluate question, cause there's no like standard against which we're judging something. Um, and then kind of like the example you gave, I can see arguments for understand, apply and analyze."

Interviewee J narrowed her potential sorting to between the cognitive processes of understand, apply, and analyze (lines 10-16; coded as Unconfidence) based on the context in which she assumed students would be presented this item (lines 10-12; coded as Assumptions of Student Reasoning or Learning Context.) The conditions in which she imagines the item being presented (lines 10-12 and 14-16; coded as Conditional) removed the cognitive processes of remember, evaluate, and create from consideration (lines 11-14; coded as Elimination.)

Then Interviewee J expanded on the arguments she could see for each of those cognitive processes, saying,

> “Um, understand because the goal here is to sort of just see, I guess it's about, it's about a relationship is to see the connection between the fractional numbers. And I guess, yeah, I might say that this is an understand question because I'm expecting, um, students to see that the, the numbers that they're given are very close to the integers and the difference is, um, probably due to some sort of measurement error. Um, and so that's, that, that might be sort of my number one is that this is conceptual and understand. Um, I could see an argument for, just an argument for analyzing would be, um, [the students] trying to decide like how, how big of a difference sort of is close enough. Um, and so then you're looking at like, what, like what do those three numbers mean in relationship to each other? And then I could see apply because the student has probably, perhaps the student has in class done a similar example, and now they're using the same method to look at this new number.”

Interviewee J started by looking once more at what the item was trying to elicit (lines 17-19; coded as Item Forensics.) Because Interviewee J sorted item Q3 as requiring any of the cognitive processes understand, apply, and analyze (lines 19-31; coded as Conditional) depending on the way she assumed the item might be presented to students and the lines of reasoning with which she assumes students might engage when approaching this item (lines 20-22 and 26-31; coded as Assumptions of Student Reasoning or Learning Context,) this excerpt was deductively coded as containing the *in vivo* codes Understand, Apply, and Analyze. Interviewee J's use of uncertain language was coded as Unconfidence (lines 17, 19, 22-31.)

When asked about how she was interacting with the BCT to make these determinations, Interviewee J said,

> 31 "So, that's a good question. Um, I definitely, with the knowledge
> 32 dimension, so I'm, I, um, like you mentioned, I'm more familiar with
> 33 the cognitive levels and less familiar with the knowledge dimension.
> 34 Um, and so that one, I am kind of reading these again and then for the
> 35 cognitive processes, I am looking at the verbs. Um, but I, yeah, I'm
> 36 also relying a lot on my, the sort of definitions that I've made for
> 37 myself about what I think what I think students are doing. And so,
> 38 you know, where I, where I sort of did that is especially on the
> 39 analyze, um, the question I'm sort of asking myself is like, 'can our
> 40 students,' students are looking for what pieces are important. And for
> 41 apply. I'm thinking about the, um, probably the like lower level of
> 42 apply, which, which is we've done something similar before and now
> 43 you're doing it, but with a different, like a different set of numbers."

Interviewee J uses item Q3 (lines 39-40; coded as Item Forensics) as an example to explain how she relies much more on the BCT for the Knowledge Dimension because she was not used to using a bidimensional version of Bloom's taxonomy (lines 31-35; coded as Interaction with the BCT) but that for the Cognitive Process Dimension she does refer back to her previous experience with Bloom's taxonomy (lines 35-43; coded as Internalized Prior Experience) and the way she applies that to her assumptions of how material like this might be presented to students (lines 40-43; coded as Assumptions of Student Reasoning or Learning Context.)

From these excerpts Interviewee J was coded as sorting item Q3 as existing at the three BCT intersections of Apply Conceptual Knowledge, Apply Conceptual Knowledge, and Analyze Conceptual Knowledge, a visualization of which can be seen in Figure 4.4.

Interviewee K

Interviewee K began his sorting by considering the nature of the question (lines 1-6; coded as Item Forensics,) saying,

> “Okay. So, so I guess in the previous question, in the 3.15, or in the sample one, it would have been on empirical formulas. Right. And so then students would have walked through that and gone through that and done all the calculations. And so now it's asking, so we don't have exact integers or exac- exactly integer ratios for your CH and O, but instead we have these decimals. Um, so let me see.”

Then he moved onto the Cognitive Process Dimension, saying,

> “Um, um, I would probably say, uh, evaluate, so evaluate is ‘students are asked to based on criteria and standards, judge assess, assign value, critique examine test, check, discriminate, justify.’ So I would probably say evaluate because, um, if for instance, students had to do, uh, the question and then they're getting 2.98 and 7.91, then they would have to kind of justify, I guess, to themselves and to whoever the instructor is why it can be rounded up to three why it can be rounded up to eight? Uh, why the 1.00 is the exac- is exactly 1.00. And so I would think that justifying, it would be part of it because this is something that comes up, you know, like when you're looking at these empirical formulas, you know, like if it's 2.33, then you know, like a student would have to justify, ‘well, I went two-and-one-third versus rounding it down to two.’ And so I think, I think I would say evaluate and, um, um, uh, I suppose I think I'd probably put it with procedural

knowledge. So I guess evaluate with procedural knowledge because students are conducting the empirical formula determination procedure. And so then you kinda have to understand, or I guess justify as well the different parts of the procedure and how are you making your decisions as you go along. And so I think I would say evaluate, uh, procedural knowledge."

Because Interviewee K concluded that item Q3 should be sorted as existing at the BCT intersection of Evaluate Procedural Knowledge, this excerpt was assigned the *a priori* codes of Evaluate and Procedural Knowledge using directed content analysis. Interviewee K made this determination based on both the reasoning he assumed students would use when approaching this item (lines 11-19 and 22-25; coded as Assumptions of Student Reasoning or Learning Context) and the Cognitive Process Dimension heading descriptions (lines 7-9; coded as Interaction with BCT.) This excerpt was also coded as Unconfidence to reflect Interviewee K's uncertain language in sorting this item (lines 7, 10, 12, 15, 19-21, 23, and 25.)

From these excerpts Interviewee K was coded as sorting item Q3 as existing at the BCT intersection of Analyze Conceptual Knowledge, a visualization of which can be seen in Figure 4.4.

Interviewee M

Interviewee M also began her sorting by considering the nature of the question as well as asking clarifying questions about the BCT (lines 7-9; coded as Interaction with BCT) saying,

> "Um I guess I can say from the get, I like don't feel good about trying to categorize this…I wish this wasn't being captured in interviews. I'm trying to like, remember how we even do this stuff and why I like cannot remember for the life of me, why this is important, but…I can like sort of try and reason about this. I, this is one where I would say it's probably going to depend on instruction. But I would say it, okay, so for, for procedural knowledge, um, is the concept, are the concepts that like underpin a procedure part of procedural knowledge or is that conceptual knowledge?"

In this excerpt, Interviewee M says she feels uncomfortable about sorting the item (lines; coded as Unconfidence,) but she does suggest that the item's sorting will probably rely upon the context in which it was presented to students (lines 5-6; coded as Assumptions of Student Reasoning or Learning Context.) Once the interviewer answered her question, Interviewee M went on to say,

> "…I'm, I'm comfortable with that reasoning.
>
> "So, I would say that this would fall under conceptual knowledge. I think another sort of indicator of that is that it's asking the learner to explain. Um, And so then, because you, so kindly gave me the word, explain in understand, I would probably put it there. I would probably do this as, as conceptual understand."

Because Interviewee M concluded that item Q3 should be sorted as existing at the BCT intersection of Understand Conceptual Knowledge, this excerpt was assigned the *a priori* codes of Understand and Conceptual Knowledge using directed content analysis. Interviewee M made this determination by keying into a verb listed in both the tool and the item, without further consideration (lines 12-15; coded as Pattern-Matching.) Interestingly, Interviewee M used her instinct about the necessary cognitive process to triangulate her sorting along the Knowledge Dimension. This excerpt (lines 12-16) was also coded as Unconfidence to reflect Interviewee M's uncertain language in sorting this item.

When asked if she was relying more on the her interactions with the BCT or her experience teaching this material to make these determinations, Interviewee M said,

> "Uh, I would say it's both. I would say because I'm still familiarizing myself with the tool, um, I end up having to, like, reread the definitions over and over again, to kind of get a sense of how to apply them. Um, I would say that, like the reason that the second question you sent, um, [Item Q8] I was more confident in my ability to reason about whether that should live in this tool, just because I had such a vivid memory of trying to teach that thing and how students, how I remember students interacting with it. Um, and I don't have that with the first or the third questions you sent. So in the absence of that knowledge, I sort of just like pick where I think it would fit based on the definitions provided in the tool."

In this excerpt, Interviewee M gives some insight into how she is using the BCT. She relies on the tool (lines 20-23 and 29-30; coded as Interaction with BCT) more when she does not have rich experience teaching the material on which to refer (lines 23-29; coded as Assumptions of Student Reasoning or Learning Context.)

From these excerpts Interviewee M was coded as sorting item Q3 as existing at the BCT intersection of Understand Conceptual Knowledge, a visualization of which can be seen in Figure 4.4.

Item Q4

Item Q4 was sorted by only interviewee H, as his first item.

Interviewee H

Interviewee H began his sorting of item Q4 along the Knowledge Dimension, structuring it the same way the interviewer did (lines 1-2; coded as Interaction with BCT) saying,

> Okay. So if we do this, like you were saying with using the knowledge levels first I could see a case for it in both conceptual knowledge, as well as procedural knowledge. Because implicitly they're going to have to work through…the calculation to result in a higher mass. So the st- they would have to be recalling steps, not only of the procedure, but also the calculation, but I think I would probably land it more in the conceptual knowledge because they are taking basic concepts and combining them together.

Because Interviewee H concluded that item Q4 required both factual and conceptual knowledge, but leaned towards the less abstract conceptual knowledge, (line 6; coded as Unconfidence) this excerpt was assigned both *a priori* codes Conceptual Knowledge and Procedural Knowledge using directed content analysis. Interviewee H sorted item Q4 based on the reasoning he assumed students would undergo when tasked with this item (lines 3-8; coded as Assumptions of Student Reasoning or Learning Context.)

Moving on to sort item Q4 along the Cognitive Process Dimension, Interviewee H said,

> And then for the, the other axis I would probably tend to put it into the analyze box because [the students are] trying to ascertain the

> parts, you know, how each thing affects the overall. They have to construct an outline and integrate that into, across, to answer the question.
>
> They're not, it's, lack of a better term, it's more complex than just understanding because you're doing multiple parts. So you're not really clarifying, you're doing more than that. So that was from the understand, same thing with apply that it may be depending on the context and how much like you said, where this was given if they were, if it was after they just ran those calculations, I could see where it would be, uh, in the apply level, but if it was just given standalone I think it would probably fall more under the analyze section.

Because Interviewee H sorted item Q4 as requiring either apply or analyze cognitive processes depending on when this item was presented to students (lines 18-21; coded as both Conditional and Assumptions of Student Reasoning or Learning Context,) this excerpt was deductively coded as both Apply and Analyze, but his indecision was coded as Unconfidence (lines 9 and 18-22.) Interviewee H sorted item Q4 as requiring either cognitive process by citing the reasoning he assumed students would use when approaching this item (lines 10-17; coded as Assumptions of Student Reasoning or Learning Context.) By comparing this line of reasoning to the Cognitive Process Dimension heading descriptors (lines 16-18; coded as Interaction with BCT) he was able to remove Understand from consideration (lines 15-22; coded as Elimination) based on the number of actions students would have to take (lines 11-17; coded as Complexity.)

From these excerpts Interviewee H was coded as sorting item Q4 in all four BCT intersections at the interfaces of Apply and Analyze and Conceptual

Knowledge and Procedural Knowledge, a visualization of which can be seen in Figure 4.5.

Item Q5

Item Q5 was sorted by the two interviewees A and B. Item Q5 was a middle item for interviewee A and the last question sorted by interviewee B.

Interviewee A

Interviewee A began his sorting of item Q5 by discussing the futility of "blooming" items just from the instructor or researcher's perspective (lines 2-5; coded as Assumptions of Student Reasoning or Learning Context.)

> "Well, I mean, my first thought is that most students would probably, you know, without, with an absence of student work…which by the way, as, as a complete aside, I always wonder why people do blooming in the absence of the actual student work, because it's the student work that tells you how they've actually engaged."

Then Interviewee A reconsidered how item Q5 was written, saying,

> "…My first thought about what a student would do is that, well, the response is given in the stem of the question you know, it gives in the first sentence you know, that this is a test for water, water quality and here is an experimental result. And then students might conclude that the water quality is poor. Wait, is that true? 'One common test for water quality…' Oh, Oh, well, so it doesn't actually say that decreased oxygen equals bad water. So it's just a test for water quality involving oxygen. So students would have to have some type of requisite knowledge that oxygen in the water is a good thing or that pollutants in the water, um I don't know give rise to lower oxygen, decreased oxygen levels. And so they would have to, they would have to come to that conclusion and well, the way that the question is stated, I'm not sure they need to actually have that as a piece of knowledge in their brain. To me, the question kind of leads them there in, in, in one capacity. And so I'm not even sure if this is remember, if a student

goes, you know, just states that the water, ‘we can conclude that the water is of poor quality.’ Um, I could definitely see them coming to that answer with no cognitive engagement, by the way that the question is stated.

“Now another student might go, ‘Oh, okay. I need to unlock knowledge about pollutants and what those pollutants do to oxygen levels’ and things like that. And so I would think that that is some factual knowledge that they need to unlock. I'm not entirely sure that there needs to be some sort of a concept here to answer the question unless it was phrased differently. But a student might also rely on something that they've learned about oxygen solubility and how pollutants affect that actually have a conceptual model of all of that. But again, the question isn't, isn't asking to elicit any of that, it's just, ‘what can we conclude about that?’ So there is, there might be a concept at play here, but probably a student would remember something about pollutants and oxygen, and that would just be a factual type of knowledge. Either previous knowledge or context of the question.”

In this excerpt, interviewee A focuses on what the question is asking of students (lines 6-13, 23-24, 29-31, and 33-35; coded as Item Forensics) and on the line of reasoning students are anticipated to use when approaching this item (lines 13-29, 31-33, and 36-38; coded as Assumptions of Student Reasoning or Learning Context.) At first, he misreads the question, but even after correcting himself, he still thinks the question might not require any cognitive engagement from students whatsoever (lines 6-13, 15-20, and 28-37; coded as Unconfidence.) He does concede that a student may go so far as to “unlock” factual knowledge, but dismisses that the question requires conceptual knowledge (lines 20-22, 29-31, and 35-38; coded as Elimination.)

Narrowing his sortings along the BCT, Interviewee A goes on to say,

> 39 “Now, you know, if we're going with, [the students are] actually using
> 40 some knowledge and not taking a guess based off the context of the
> 41 question among the Bloom levels, well, if it's factual knowledge,
> 42 they're remembering it. If it's conceptual knowledge, then they're
> 43 understanding. I don't really see, I don't really see how application
> 44 could be at play here, given how the question is worded. It's just asked
> 45 ‘what do you conclude?’ It's not probing any deeper for anything that I
> 46 think is a higher-order, cognitive engagement.”

Because Interviewee A concluded that item Q5 should be sorted as existing at the BCT intersection of Remember Factual Knowledge, this excerpt was assigned the *a priori* codes of Remember and Factual Knowledge using directed content analysis. He made this determination by continuing to consider the reasoning he assumed students would undergo when approaching this item (lines 39-43; coded as Assumptions of Student Reasoning or Learning Context) and what he thinks the item is trying to elicit (lines 44-46; coded as Item Forensics,) which also helps him to remove the cognitive process of apply from consideration (lines 43-46; coded as Elimination.) Interviewee A does discuss the potential intersection of Understanding Conceptual Knowledge, but given his elimination of those categories in the previous excerpt, Interviewee A was not coded as double sorting item Q5 nor was it coded as Conditional (lines 39-46; coded as Unconfidence.)

Asked for any final thoughts on item Q5, interviewee A considers the nature of the question (lines 46-53; coded as Item Forensics,) saying, “It was an interesting one because I definitely even, even though I read it out loud, I jumped to a conclusion about what content was actually there. But yeah, I, to me, it it's one that masquerades potentially as something, a lot more higher-order than maybe what students might do to give a response that an instructor would go, ‘Oh, yay. They understand!’ but the processes there might not be aligned with what the instructor might have hoped in the construction of the question.”

From these excerpts Interviewee A was coded as sorting item Q5 as existing at the most upper-left BCT intersection of Remember Factual Knowledge, a visualization of which can be seen in Figure 4.6.

Interviewee B

Interviewee B began her sorting of item Q5 along the Cognitive Process Dimension, but like interviewee A, by considering the phrasing of the question, saying,

“Okay. So the way this is worded leads me to believe that the students are not familiar with water quality tests in general. But they may be familiar with dissolved oxygen, that concept of what that means. Okay. So let me think now what are we doing here? So, goodness. I think this might fall into analysis for me. Because we're given a scenario, we're basically given data. I mean, it's, it's not data, it's relative data. Right. And it's asking the student to analyze, to think, ‘what do you know about this system and what can you, what can you say based on these results?’ Right. So that's almost like a classic analysis task in my mind and, and definitely probably more conceptual at this point based on what I know, but, but this, yeah, like I said, not a chemist, so maybe it's more procedural than I'm thinking. But it feels like an analysis task to me.”

Because Interviewee B sorted item Q5 as requiring only the analyze cognitive process, this excerpt was assigned the *a priori* code Analyze using directed content analysis. Interviewee B sorted item Q5 as requiring the cognitive process of analyze by citing both the context of how the item would be presented in a class (lines 1-3; coded as Assumptions of Student Reasoning or Learning Context) and the reasoning she thought the item was trying to elicit (lines 1-2 and 5-9; coded as Item Forensics,) as well as her previous experience with Bloom's taxonomy (lines 9-13; coded as Internalized Prior Experience.) This excerpt was also coded as Unconfidence to reflect Interviewee B's uncertain language in sorting this item (lines 4-5 and 9-13.)

When asked about how she was interacting with the BCT to make these determinations, Interviewee B said,

> "…So I'm kind of um I'm marrying your tool with my experiences using Bloom's and so are there keywords? Yes. I'm looking kind of for verbs, right, which is a common thing. I'm trying to get a sense of what is the task at hand? Like what are we trying to do? So here, when we say, 'what can we conclude?' Right. That's like, 'based on your analysis, what would you say is happening?' Like you're making a statement, you're making a claim. Right. And so my brain is like, 'okay, make a claim.' Right. So that that's part of the analysis. It could be I mean, if, if you're thinking about a claim differently, that could be a creative generative task, right. So that it falls in that upper, upper cognitive skill level, you know, area for me that we're asking students to really think deeply about this, these pieces of data, analyze it and come to a conclusion or make a claim based on, based on that information. So I guess that's how I'm thinking about it and interacting with the tool is, is thinking about it in that way. So honing in on the verb and then trying to suss out, like, what is, what is it that we're actually asking, right. Because we can use lots of verbs that are, that would fall at different levels, depending on how, on the task it's tied to. And so that's, that's how I'm seeing this is 'here's the verb. Here's what we're asking the students to do. Here's the, the broader task.' And that tells me, 'okay, this is not an absolute remember task. This is this other type of a task,' I think."

In this excerpt, interviewee B discusses referring to her previous experience with Bloom's taxonomy (lines 14-16 and 20-31; coded as Internalized Prior Experience,) while utilizing the BCT (lines 14-16, 26-35; coded as Interaction with the BCT) and relates it back to item Q5 via the line of reasoning she assumes a student would use when presented with this item (lines 16-17 and 19-26; coded as Assumptions of Student Reasoning or Learning Context) and the reasoning she thought the item was trying to elicit (lines 17-19, 24-26, and 29-35; coded as Item Forensics.)

While Interviewee B did start considering the Knowledge Dimension while sorting along the Cognitive Process Dimension (lines 1-13,) in the following excerpt, she narrows her sorting of the Knowledge Dimension while continuing to consider how she interacts with the BCT (lines 36-41; coded as Interaction with BCT,) saying,

"Yeah, I would use the tool more for that looking at like, okay. You know factual knowledge, am I asking, am I asking for recall of like terminology definitions, et cetera? No, I'm not in this question. Okay. So now I'm thinking about, I know it's not metacognitive cause I'm not asking them to explain how they know this to be so, right. So I'm looking at between procedural and conceptual knowledge. I'm trying to understand, you know, this is where I would fall short because I'm not sure is this a common skill set that, that the students have been taught if, so that puts it into procedural knowledge, but are they applying conceptual understanding, which is where I tend to think it is. Is that, 'do you understand what's happening when oxygen is dissolved in water and you've got other pollutants and those interact,' those not interactions, but reactions–good job–um happening between oxygen and those other elements. To me, that's what I think is much

51 more probably the conceptual side of things is where I think that would fall from my perspective."

Because Interviewee B concluded that item Q5 required only conceptual knowledge, this excerpt was assigned the *a priori* code Conceptual Knowledge using directed content analysis. In this excerpt, Interviewee B relies much more on the BCT than she did for the Cognitive Process Dimension, because she was not used to using a bidimensional version of Bloom's taxonomy. Interviewee B mainly uses her understanding of what the item is trying to elicit (lines 37-40 and 46-49; coded as Item Forensics) to remove metacognitive knowledge and factual knowledge from consideration (lines 36-40; coded as Elimination.) Even thought Interviewee B can use her prior experience with Bloom's taxonomy to inform her sorting (lines 39-40, 44-46, and 49-51; coded as Internalized Prior Experience,) she admits that she is unsure of the discipline-specific knowledge that this item would need (lines 40-49; coded as Unconfidence) but that someone with more chemistry knowledge might code it differently (lines 41-51; coded as Conditional) and that chemistry students may or may not be familiar (lines 42-44; coded as Assumptions of Student Reasoning or Learning Context) with what this scenario might entail (lines 46-49; coded as Complexity.)

From these excerpts Interviewee B was coded as sorting item Q5 as existing at the BCT intersection of Analyze Conceptual Knowledge, a visualization of which can be seen in Figure 4.6.

Item Q6

Item Q6 was sorted by the 8 interviewees B, C, D, E, G, H, I, and J. Item Q6 was the first question sorted by interviewees B and E, the only item sorted by interviewees D and I, and a middle item for the remaining interviewees.

Interviewee B

Interviewee B began her sorting by considering the nature of the question (lines 1-5; coded as Item Forensics,) saying,

> "Oh, powder form. That's tricky. Anytime you've got stuff in powder. That's crazy. I wouldn't know how to do that. Okay. So we're proposing a method. This is tough for me because it could be, if the students have had, yeah, goodness. It could be any…so depending on the context of the course, right?"

In this excerpt, Interviewee B discusses her perceived uncertainty (lines 1-5; coded as Unconfidence) due to the difficulty of the item, based on her expertise as a Biology Education Research community member (lines 1-3; coded as Difficulty.) She also notes the conditional nature of sorting this item (lines 3-5; coded as Conditional,) based on the context in which it was presented to students (lines 4-5; coded as Assumptions of Student Reasoning or Learning Context.)

Expanding on the contextual/conditional nature of the item, Interviewee B went on to say, "Maybe they've already seen these examples before, in which case, this is a remember question for sure. But maybe, what the students have been exposed to is just talking about 'how we separate mixtures,' right? That's the general concept. And so we have different tools that we might use to separate mixtures. So we have some principles that we're aware of. And then we have to take each of these examples and think about, okay, 'what is blood? What do I know about blood? And which principles in terms of mixture separation would apply here.' And then that's, that's going to dictate what I would suggest. So for me, that would be an applied type question, um, based on what I just said there, and then I think it's more a conceptual knowledge or that dimension would be conceptual. Like, do you understand the concept of mixture separation? And you're demonstrating that by applying it to these new scenarios that you've never seen before. So I think that's what I would say there. I don't know. How do you feel about that?"

Because Interviewee B concluded that item Q6 should be sorted as existing at the BCT intersection of apply conceptual knowledge, this excerpt was assigned the *a priori* codes of Apply and Conceptual Knowledge using directed content analysis. Interviewee B made this determination mainy based on the reasoning with which she assumed students would engage when tasked with this item and the context in which she assumed it would be presented to students (lines 6-13 and 17-19; coded as Assumptions of Student Reasoning or Learning Context.) She also re-emphasized the conditional nature of sorting this item (lines 6-9 and 14; coded as Conditional.) This excerpt was also coded as Unconfidence to reflect Interviewee B's uncertain language in sorting this item (lines 14-16 and 19-20.)

From these excerpts Interviewee B was coded as sorting item Q6 as existing at the BCT intersection of Apply Conceptual Knowledge, a visualization of which can be seen in Figure 4.7.

Interviewee C

Interviewee C began her sorting of item Q6 along the Knowledge Dimension, saying,

> “Okay. So propose a method…So propose a method, I'm thinking procedural, because they ha- you know, but they're not really actually doing it. So this, it says- so maybe it's not procedural because they, you asking them to explain the method, but they not, I don't actually have to do a method. So I look at procedural and it says, uh, it's ‘techniques and skills.’ So I think I will leave that out. So let's, um, go up to conceptual and factual and see where it would end up. Um, hmm…so I would say this would be conceptual in terms of, um, uh, the knowledge dimension.
>
> “Right. I mean, I'm still stuck between thinking about how procedural…but, um, because, I think if you were doing it in the lab, there will be some…um, you have to have some idea of how to do this step-by-step. You know how to go about doing it. So the procedural is in my mind because I'm looking at it…I see. You know, i- can you, can you clarify procedural in terms of, um, is it, is it something that you have to be able to do, or you just have to up knowledge of how to do–?”

Interviewee C narrowed her potential sorting to between conceptual and procedural knowledge (lines 3, 6, and 11-17; coded as Unconfidence.) She first gravitated towards procedural, but thought that because students only had to propose a procedure instead of conducting it, that maybe instead the item was intended to elicit conceptual knowledge (lines 3-5 and 12-14; coded as Conditional.) She came to this conclusion mainly based on the headers of the knowledge dimension (lines 5-7 and 15-17; coded as Interaction with BCT) and the context in which she assumed the item would be presented to students (lines 12-14; coded as Assumptions of Student Reasoning or Learning Context,) but when she was unable to make a further determination, Interviewee C asked a clarifying question about the nature of procedural knowledge (lines 3-12 and 15-17; coded as Unconfidence.)

After her question was answered, Interviewee C went on to say,

> “Okay. Okay. Okay. So then, so then, um, I'll go back to between procedural and conceptual. There's…there's both of these, have to be, cause you, yes, you have to know the techniques, but then you have to know why, why are you doing it this way? Why putting, you know, if you're doing unrefined petroleum, besides just knowing the procedure, you have to know…okay, let's say you were going to use distillation or something. You have to have that knowledge of distillation and how it works. So it's somewhere between conceptual and procedural and it's okay to choose those two, right? Okay. Um, and in terms of, um, this will be an apply, I believe. Um, I will be, uh, mostly understanding, apply, but apply. I see it because you have to use that knowledge in order to, um, to carry out this experiment or, um, so I would say it's more of an apply.”

Because Interviewee C concluded that item Q6 required both conceptual and procedural knowledge, this excerpt was assigned both the *a priori* code Conceptual Knowledge as well as Procedural Knowledge using directed content analysis. Interviewee C sorted item Q6 as existing at this interface based on the reasoning she assumed students would undergo when tasked with this item (lines 19-24 and 27-28; coded as Assumptions of Student Reasoning or Learning Context,) but throughout the process she questioned both choices and whether she had to pick one (lines 17-20 and 24-27; coded as Unconfidence.)

When asked about her use of the phrase "carry out," interviewee C went on to explain that she was matching activities which she thought students were conducting with language in the Cognitive Process headers (lines 30-34; coded as Interaction with BCT) but her uncertainty was coded as Unconfidence (lines 32-34.)

30 "Based on their, their, their knowledge and, and then analyze, 'in
31 order to break material into' how those parts relate to one another. So,
32 apply/analyze. Hmm. Are we going all the way up to evaluate? And
32 that's 'based on a criterion and standards, judge, assess, assign value,
34 critique.' Okay. Yeah. So apply to evaluate."

When asked whether this question and subject matter is something that she would cover in her teaching, interviewee C explained,

35 "Mmhm a little bit. Yeah. Um, yeah, I think it's multiple components,
36 but let me just read these again in the mind– and probably think
37 about…yeah. I still leaning towards apply.
38
39 "I mean, yes. I just taught lab, um, the organic lab last semester. And
40 in that lab, I, I guess we talk about i- It did a lot of distillation, even
41 though we didn't do it in person this time, but they do a lot of
42 distillation and distillation they have to separate, um, the fractional
43 distillation and separate unknown materials. So in the, in the real lab,
44 the in-person lab they'll have to separate, um, an unknown two
45 unknown, um, components. And so they have to actually go through
46 and do it. And then they have a lot of analysis after that because they
47 have to figure out what the unknown is. Um, but they do talk about
48 pro- petroleum, maybe just briefly in terms of, just for as a fractional
49 distillation process. Um, and also separation of, uh, let's say, um, salt
50 in some other, you know, because they they have to think about, uh,
51 uh, you, maybe evaporating or dissolving this, you have sand and salt
52 together, maybe dissolving the salt, filtering those kinds of things. So

they do some filtration and things like that. So there's some procedural stuff involved, but it's also some understanding of the conceptual as to why we're doing things the way we're doing it."

In this excerpt, interviewee C explains the overlap of her experience with teaching organic lab with item Q6. She points out the multiple components of item Q6 (lines 35; coded as Complexity) and refers once again to the Cognitive Process dimension headers (lines 36; coded as Interaction with BCT) to compare against the context in which she assumes this item would be presented to students (lines 39-40; coded as Assumptions of Student Reasoning or Learning Context) but she still exhibits uncertainty in her sorting (lines 37 and 40; coded as Unconfidence.)

Asked if this contextualization helped her narrow her cognitive process dimension sortings, interviewee C discussed the difficulty of sorting (lines 56-57 and 67-69; coded as Difficulty.)

"You mean in the, in the, yeah, the co- I'm like, this is not as easy as it looks, um, between you see, that apply, or an, or which one? And an- analyze, apply and evaluate. You mean, thinking about how I w- what we talk about? Mmhm, yeah, I see all those words in apply seem to come, you know, I seem to relate to the words that's under apply because, yeah, they have to 'carry out or conduct,' and that's more, you know, what they will do in the lab. Um, and the analyze, it saying 'to break material into constituent parts and how those parts relate to each other, deconstruct, examine,' maybe, but I think it falls mostly under apply and then some analyze.

"Oh, that was not so easy because there's also you thinking about, you know, where people might fall, or when you think about the context of where this would be taught and how…"

In this excerpt, interviewee C reconsidered her Cognitive Process Dimension sorting, referring once again to the dimension headers (lines 59-65; coded as Interaction with BCT) and Pattern-Matching with the definition of the cognitive process of Apply (lines 59-62) to compare against the context in which she assumes students will be presented with this task (lines 61-62; coded as Assumptions of Student Reasoning or Learning Context.) Here, she reinforces her earlier sorting of apply, but still expresses some uncertainty of her sorting (lines 56-60 and 64-69; coded as Unconfidence)

From these excerpts Interviewee C was coded as double sorting item Q6 as existing at both the BCT intersection of Apply Conceptual Knowledge and of Apply Procedural Knowledge, a visualization of which can be seen in Figure 4.7.

Interviewee D

Interviewee D began his sorting of item Q6 along the Knowledge Dimension, saying

> "So the first thing that comes to mind is that 'propose a method' seems like a procedure. So I'm going to scroll down to your tool and re-read that, but 'algorithms, techniques, and methods,' 'propose a method.' That seems to fall very neatly into the procedural knowledge category."

Because Interviewee D concluded that item Q6 required only procedural knowledge, this excerpt was assigned the *a priori* code Procedural Knowledge using directed content analysis. Interviewee D made this determination by comparing the wording of the question (lines 1-3; coded as Item Forensics) with the headings of the Knowledge Dimension (line 3; coded as Interaction with BCT,) though he does immediately reach for his previous experience with Bloom's Taxonomy (lines 1-2; coded as Internalized Prior Taxonomy Experience.)

Moving on to sort item Q6 along the Cognitive Process Dimension, Interviewee D said,

> "So I'm going to look across now, knowing it's going to be a procedural knowledge and decide on the cognitive process dimension. And I think it's already going to be a higher-ordered skill, because there's nothing…I guess that's not true. This could be just a memory question if they had just done this in a lab, if this is, 'okay, we just did this in lab last week. Now it's on the test.' That's going to be, 'what did you do in lab last week?' And I know some professors do ask those kinds of questions of 'what did you do in lab last week? Were you paying attention? Were you there?' But we're going to assume that's not the case."

In this excerpt, Interviewee D switches between BCT dimensions (lines; coded as Interaction with BCT,) and then discusses how the way that this item were presented to students (lines 14-15; coded as Assumptions of Student Reasoning or Learning Context) would affect its sorting (lines 7-14; coded as Conditional.)

After establishing his assumptions about the context in which this item would be presented, Interviewee D went on to say,

> "And it's going to be a higher-order skill that students are going to have to know what the different components are and how to separate them, so I feel like, right off the bat, reading the definitions, analyze is

‘students are asked, in order to break into parts.’ So they'd have to break what is blood into its various parts. What is iron sulfur and what is petroleum? So that's all…each as a mixture that would need to be broken down. Once it's broken down, they need to know how to actually separate that. So I feel like analyze is sticking out right off the bat. So procedural and analyze. Yeah. I don't feel like it would be evaluate or create because there's nothing to evaluate.

“I guess I didn't read create yet.”

In this excerpt, Interviewee D re-emphasizes that he thinks this item requires higher-order cognitive processes based on the lines of thought he assumes students will use when facing this problem (lines 16-23; coded as Assumptions of Student Reasoning or Learning Context) and alludes to the more complex natures of petroleum and blood compared to the iron-sulfur mixture (lines 17-23; coded as Complexity.) He uses the Cognitive Process Dimension headers (lines 18-20; coded as Interaction with BCT) to focus in on the cognitive process of analyze, and dismisses the two rightward processes (lines 24-25; coded as Elimination,) before realizing that he hadn’t considered the cognitive process of create yet (lines 23-27; coded as Unconfidence.)

After reading the cognitive process dimension heading for create, interviewee D went back to considering what the question is asking, (lines 34-38; coded as Item Forensics,) saying,

“I feel like create, it's really going to depend on what this is in context of. I feel like if this was a test question, I'm still comfortable in my analyze, but okay. We're gonna break this in, and I'm proposing, like, a hypothetical thing that may or may not work, but if this was like an actual, like ‘design a lab experiment and go and get the materials and actually do this experiment,’ I feel like that would be create. I feel like, I feel like it depends on what's on either side of this. So students actually need to think about materials and actually do like conclusions. And I feel like that's what create would be, is like, 'Oh, now they're actually doing this' versus I think, breaking down into blood and saying how to separate that, I'd be like, 'it's just analyze.'"

In this excerpt, interviewee D creates a hypothetical situation in which he could see item Q6 requiring the cognitive process of create. He expands on the types of contexts in which he could see this item being presented to students and how that would conditionally affect the sorting of item Q6 (lines 28-38; coded as both Assumptions of Student Reasoning or Learning Context and Conditional.)

Contemplating the nature of the Cognitive Process Dimension (lines 39-41; coded as Interaction with BCT,) interviewee D continued,

"But like, I know it's right in-between, but what's weird is evaluate's like not in between those two of them. There's not like, a middle ground between…I think it depends on, yeah, just if it's a question of like, 'how would you separate blood?' I could answer that on the test without really creating that much new knowledge versus, 'we went to a crime scene and there's these three mixtures and we need to figure out how to separate them and identify like, the suspect.' I feel like it would be like, 'Oh, now you're getting to create where they have to create like a scenario and do it and analyze it and give you a result of who the suspect is.' So I feel like I could see either depending how the professor writes this, it being either one."

In this excerpt, Interviewee D continues to differentiate what line of reasoning on the student's part (lines 41-48; coded as Assumptions of Student Reasoning or Learning Context) would delineate this item from being a create or analyze item (lines 48-49; coded as Conditional,) but his uncertainty was coded as Unconfidence (lines 39-41 and 45-49.)

When asked if the conditionality of the cognitive processes necessary to answer this item caused any change to his sorting of the Knowledge Dimension, interviewee D said,

> 50 “I feel like it's still procedural, unless the teacher went over it. I feel
> 51 like…I guess like you did with your examples, the teacher went over
> 52 like, ‘okay, here's the components of blood and here's how we would
> 53 separate that.’ And now you just have to kind of piece together, then
> 54 it's going to be like, ‘okay, well I know what the,’ but I feel like…’skills
> 55 specific to the interrelationship’...Yeah, I guess it's both on the line,
> 56 between there. So the skills would be the actual separating. And then
> 57 in order to separate it, you have to know the ‘interrelationships among
> 58 components.’ So neither, both, in between [procedural knowledge and
> 59 conceptual knowledge].”

In this excerpt, interviewee D reconsiders the Knowledge Dimension of the BCT, once again providing scenarios (lines 50-56 and 58-59; coded as Conditional) in which he could see the item eliciting Conceptual Knowledge or Procedural Knowledge (lines 50-51, 54-56, and 58-59; coded as Unconfidence) based on what skills students would need to exhibit (lines 53-58; coded as Assumptions of Student Reasoning or Learning Context.) He bases these skills on the headers of the Cognitive Process Dimension (lines 54-59; coded as Interaction with BCT,) but discusses how this item requires cumulative knowledge–that of the various separation procedures, but also of the principles upon which those procedures are based (lines 51-58; coded as Complexity.)

Moving on to construct some hypotheticals about what the item could be asking for (lines 61-63 and 67-84, ; coded as Item Forensics,) Interviewee D said,

> “I feel like you could ask, let's see, procedural analyze question, let's see if that's right– ‘Students’...Yeah. Yeah. So if you. Yeah, yeah. Procedural and analyze, and the question would be like, ‘here are the components of blood.’ So it gives them the components of blood and then be like, ‘okay, how would you separate this?’ And then that would be using procedural knowledge and all of the skills of like, ‘I would go get the sep funnel and I would sort it out by densities and I would whatever,’ probably a bad example for blood, because that's not going to really separate by density, but you know my point, but like if you gave them the components in the question you could, or in the directions, I guess, and still ask that same question, it would just be procedural doing it. But if they had to know what the components are, then it's both a conceptual because they have to know what it is and it's still getting out skills. So then it's like a two-part question.
>
> “And how would you grade that of, are you looking at a) can they break down into the different layers of blood and like, are they using the right skills? Is this a lab question or is this a lecture question? And the lecture you probably care more about can they identify the components and then know how to separate, that is a backseat, where in a lab, you probably care less that they missed a component and more that they actually can separate things out successfully. If it's both, what do you weigh? I feel like it becomes, a professor is weighing on which category do they care about, the skill and the lab, or do they care that [the students] know what the components are of petroleum? What the components are of blood?”

This hypothetical helps Interviewee D articulate the difference between the various interfaces he’s been considering–to Interviewee D, it matters if a student were given the components of each mixture, or if they had to determine them (lines 70-72; coded as Difficulty,) and whether the student has to have knowledge of separation procedures, or just the principles that would allow different components to be separated. To help him articulate this, Interviewee D uses the headings of both dimensions of the BCT (lines 59-61; coded as Interaction with BCT) to isolate the skills and knowledge he thinks exists at each intersection (lines 63-72; coded as Assumptions of Student Reasoning or Learning Context) and then matches that to the various ways he thinks the item could be presented to students (lines 60-84; coded as Conditional.)

When asked if he thought all parts of the item existed at the same BCT intersections, interviewee D said,

> 85 “I do think you would get lower things on iron-sulfur than you would
> 86 on blood…I feel like powder, you just have to find which liquid one is
> 87 soluble in and one isn't. And, done…Yeah. Dissolving could be a skill, I
> 88 guess, but that doesn't really… whereas like blood, I feel like there are
> 89 skills involved in separating blood. That is a multi-step process.
> 90 Whereas a powder could be solubility. I don't know, off the top of my
> 91 head what iron and sulfur are different solubilities in, but that should
> 92 be, that should be solvable.”

In this excerpt, interviewee D mainly refers to the number of components in each mixture and the number of steps involved in separating each mixture (lines 86-92; coded as Complexity) to say that he thinks the first task, of separating an iron-sulfur mixture in powder form, might require less knowledge and cognitive processes than the other two facets of the question (lines 85-86; coded as Item Forensics.)

Asked if he had anything else he wanted to say about this item, Interviewee D went on to emphasize the importance of context when sorting items (lines 94-101; coded as Conditional,) saying,

> "I think it just shows that it's hard to do this for, just like, one question fits everything, because based on the mixture, changes how much they have to think about this, whether it's like, 'okay, this is salt and a metal' or whatever. If you have some things that are clearly different, I can use a magnet and separate that. I forget what we did in the lab, but it's something where like, they use a magnet and it just comes right out. And I was like, 'okay, like, that's an easy mixture.' But if you do like blood, does that increase the difficulty, which also increases the amount of skills, and that's where it's harder to do. Like, just because these words are in the question stem doesn't mean it's always going to fit the same category."

He once again discusses the relationship between difficulty (lines 93-94 and 99-101; coded as Difficulty) and complexity (lines 93-103; coded as Complexity) in this item, and emphasizes that one sorting tactic will not work for all items (lines 93-94 and 100-103; coded as Unconfidence.)

From these excerpts, interviewee D was double coded as sorting item Q6 at the BCT intersections of Analyze Procedural Knowledge and Create Procedural Knowledge, a visualization of which can be seen in Figure 4.7.

Interviewee E

Interviewee E began her sorting of item Q6 along the Knowledge Dimension, saying,

> "Okay. So there's two, two mixtures. Um, so gosh, I mean, if I'm doing it as, systematic way, um, I guess I can follow what you did and start with the knowledge dimension. Um, so some of it, I mean, some of it is

factual, right? Because you're in, in terms of what the possible answers are, you're, you may be wanting to, and then I guess this goes into the cognitive process, but you need to remember what, um, unrefined petroleum and iron sulfur in a mixture is, right? Um, so then I could keep going. Um, is it conceptual? Um, I don't know if I think that this is conceptual. I might, I might say that it's not, um, let's see. Is it procedural? This could potentially be procedural, I guess, if you're thinking about…um, I mean, it's proposing, so I'm proposing something. So it kind of depends on if, if there was sort of a procedure. Um, I don't think it, I don't think procedural has to involve numbers. I don't think it has to involve an algorithm, or an equation or anything like that. I think it can be, um, uh, you know, 'I was shown this example of how to do this in class and can I apply that same process or thinking process.' So maybe procedural, in a sense, it does sort of feel, um, but the, the fact that you have to propose a method, um, maybe it takes it from what might be factual to more procedural. It's not metacognitive for sure. So, um, so I might come back to that. So then there's remember, you certainly have to remember some things. Um, you also, not just remember, I guess if you remember what an iron sulfur is and unref- or what iron sulfur mixture in powder form is and what unrefined petroleum is, does that extend to understanding? Um, maybe when you have to propose a method, like you have to understand them well enough to know what method might work. Um, so, so it, it could extend up to understand and possibly apply, but I don't know that I, I guess as you said, I think it depends on what they were shown in class. So I don't know that I think it goes beyond apply to analyze, cause I don't think there's any information here. There's what, there's factual things that you carry in your head, like, what are these different kinds of, um, mixtures and then remembering kind of, what it is. So it feels like you could easily answer this with a remember, um, if you had this set of things, depending on how it was presented to you, you know, and, and what you did, um, not clear whether or not it's quite to an understand and, and apply. It may be feeling a little bit on the lower ends to me."

Because Interviewee E concluded that item Q6 should be sorted as existing at the BCT intersections of Understand Procedural Knowledge and Apply Procedural Knowledge this excerpt was assigned the *a priori* codes of Understand, Apply, and Procedural Knowledge using directed content analysis. Interviewee E made this determination by moving down the Knowledge Dimension and across the Cognitive Process Dimension, and considering each category (lines 2-10, 19-21, 27-29, and 36-37; coded as Interaction with BCT) which allowed her to remove metacognitive knowledge and the analyze cognitive process from consideration (lines 19-20, 29-30, and 35-37; coded as Elimination.) Her sorting depended on (lines 3-5, 10-12, 27-30, and 34-35; coded as Conditional) all the things students would need to keep in mind to answer the item (lines 1-7, 12-26, and 30-35; coded as Complexity,) what ways she assumed students might interact with the item (lines 4-7, 11-18, 21-35; coded as Assumptions of Student Reasoning or Learning Context) and what the item might be asking of students (lines 1, 18, and 30; coded as Item Forensics.) Interestingly, while she was at first uncertain whether this item would elicit conceptual knowledge, Interviewee E checked her Knowledge Dimension sorting against her thoughts about the Cognitive Process Dimension (lines 1-30 and 33-37; coded as Unconfidence) and referred to her previous experience with Bloom's taxonomy to triangulate her sorting (lines 13-20, 29-30, and 36-37; coded as Internalized Prior Experience.)

Asked if she had considered the cognitive processes of evaluate and create, Interviewee E said, "Yeah. I mean, to me, evaluate, uh, okay, 'students were asked to, to, based on criteria and standards, judge, assess, um, value,' yeah, so you, you would have to present like, here's possible solutions, maybe decide between the solutions they're being asked to propose, the solution, and, I'm inferring, um, from this and maybe right or wrong that, that the students maybe were taught that there's, to identify particular kinds of mixtures. And then from those mixtures, 'here's certain things that you can do to separate mixtures,' right? So it's almost, it almost could be a matching, but they have to remember what to match to, you know, it's like one step away from what might be even a matching question in that sense. Um, so it definitely doesn't feel like evaluate to me and I don't know that they're really given anything to analyze either, you know, like, um, for create, they're there'- they're not being asked to, I guess, if they were never given any notion of what ways you might separate mixtures, if they had no idea and they had to just come up with something out of thin air, that could be a create. Like, if I was, if I didn't know anything about how chemists separate, um, you know, 'iron-sulfur mixtures in powder form,' then I might go, 'okay, what do I know about iron?' And I start brainstorming. And then that, to me, feels more like create, um, I guess, given what I know about chemistry questions. I'm assuming it's not that, right? Like, we don't often put students in that position. And so I'm, I'm, I'm still kind of sticking with the lower ends. Like remember, I definitely remember. I think, I think it's almost like you always remember, it's always there. It's just like, 'is it only remember, or do you go beyond that?' That's kind of the way I feel about it, but, um, it's, it's understand and potentially apply, but I don't feel like it goes above that. And if it were, if it were to be create, it would have to be with the understanding that they weren't given any like, options in class of, of ways to separate mixtures."

In this excerpt, Interviewee E compared her personal definition of evaluate (lines 38-42, 45-58, and 61-63; coded as Internalized Prior Experience) with that of the BCT (lines 38-39; coded as Interaction with BCT) and provided scenarios (lines 40-58 and 64-67; coded as Conditional) where this item could require the cognitive processes of Evaluate or create (lines 40-48 and 51-62; coded as Assumptions of Student Reasoning or Learning Context,) but explicitly states that those scenarios are unlikely (lines 38, 42-65; coded as Unconfidence) in order to remove those cognitive processes from consideration (lines 48-50 and 58-59; coded as Elimination.)

When asked whether this question and subject matter is something that she would cover in her teaching, interviewee E explained,

> "No, no, to me this feels like a gen chem question or maybe an orgo I, but not at [institution.] And I never, I don't, I've taught orgo I lecture, I've never taught orgo I lab. And this is, this, a lab question, I don't know, maybe not this, this kind of feels like mixtures, feels like understanding what a mixture is. It feels like, in general chemistry thing to me, or even a high school thing. So I wouldn't expect, this is not a question that I would have given or experienced as, in my instructional experience, but I don't teach at those levels. I don't teach those courses So…it"s weird because when I think I, you know, when I hear 'mixture,' that automatically makes me think of back when you're learning about the difference between solids and liquids and, and what a mixture is and what a pure compound is and things like that. So that feels very general chemistry. But then, yeah, I guess if you're talking about unrefined petroleum, um, you know, then I remember even, you know, at a master's level class in polymer chemistry, where we're talking about refining and, and other things like that and, and, you know, so you could see, but I don't think that this particular question would be. yet it does really feel general chemistry question or possibly, depending, possibly like earlier, organic than I teach."

From these excerpts Interviewee E was coded as sorting item Q6 as existing at the BCT intersections of Understand Procedural Knowledge, and of Apply Procedural Knowledge, a visualization of which can be seen in Figure 4.7.

Interviewee G

Interviewee G began his sorting by considering the nature of the question saying,

> "Yes. Yes. This is usually being taught in the lower secondary chemistry. You know, when you talk about mixtures and the iron and sulfur example, is a standard one, being useful separating, but, uh, uh, things like blood or, uh, unrefined petroleum, uh, mo–more complex, of course.
>
> "Uh, of course it depends on what your knowledge you have. Uh, what is your biology knowledge, knowledge about the composition of blood and, uh, uh, this, uh, requires of course knowledge, which, uh, you must have to do that. If you don't have that, you might have difficulties with it. Although, uh, people have some, have some experience with blood through blood analysis. Uh, we are usually doing, uh, uh, in a microbiology lab. And so we, there are things like, uh, blood cells and white cells. And, uh, we might have some idea about this kind of things, but you did, you didn't, you need a lot of knowledge about blood. And so this is a medical knowledge, biology knowledge, which is there, uh, and also, uh, because you have a lot of constituents, you might need a number of methods, separating techniques. Again, again, that depends, you know, you might be needing centrifusion or you might be needing chromatography, uh, or different, other you know, using some solvent to separate the things, uh, the unrefined petroleum. Again, it's a more familiar item because we know we have this distillation column by which we separate the various, uh, portions of constituents of uh, petroleum.
>
> "I mean, for myself they are, are different. And I find if I, if you ask me to do that now. I have to guess about the blood. I know, I know a lot about, uh, unrefined petroleum and the other one is a simple situation. So I, I [wouldn't] place them in the same box."

Interviewee G was the only participant who explicitly sorted each of the three tasks of item Q6 at separate intersections (lines 14-21 and 26-29; coded as Unconfidence.) In this excerpt, Interviewee G explains that different intersections are necessary based on (lines 7-24; coded as Conditional) the familiarity a student has with the material, and the complexity of the mixture (lines 1-5, 13-19, 23-24, and 28; coded as Complexity.) He provides assumptions of both the learning contexts in which students might see this item and the lines of reasoning students may employ when tasked with it (lines 1-18 and 22-24; coded as Assumptions of Student Reasoning or Learning Context.)

Specifying the sorting for what he considered to be the easiest task, of separating an iron-sulfur mixture in powder form, Interviewee G said,

> 30 “You're like, okay. With, uh, you start with easier one, it's just
> 31 knowledge remembering, you know, about the magnetic property of
> 32 iron. But usually this is an example, a standard example, usually. And
> 33 so it's just a goal of knowledge. So it goes, it goes factual knowledge.
> 34 And excuse me, is- it goes to the upper, you know, uh, top left box, the
> 35 simplest one.”

Because Interviewee G concluded that this facet of item Q6 should be sorted as existing at the upper leftmost box (lines 33-35; coded as Interaction with BCT,) the BCT intersection of Remember Factual Knowledge, this excerpt was assigned the *a priori* codes of Remember and Factual Knowledge using directed content analysis. Interviewee G made this determination based on what he perceived to be the relative ease of the task (lines 30-35; coded as both Complexity and Difficulty) and how he assumed students would answer it (lines 31-32; coded as Assumptions of Student Reasoning or Learning Context.)

Specifying the sorting for separating a sample of blood, Interviewee G said,

> "So if I go to the blood test, uh, I will have to, usually you need to analyze, even to evaluate, but, and you have a, you need here, some factual knowledge, procedural knowledge, some conceptual knowledge. I don't know, you need some metacognitive knowledge here for that, but so I will place, to uh procedural and evaluate, the box there, but I'm talking for myself. Now, if I was more knowledge about blood. Yeah, it could be, it could be, top left box as well. Well, for a microbiologist, you know, it's, uh, everyday routine."

Because Interviewee G concluded that this facet of item Q6 should be sorted as existing at the BCT intersection of Evaluate Procedural Knowledge (lines 37-42; coded as Interaction with BCT,) this excerpt was assigned the *a priori* codes of Evaluate and Procedural Knowledge using directed content analysis. Interviewee G made this determination based on his previous experience with Bloom's Taxonomy (lines 36-37; coded as Internalized Prior Taxonomy Experience,) which he also uses to remove metacognitive knowledge from consideration (line 39; coded as Elimination.) Interviewee G is clear that these are his personal sortings (lines 41-42; coded as Unconfidence) and that someone with different experiences might have a different sorting (lines 40-43; coded as Conditional.) For example, Interviewee G pointed out that depending on the familiarity someone has with separating blood (as in a microbiology student versus a microbiology professional) that this item may also fall into the upper leftmost BCT intersection of Remember Factual

Knowledge. As this was his sorting for the previous facet of item Q6, this intersection was already included in the sorting of interviewee G.

Specifying the sorting for separating unrefined petroleum, Interviewee G said,

> “Unrefined petroleum, uh, uh, depends how, how, how much you have understood, uh, the distillation, uh, procedure. I recall a chemist who was working in the state chemic- chemistry labs to whom we were sending…some of our students. So I sent this student to him, and, uh, when I discussed about students' performance and, uh, demonstrated knowledge, demonstrated difficulties here in his placement, in the chemistry lab, he said, ‘I found that these students cannot connect between distillation, as it used in physical chemistry, is that it's as it's used in organic chemistry,’ uh, he considers to be two different things. You know, you do distillation in physical chemistry for a different purpose, and you do that in organic chemistry for different purpose, you know, organic chemistry, you do that for, just for separating, but, you know, organic chemistry, you, you draw graphs, diagrams, and, uh, the student cannot make the connection between it. He thinks that two different things. So I have, you have a lot of conceptual understanding here, which applies also to the blood situation. You have concepts, uh, there, and you have a more complex mixtures, Blood, it's also, it's a biological material and there is the biology dimension to it.
>
> “If I am, unfamiliar with it. Sure. It's procedural knowledge. There you need understanding. You need to apply, uh, to analyze even to evaluate. I should think…Well, uh, that depends, depends on how familiar I am. If I take myself as a, because I'm been teaching physical chemistry laboratories during my career, uh, I should, uh, place that to the application, and to procedural again.”

Because Interviewee G concluded that this facet of item Q6 should be sorted as existing at the BCT intersection of Apply Procedural Knowledge, this excerpt was assigned the *a priori* codes of Apply and Procedural Knowledge using directed content analysis. Interviewee G made this determination based once again on the familiarity a student has with the mixture (lines 44-45 and 50-62; coded as Complexity,) and its necessary separation techniques (lines 53-57; coded as Assumptions of Student Reasoning or Learning Context.) Interviewee G also pointed out that depending on the familiarity someone has with the processes of distillation, that this item may also fall into BCT intersection of Evaluate Procedural Knowledge (lines 44-45 and 64-69; coded as Conditional.) As this was his sorting for the previous facet of item Q6, this intersection was already included in the sorting of interviewee G.

From these excerpts, interviewee G was triple coded as sorting item Q6 at the BCT intersections of Remember Factual Knowledge, Apply Procedural Knowledge, and Evaluate Procedural Knowledge, a visualization of which can be seen in Figure 4.7.

Interviewee H

Interviewee H began his sorting of item Q6 along the Knowledge Dimension (lines 1-2; coded as Interaction with BCT,) saying,

"Okay. So we have a multi-part question. So if we start with the knowledge levels my first instinct is because you're asking [the students] to propose a method, would be under procedural. So they would have to have a list of steps in which to do those, but they also have to be aware that those are all different. So they're going to have three in some sense, different methods. They wouldn't, well, not one would follow the same, so they have to be aware of that as well. So like I said, I would probably put it in the procedural knowledge, but then it would also float a little bit towards the conceptual, um, level. Cause they're not asked to think about what they're thinking about. So I wouldn't put it in the metacog- metacognitive. But yeah, I think procedural would be where I put it. And then, Because it's multi-part, I would probably put it in…there's three distinct and different mixtures. I would probably put it into apply cause they would have to um think through the process multiple times and understand and predict how those things are going to occur. Um again, floating a little bit towards the analyze, but probably if I was to classify this in one place, it would be procedural apply."

Because Interviewee H concluded that item Q6 should be sorted as existing at the BCT intersection of Apply Procedural Knowledge, this excerpt was assigned the *a priori* codes of Apply and Procedural Knowledge using directed content analysis. Interviewee H made this determination first considering the nature of the item (lines 1-3 and 9-12; coded as Item Forensics) and then thinking about what line of reasoning he assumed students would use when given this item (lines 3-7 and 14-16; coded as Assumptions of Student Reasoning or Learning Context.) Spreading out from the Procedural Knowledge category, he is able to remove metacognitive knowledge from consideration (lines 9-12; coded as Elimination.) He does recognize that the multi-part question might require multiple methods, (lines 1, 4-7, and 12-16; coded as Complexity) but he does not suggest that these methods exist at different BCT intersections instead that the process of thinking through the problem multiple times cause the item to require the cognitive process of apply. While Interviewee H does suggest that the item may not fit neatly in one intersection (lines 8-18; coded as Unconfidence,) he does say that if he had to pick one intersection, that he could narrow it down.

When asked about how he was interacting with the BCT to make these determinations, Interviewee H said,

> "So I'm primarily looking at the descriptions in the, in the tool um and thinking about how I would solve the problem and what, which of those descriptions most aptly applied to the way I would solve the problem. Probably, I would probably say both of that. That's what I'm looking at specifically the, the descriptions and not necessarily the, my history of Bloom's taxonomy."

From these excerpts Interviewee H was coded as sorting item Q6 as existing at the BCT intersection of Apply Procedural Knowledge, a visualization of which can be seen in Figure 4.7.

Interviewee I

Interviewee I began his sorting of item Q6 by deferring to his preferred method of sorting items by cognitive level (lines 1-4; coded as Internalized Prior Taxonomy Experience.)

> "Oh, cool. Ooh. Okay. This is interesting. So if you use the 3D language, just as planning, conduct investigations, which you almost never assess cause, oh my God. Can you imagine grading this? Oh, it'd be rough. Okay. I'm going to, I'm going to try to behave here.
>
> "As, as, was the case with your earlier example, it would definitely depend on what had happened in class. So if this was an example or something very, very similar to this, that would be an easy sort of jump for students to make then this might well be more in the you know, sort of recalling and reconstituting something you've heard before. But if you hadn't, if you'd only heard about like, basics of separation strategies, what is it?"

In this excerpt, Interviewee I first remarks on the difficulty of grading an item like this (lines 2-4 and 7-9; coded as both Difficulty and Complexity.)

Then, Interviewee I said that sorting this item along the BCT would depend (lines 6-12; coded as Conditional) on how the item was presented to students (lines 6-12; coded as Assumptions of Student Reasoning or Learning Context.)

Moving on to sort item Q6 along the Cognitive Process Dimension, Interviewee I said,

> “So you, if you're designing a car-, it looks like this is create-ish. If it's actually something they're having to plan and conduct, or maybe not conduct, plan from first principles, in which case it sounds like that would be in the ‘create’ den. That said, these aren't mutually exclusive.
>
> “Um, so, when you are planning an investigation, you are going to hopefully be evaluating your proposed investigation against other alternatives. You'll hopefully have a reason why you've chosen the thing that you have. So that certainly falls under this evaluate thing, ‘constituent parts and how they relate to one another.’ That's, there's a lot of overlap here. Okay. So if you think about how these things in a, in a mixture might be interfacing with one another, you may need to do that to consider what sort of reasonable separation strategy you would need. There's probably also things you are gonna end up recalling. I find this much easier to put in a box with the 3D stuff, but it's what I'm used to. So that could be why.
>
> “Um, so potentially it would be create, but I mean, there's the other stuff under the hood as well. I think, I don't think you can…’create a coherent or functional whole,’ if I simulate or draft, there's a lot there without looking for alternatives, without evaluating the reasonableness of those alternatives, based on your knowledge of the domain, maybe past experimental things, you've done, whatever.”

Because Interviewee I concluded that item Q6 required only the create cognitive process, this excerpt was assigned the *a priori* code Create using directed content analysis. Interviewee I came to this determination based on the lines of reasoning he assumed students would use when approaching this item (lines 13-28 and 32-36; coded as Assumptions of Student Reasoning or Learning Context) but displays uncertainty throughout the excerpt (lines 13-22 and 25-36; coded as Unconfidence) based on the amount of steps a student would need to undertake (lines 13-36; coded as both Conditional and Complexity.) Interviewee I cross-checks his instinctive sorting of item Q6 against both the Cognitive Process Dimension headers (lines 22-24; coded as Interaction with BCT) and against his preferred sorting dynamic (lines 15 and 28-29; coded as Internalized Prior Taxonomy Experience)

Considering the Knowledge Dimension, Interviewee I said,

"But the knowledge, well, I mean, I said, see it said 'specific to chemistry.' I wonder what, how specific they have to be. Certainly, if you are actually proposing a method that a human would need to, do you need to have some knowledge of how these sorts of procedures interface? I don't know how this, I mean, the reason why nobody assesses these sorts of things, um, except for in very scaffolded ways on labs, at least as far as I've seen is that if you ask something like this to a lecture, students are going to be like, uh, 'either I repeat the thing that you've told me before, or I'm going to sort of make something up from something I've experienced in the past' and so their, their procedural knowledge is going to be potentially problematic. If they've never done anything like this before, to relationships between basic components in this mixture, blood, my god, this it'd be terrible to separate.

"I mean, I suppose you have to think about if you're going to talk about solubility, to what extent do you need to justify your proposed method? Because I think if you're just proposing and not justifying, this is kind of not the most useful question because the meat of the proposition is actually the criteria you're using to decide why one method is more suitable than all other alternatives or than the other alternatives you've thought of. It would seem to me- Oh, I, I mean, I would think that you would be connecting up knowledge elements and if that's what ‘conceptual knowledge means,’ I would think that you would be doing that, if you are evaluating some methodological proposal against other alternatives and thinking about the constraints and affordances of that method. I mean, if you're talking about some sort of separation protocol, you're thinking about, I mean, something to do with how these things interact with one another and how you might tease them apart.”

Interviewee I begins this excerpt by questioning the rigidity of the BCT (lines 37-38 and 58-61; coded as Interaction with BCT,) then expands on why this item might be complicated to grade except in very scaffolded situations (lines 38-43 and 48-66; coded as both Complexity and Difficulty) even in his preferred sorting dynamic (lines 41-43 and 54-55; coded as Internalized Prior Taxonomy Experience) even giving a couple of potential answers that students might provide (lines 38-48 and 54-64; coded as Conditional) that, according to him, make this a less useful item for measurement. Because Interviewee I tentatively (lines 37-66; coded as Unconfidence) concluded that item Q6 required only conceptual knowledge, this excerpt was assigned the *a priori* code Conceptual Knowledge using directed content analysis. Interviewee I made this determination based on the lines of reasoning in which he assumed students would engage when tasked with the item (lines 38-66; coded as Assumptions of Student Reasoning or Learning Context,) and what the item is asking students to do (lines 49-50; coded as Item Forensics.)

Asked if he thought each of the three tasks of item Q6 might require separate cognitive processes, interviewee I admitted he had misread the item (lines 67-70; coded as Unconfidence,) saying,

"You know, and I didn't even read that right. Because I thought [the blood, unrefined petroleum, and iron-sulfur mixture] were all together. And I was like, 'why, why have you put blood in the oil? Is this some sort of weird analogy?' Okay, this makes sense. Now, I'm sorry, it's been a day. That's just, really it really depends what you want here. I mean, if like, if you just want recall of a method that they heard, 'separate something pretty close to blood,' and so they're just going to be like, 'yeah, use HPLC.' Then that's not all that useful. If you want to think about the constraints and affordances of different methods and why they work the way they do, which I would argue is a whole lot more useful, then that's going to have to be much more scaffolded and that's going to definitely bump it up the scale in terms of levels. So sure. I mean, I don't, but were you asking, is it necessarily going to be at a different level depending on the course? Cause I don't think that's the case. I don't know. It would depend on how they frame the question, and it would depend on their general chemistry learning environment, I expect. So for example, our environment here is really, you know, even though we've got active stuff going on, it still assesses dumb shit. So I mean, they would probably be looking for some sort of definition, cause they've got like hard-, soft-mass spec, and like the first unit, and they, kids just look at lines on paper and tell them what they want to hear. There's nothing at all about why you would have a particular interpretation over other alternatives. And so if they hadn't heard something that would come to mind when they were thinking of separating, you know, iron and sulfur, although you would think it is just, we're talking about shavings, powder, big chunks, big chunks is easy, you've got an intuitive resource there. So yes, that would not be problematic if you're thinking about separating blood or unrefined petroleum. Then if you're talking about [our university,] the students are probably going to give you gobbledy-goop and not be willing to try to connect knowledge elements to explain how and why a particular protocol might work, cause that's all they've ever seen, and all they've ever been assessed on. It's not a function of the students. I think it's a function of the emphasis of the environment."

In this excerpt, Interviewee I discusses that the sorting of the item necessarily depends on what the instructor is hoping to measure (lines 71-78 and 81-83 and 89-95; coded as Conditional,) as well as the classroom environment, and that making the item more useful for measurement would require more scaffolding on the part of the instructor (lines 72-83 and 91-93; coded as both Complexity and Internalized Prior Taxonomy Experience.) Again, he provides a couple of scenarios of ways students might approach the item (lines 73-74 and 81-100; coded as Assumptions of Student Reasoning or Learning Context).

From these excerpts Interviewee I was coded as sorting item Q6 as existing at the BCT intersection of Create Conceptual Knowledge, a visualization of which can be seen in Figure 4.7.

Interviewee J

Interviewee J began her sorting of item Q6 along the Cognitive Process Dimension (lines 1-2; coded as Interaction with BCT,) saying,

"Hmm. This is a good question. Um, okay. So in terms of knowledge dimension, I think that, um, this is, well, it draws on a couple of things, um, but there's, uh, an argument for procedural knowledge here because the, um, the student is asked to kind of like, visualize what that student would do in a lab setting. And so then, understanding how, like the actual skills, sort of the tools that would be available are, uh, is, is really necessary. Um, I think also that there's going to be some factual knowledge about what these, what is included in these mixtures or, or the student need to gain that in order to be able to answer the question and also some conceptual knowledge about how to apply things like, um, intermolecular forces or a difference in boiling points or difference in other properties in order to be able to choose the right method. So that one is really complicated actually. Um, but I can see, I can see arguments for procedural, conceptual, or factual."

Because Interviewee J could see arguments for any of the first three knowledge dimension categories (lines 3-15; coded as Unconfidence,) this excerpt was deductively coded as containing the *in vivo* codes Factual Knowledge, Conceptual Knowledge, and Procedural Knowledge. She made this assumption primarily on her assumptions of what knowledge this item would elicit from students (lines 5-13; coded as Assumptions of Student Reasoning or Learning Context,) and how many things (lines 2-14; coded as Complexity) a student is being asked to keep track of at a time (lines 4-5; coded as Item Forensics.)

Moving on to sort item Q6 along the Cognitive Process Dimension, Interviewee J said,

"Um, in terms of the, um, cognitive processes, I would, I, I would code this as a create question, um, because students are asked to devise a method or develop a method, uh, propose a method, what is the actual word? And so there, they have to sort of start from, from nothing. Um, yeah, I'm sort of thinking about like, if there, if, if there were more context, if, you know, if the instructor had sort of already explained it's, it's hard for me to believe that the instructor sort of has already explained how to do all three of these things, given that this is this multi-part question. Um, but in that case, if we had talked about how to do very similar things, then again, it would be something like an apply question, but I think the strongest argument here is create."

Because Interviewee J sorted item Q6 as requiring either apply or create cognitive processes depending on her assumptions of how this item would be presented to students (lines 19-25; coded as Assumptions of Student Reasoning or Learning Context) and what the item is actually asking of students (lines 17-19; coded as Item Forensics,) this excerpt was deductively coded as both Apply and Create. However, Interviewee J's uncertainty was coded as Unconfidence (lines 16-26.)

Asked if she thought each of the three tasks of item Q6 might require separate cognitive processes, interviewee J said,

> "In each case, the student is asked to propose a method. And so that, I would argue is kind of fundamentally create, like that is a create-level process. Um, again, whether it's better defined as a lower-level might depend on the student's familiarity with any of those mixtures. Um, but it doesn't change it very much for me because yeah, again, the, the, the task is still the same."

In this excerpt, interviewee J posits that the task of each part of the item is the same (lines 28 and 33; coded as Item Forensics,) and therefore all require the same cognitive process (lines 28-30; coded as Internalized Prior Experience,) unless the student were very familiar with the mixture (lines 30-32; coded as Conditional) possibly because of how the material was presented to them (lines 29-31; coded as Assumptions of Student Reasoning or Learning Context.)

Asked if she thought each of the three tasks of item Q6 might elicit different Knowledge Dimension categories, interviewee J said, “That's a good question. Um, that would depend on the context, I think, um, because the, so for example, I'm, I'm sort of picturing this being as part of an analytical chemistry course, or maybe because of the blood example, a biochemistry course. And so then, um, where the students may like, there, there may be more context here where it's like, ‘Oh, well, this is sort of obviously the method,’ like if the procedure is kind of already chosen for you then, um, that would change whether or not the student had to provide that. Um, but just sort of in isolation, imagining that, you know, you turn to that page on your exam: ‘propose a method to separate unrefined petroleum,’ um, you're still going to have to think about what procedures are available to you. And so then, that does- that doesn't change that part. And you're still going to have to think about the components of those mixtures, and you're still gonna have to think about what connects those in order to come up with a good choice.”

In this excerpt, interviewee J says that a difference in knowledge levels would depend on what kind of course this item was presented in (lines 34-48; coded as Assumptions of Student Reasoning or Learning Context.)

From these excerpts interviewee J was double coded as sorting item Q6 at both the BCT intersection of Remember Factual Knowledge and of Apply Factual Knowledge, a visualization of which can be seen in Figure 4.7.

Item Q7

Item Q7 was sorted by the 6 interviewees C, E, L, M, N, and O. Item Q7 was the only question sorted by interviewee O, the first question sorted by interviewee N, the last question sorted by interviewees C and M, and a middle item for interviewees E and L.

Interviewee C

Interviewee C began her sorting of item Q7 along the knowledge dimension, saying,

> “Okay. So let's see, look at [the tool] again. So this, uh, this is looking pretty conceptual to me. Okay. Um, because you know, it's talking, so they have to have that knowledge of, um, like you said, how things dissolve, why did it dissolve faster or slower, or what, what, what factors lead to, um, dissolving. Um, so I think it's a conceptual is, is part of a theory, you know, uh, or an idea of how in ca- how how chemicals work.”

Because Interviewee C concluded that item Q7 required only conceptual knowledge, this excerpt was assigned the *a priori* code Conceptual Knowledge using directed content analysis. After consulting the tool (line 1; coded as Interaction with BCT,) Interviewee C came to this determination based on the knowledge she assumed a student would have access to when presented with this item (lines 3-7; coded as Assumptions of Student Reasoning or Learning Context,) but her indecision was coded as Unconfidence (lines 2-3 and 5-6.)

Moving on to sort item Q7 along the Cognitive Process Dimension, Interviewee C said,

> “Um, and in terms of the, uh, cognitive process, this looks like understand, um, because it had to describe or explain, um, and let me see. Yeah. I don't see the apply so, I think it's yeah, yeah, I think it's understand. Yeah, because I don't see the apply coming in with, based on the definitions here, then they don't have to calculate anything or execute anything. Um, and under analyze. So can [an item] can be understanding and analyze skipping over apply?”

Because Interviewee C sorted item Q7 as requiring either understand or analyze cognitive processes, this excerpt was deductively coded as both Understand and Analyze. Interviewee C started her sorting of item Q7 by keying into a verb listed in both the tool and the item (lines 12-15; coded as Pattern-Matching) before expanding her consideration to the cognitive processes of apply and analyze. Interviewee C removed the cognitive process of apply from consideration (lines 10-13; coded as Elimination) by further examining the cognitive process dimension headers (lines 11-13; coded as Interaction with BCT) but she was unsure of one's ability to code at non-adjacent intersections (lines 8, 10, and 13-14; coded as Unconfidence.)

After being told that she does not have to use adjoining intersections, Interviewee C expanded on her choice of the cognitive process of analyze, saying,

> "Okay. Cause I was like, oh, you have to be in this particular, um, range. But, um, there's also, I think, analyze in [the item] as well, because, um, if you analyze part talks about, um, comparing and contrasting, and, but this [item the students] have to compare, they have to look at the cube sugar compared to table sugar, look at the temperature and see what's the difference between those two? What makes them different? How would that change? Whether it, it dissolves faster or slower. So I think there's some comparison analysis going on. So understanding and, and, and, and, analyze."

Interviewee C bases her sorting of item Q7 into the cognitive process of analyze based on what she thinks the question is asking (lines 18-22; coded as Item Forensics,) and on the cognitive process dimension headers (lines 15-18; coded as Interaction with BCT.)

Moving right along the BCT, interviewee C considered the cognitive process of Evaluation, saying,

> "Oh, I did not go up to evaluate. Um, So even though the word 'justify' is under evaluate, [the students are] not asked to really justify the answer. They just asked to just say, you know, I guess you circle what answer is why, there's no reason, you don't have to give a reasoning for why they're doing it. So I don't think it would be evaluate necessarily. But yeah. Um, it's not create, so yeah. I will go with, you know, some understanding and also analyze.
>
> "Yeah. And, and, and do you think, I guess in my mind, is, does the question have to be, um, have a, you know, that hierarchy like it has to have, you know, it had the understand component, can you skip over apply and have an understanding and analyze is that possible or, you know, there's a higher level, but you skipping over?"

Once again, Interviewee C removed the cognitive process of evaluate from consideration (lines 25-29 ; coded as Elimination) by further examining the cognitive process dimension headers (lines 24-25 and 32-36; coded as Interaction with BCT) and what she thought the item was asking (lines 25-28; coded as Item Forensics,) but she was unsure of one's ability to code at non-adjacent intersections (lines 25-26, 28-29, and 32-36 ; coded as Unconfidence.)

Following up on that question to enquire whether all items should therefore be coded as including Remember Factual Knowledge, Interviewee C went on to say, "Yeah. Cause I mean, I know no once I said I did not include remember, but you know, I do tell my students, I say something you just can't, you can't skip over. I mean, if you need to learn the amino acids, you just got to learn them. I mean, I can't because you can't move on without it. So you just have to know that, you know, basic remembering part, but I don't know, you know, the questions that we looked at. Um, I like even that first question, um, [item Q1] there's some remembering there. Um, but I didn't, I didn't, I guess I didn't think about that at that point, but yeah, I think, yeah, there's some remembering going on because there's a basic thing that you need to know before you could even think about analyzing or, you know, um, and even, I don't know, that there's a big remember factor in the, in the, in the last question though, but, um, because it's not something I think you have, but I think in any, in all of the problems, there's probably some small level of remembering going on."

When asked if she was sorting based on the highest or the minimum cognitive process which a student might need to answer the item, interviewee C said,

"Okay. Yeah. Okay. Yeah. So you think what? The highest, yeah, I was still still stick with what I said then. Okay."

From these excerpts interviewee C was double coded as sorting item Q7 at both the BCT intersection of Understand Conceptual Knowledge and of Analyze Conceptual Knowledge, a visualization of which can be seen in Figure 4.8.

Interviewee E

Interviewee E started her sorting of item Q7 by looking at what the item was trying to elicit (lines 1-6; coded as Item Forensics.)

"Okay. So I guess before I'm even looking at the, back at the taxonomy, um, so you're given some information that you have to

> process and understand and begin with, um, equal amount of water. Okay. So something is held constant. The same amount of sugar is added. Okay. So sugar and water is constant, um, in which, okay, so it's the form of the sugar and the temperature in the thing."

From here, Interviewee E focused on the Knowledge Dimension, saying,

> "There's two things that seem to be in [the item] that, um, so it's beyond factual. There's some facts in here, but it's, it's somewhat conceptual. Um, you, uh, you have to understand dissolution, I guess, in a conceptual way or relating those things, but it does really feel somewhat procedural to me. Um, because you're, you're, it's kind of a problem. And so you're going to think about what are the, you know, there's a process of, do I have to understand what do I have to understand to solve this problem? Are there equations I can use? Are there theories or facts or things that I need to remember to solve this problem? Um, it's not metacognitive, so I'm going to go with procedural on this one, um, and thinking about it that way."

Because Interviewee E concluded that item Q7 required only procedural knowledge, this excerpt was assigned the *a priori* code Procedural Knowledge using directed content analysis. Interviewee E came to this determination based on what she thought the item was asking students to do (lines 7-8 and 11-13; coded as Item Forensics.) She was able to remove the knowledge dimension category of factual knowledge from consideration (lines 7-8 and 16; coded as Elimination) based on the number of tasks which students would have to accomplish to finish the item (lines 7-10; coded as Complexity,) and she decided between conceptual and procedural knowledge based on the reasoning with which she assumed students would approach the item (lines 9, 11-16; coded as Assumptions of Student Reasoning or Learning

Context.) Throughout this excerpt, Interviewee E used her previous experience of Bloom's taxonomy to assist her (lines 7-8, 10-13; coded as Internalized Prior Experience,) and displayed little uncertainty in her sorting (lines 9-12 and 16-17; coded as Unconfidence.)

Interviewee E next moved across the cognitive process dimension (lines 18 and 20-22; coded as Interaction with BCT,) removing cognitive processes from consideration (lines 18-19 and 24-29; coded as Elimination) as she decided the item required more and different cognitive processes.

> "Um, remember, you do have to remember some things, but I don't think it's a remember thing. You certainly have to understand things. You certainly have to apply things. Um, and, um, let me reread and, um, analyze to think about that. So 'compare, contrast, deconstruct, parse, distinguish, reconstruct, um, attribute compare.' I mean, there's a comparison here. You're comparing temperatures and you're comparing, but it, it, I don't know if it, I mean, it's a little bit of analyze, but it's not analyze in the same way that 'analyze a table' is. So I don't know if it's quite analyze and it's not evaluate, you're not being given a solution and saying, 'is this correct?' And it's, doesn't, it's definitely doesn't feel like create to me because the options are all given to them. They're not, it's not open-ended and it's not, um, so to me it feels like procedural apply. Um, so I'm assuming that they would given, be, given somewhat similar problems or seeing something in their textbook in class that's somewhat related. Um, and, and maybe it's extending a little bit from what they've seen, but that's what it feels like to me."

Because Interviewee E concluded that item Q7 required only the apply cognitive process, this excerpt was assigned the *a priori* code Apply using directed content analysis. Interviewee E made this determination by considering what the item was asking students to accomplish (lines 22-29;

coded as Item Forensics,) and what lines of reasoning she assumed students would undertake when presented with this item (lines 18-20, 23-24, and 30-34; coded as Assumptions of Student Reasoning or Learning Context.) Interviewee E discussed how one could pattern-match the wording of the item with the BCT Cognitive Process of Analyze (lines 20-26; coded as Pattern-Matching,) but dismissed this. Interviewee E continues to rely on her prior experience with Bloom's taxonomy (lines 18-20 and 22-34; coded as Internalized Prior Experience) but does express some uncertainty in the context of this sorting (lines 18-19, 24-34; coded as Unconfidence.)

From these excerpts Interviewee E was coded as sorting item Q7 as existing at the BCT intersection of Apply Procedural Knowledge, a visualization of which can be seen in Figure 4.8.

Interviewee L

Interviewee L also started his sorting of item Q7 by looking at what the item was trying to elicit (lines 1-4; coded as Item Forensics.)

"So the question is asking in which container is dissolution fastest. Um, and it's saying that they all contain the same, uh, the same amount of water, and they're adding the same amount of sugar, just in different, um, physical states and at different temperatures."

From here, Interviewee L likewise pivoted to the Knowledge Dimension. As he worked his way up the Knowledge Dimension, he used the heading descriptions (lines 5-7; coded as Interaction with BCT) to remove metacognitive and procedural knowledge from consideration (lines 5-10; coded as Elimination.)

“Okay. Um, so again, metacognitive, nope. They're not being asked to think about their thinking, um, procedural, ‘the skills specific to chemistry and the criteria for appropriate practice.’ This is not a procedural question, per se. This seems to be conceptual again. Um, this isn't a, what what's, mm they're not being asked to anything regarding certain steps. They're asked what uh container i- in which container is dissolution the fastest. So they're being asked a conceptual question about, um, a process, uh, in chemistry, um, factual, um, students, most the fundamentals, this a fundamental, this is definitely a conceptual question. Um, um, the relationship between solubility, surface area, and temperature, um, would be what this is testing that, that relationship.”

Because Interviewee L concluded that item Q7 required only conceptual knowledge, this excerpt was assigned the *a priori* code Conceptual Knowledge using directed content analysis. Interviewee L mainly compared what he thought the item was asking of students (lines 5-16; coded as Item Forensics,) to the Knowledge Dimension category headings to triangulate his sorting, but did show a little uncertainty in his determination (line 8; coded as Unconfidence.)

Using the same process, Interviewee L moved leftward along the Cognitive Process Dimension to sort item Q7, saying,

“Um, so now, which cognitive process, so conceptual again, uh, cognitive, uh, cognitive, uh, processes, let's start with create, they're not being asked to create anything, evaluate, um, it's not, they're not necessarily being asked, uh, to, um, ‘students are asked,’ heck this is going to bother me, ‘students are asked to, based on criteria and standards, judge assess assign value, critique, detect, examine, test check, discriminate, justify, recommend, appraise or grade,’ um, there ‘in which container is dissolution the fastest.’ They are being asked to, um, examine, which would be assess, which would be the fastest.

“Um, so I think that's an evaluate, but let me just look through these again, they're not being asked to simply recall something, they're not being asked to, um, clarify, you could say ‘conclude,’ but I don't think that's the meaning of understand, ‘to conclude,’ um, that one of, you're you're trying to differentiate here. You're not trying to conclude something.

“Um, I'll leave that, um, ‘students are asked within a given context to employ, calculate, carry out, change conduct, execute, implement, infer, modify, operate, predict, use y-, uh, utilize or substitute.’ Um, this could be a ‘predict in which container is dissolution the fastest.’ They're predicting which one would be the fastest based on the application of, uh, understanding the relationship between these concepts. Um, analyze ‘students are asked to break material into its constituent parts.’ Um, it's already broken up to me. Um, so they're not being asked to break up anything they're trying, they're asked how things ar- based on how things relate to each other, which would be the fastest. Um, so I- the conclusion is not actually applying something to come to an answer, they're, they're, they're applying something, um, uh, a set of concepts, but then evaluating that interrelationship. So again, I think I would stick with, uh, conceptual knowledge evaluate.

“I, there might be, I, there is some apply in there, but it's a question of does, does evaluate no- ‘based on criteria and standards.’ So it's based on some set of standards, these standards being that lower surface area and higher temperature create a faster rate of dissolution. Um, that's the criteria and they're being asked to, uh, ‘assign, detect discriminate,’ discriminate, I think is the key word they're discriminating between these things to determine which would be, uh, uh, the fastest. So I think this is a pretty solid evaluate.”

Because Interviewee L concluded that item Q7 required only the evaluate cognitive process, this excerpt was assigned the *a priori* code Evaluate using directed content analysis. Once again, interviewee L mainly compared what he thought the item was asking of students (lines 18-20, 24-

25, 28-29, 37-46, and 49-54; coded as Item Forensics,) to the cognitive process headings (lines 18-24, 27-37, 40-41, and 50-55; coded as Interaction with BCT) to triangulate his sorting, removing the cognitive processes of remember, analyze, and create from consideration (lines 18-20, 28-32 and 41-42; coded as Elimination.) Once again, he did display a bit of uncertainty in his sorting (lines 20, 27, 29, 37, 46-47, 49, 54, and 56; coded as Unconfidence.)

Asked to follow-up on his comments about the BCT Cognitive Process Dimension headers, Interviewee L said,

> "Um, they may be using con- they're using conceptual knowledge to discriminate. Um, I think it's really easy and this is why I have issues with Bloom's taxonomy. It's so easy to say, 'well, they're applying their understanding of this concept in order to evaluate,' but ultimately this question, the main point of this question is asking students to discriminate between these different samples to determine which would be, which would dissolve the fastest.
>
> "Okay. Conclude, um, 'students are asked to clarify, conclude.' So I think in the moment I was like, 'well, conclude if you're coming to an answer you're concluding, that's the answer.' So I, I, the-, the word 'conclude' in my mind needs to be operationalized a little bit better. Um, so let me reread this question th- this definition 'understand: students are asked to clarify, conclude, describe, explain, generalize, illustrate, interpret, paraphrase, represent, restate, summarize, symbolize, or exemplify in order to draw meaning from,' um, so I should have finished reading the part 'in order to draw meaning from,' so 'students are asked to,' if we just plug in the one word, 'students are asked to conclude in order to draw meaning from,' and I'm assuming that's a correct way to read this definition. Okay. Uh, so 'students are asked to conclude in order to draw meaning from,' I, I, I, this might just be me and my never-been-good-at-English brain. I, that doesn't make a whole lot of sense to me, that sentence 'to conclude,' 'students are asked to conclude in order to draw meaning from,' and when I think of the definition for conclude, I think, uh, of of come to a, a, a final understanding. I'm trying to not use, 'conclude,' um, trying to

come to some, some final understanding of something or a final representation of something. Um, so, so I think all questions, you, you come to a conclusion, so you conclude something, um, or your conclusion could be just that you, you, you don't know, but that is a conclusion. It's what you're coming to at the end of something. So I read that as being applied to all of these, if you conclude something."

In this excerpt, Interviewee L more closely examines the cognitive processes of understand and how it relates to the verb "conclude" echoing similar thoughts about this heading as Interviewee I did for item Q6.

From these excerpts Interviewee L was coded as sorting item Q7 as existing at the BCT intersection of Evaluate Conceptual Knowledge, a visualization of which can be seen in Figure 4.8.

Interviewee M

Interviewee M likewise began her sorting by considering the nature of the question (lines 1-2 and 5-17; coded as Interaction with BCT) saying,

"You know, that that item did not end the way I thought it was going to. Okay. Alright. Um, the same amount of sugar is added to, which is, Is? I'm like very nervous to posit what I think is the correct answer here, but I'm just going to not be nervous because whatever. Would we think, uh, cause I can't classify it, if I have no idea what it's targeting. Would we argue that there are like two factors we're considering in dissolving right now, which would be like surface area-to-water and then temperature? Okay. So we would say powdered sugar at 10 degrees. Okay, cool. Cool. Please don't tell my students that was hard as it was. Okay. So I, uh, that's good. I like being surprised by things and that question surprised me. I think I- I think the equal amounts of sugar had me thinking like, well, I don't know, maybe something with like saturation or something with energy, cause we talk about energy or really any time we talk about sugar, it's usually to contrast with, um, something that's going to be electrolytic in solution, and so I don't know the speed, like they got, you got me. Alright. So I would say that it, now that I reflect, I like it's kind of spicy."

In this excerpt, Interviewee M discusses her uncertainty of how to answer the item (lines 3-10; coded as Difficulty) and that sorting the item is not possible without it (lines 5-9; coded as Unconfidence.)

Moving on to consider the Knowledge Dimension, Interviewee M said,

"I would still say that it's conceptual knowledge, cause it's, requiring them to think about factors that are gonna affect the solution."

Because Interviewee M concluded that item Q7 required only conceptual knowledge, this excerpt was assigned the *a priori* code Conceptual Knowledge using directed content analysis.

Then Interviewee M considered which Cognitive Processes the item is asking students to perform (lines 20-21; coded as Item Forensics,) saying,

"Um, I would probably put this into apply because it's asking them to make a prediction. Um, I think it would probably only work in apply in like a unit where we've spent a bit of time talking about solutions. I could see it being like a little higher than that if we did it at the beginning of a unit on solutions where we were having them just try and use prior knowledge, um, to think about it, where, like, they could probably draw on whatever. Um, then I could see that being a little bit higher, like analyze or evaluate, where they had to make a prediction without necessarily having the knowledge to do so. Um, but after a unit on this, um, solutions and dissolution, I would say apply probably makes the most sense here."

Because Interviewee M sorted item Q7 as requiring any of the cognitive processes apply, analyze, or evaluate (lines 20-30; coded as Conditional) depending on the way she assumed the item might be presented to students (lines 21-29; coded as Assumptions of Student Reasoning or Learning Context,) this excerpt was deductively coded as containing the *in vivo* codes Apply, Analyze, and Evaluate. Interviewee M's use of uncertain language was coded as Unconfidence (lines 20-23 and 26-30.)

From these excerpts Interviewee M was coded as sorting item Q7 as existing at the three BCT intersections of Apply Conceptual Knowledge, Analyze Conceptual Knowledge, and Evaluate Conceptual Knowledge, a visualization of which can be seen in Figure 4.8.

Interviewee N

Interviewee N began her sorting of item Q7 along the knowledge dimension, saying,

> "Okay, so let's see. I'd say, powdered sugar at 10° C, that would be my guess, um educated guess. So then, what am I doing here? I guess this is probably somewhere between conceptual and factual."

Because Interviewee N concluded that item Q7 required either factual or conceptual knowledge, this excerpt was assigned the *a priori* codes Factual Knowledge and Conceptual Knowledge using directed content analysis. After answering this item for herself, Interviewee N based her sorting on how she was interfacing with the item (lines 2-3; coded as Unconfidence.)

Moving on to the Cognitive Process Dimension, Interviewee N said,

> "And then it's probably much more of a kind of apply/analyze sitch than [item Q8] because we have to kind of consider…Well, it also relies on some kind of implicit knowledge about what's the difference

between cubed, table, and powdered sugar, um which could be…But assuming that they have, that they bring that knowledge in with them, then they have to kind of think like, what's the smallest, or the, like, they probably wouldn't think of in this these terms, but the, like a, surface-area-to-volume ratio. And if we have the highest surface-area-to-volume ratio, that's w– in powdered sugar, and then at the higher temperature, then we'll see the dissolution happen much more quickly. But they would probably think about it much more from the point of view of, 'what's smalles–' just like particle size, and maybe an interface between particle size and the solution. Um, so anyway, yeah, I'd say they're uh applying or analyzing factual/conceptual knowledge. Yeah."

In this excerpt, Interviewee N suggests that item Q7 exists at the interfaces between the cognitive processes apply and analyze as well as between Factual Knowledge and Conceptual Knowledge, though she reconsiders the Knowledge Dimension in the next excerpt. Because Interviewee N concluded that item Q7 existed at the interface between the cognitive processes of apply and analyze, this excerpt was assigned the *a priori* codes Apply and Analyze, using directed content analysis. Interviewee N sorted item Q7 as existing at this interface based on the reasoning she assumed students would undergo when tasked with this item (lines 7-16; coded as Assumptions of Student Reasoning or Learning Context,) and on what the item was asking students to accomplish (lines 5-7; coded as Item Forensics.) Interviewee N's uncertain language was coded as Unconfidence (lines 4, 9-10, and 14-17.)

Asked what she considered to be the difference between the interfaces of apply versus analyze and factual knowledge versus conceptual knowledge, Interviewee N explained, "Hmm. Maybe between apply and analyze, analyze is maybe more of a novel situation than apply, maybe the context is more familiar, in an apply situation versus an analyze situation. And then between factual and conceptual…Yeah, now that I'm thinking about this, this might be much more, more on the end of conceptual, than factual. Like, factual's just like, yeah. Term– like, you have 'terminology, technical definitions, specific details,' versus conceptual, 'interrelationships, classifications, categories, principles, generalizations.' So, here it's kind of we're like, applying a principle. Yeah, I said it: 'applying a principle.' Maybe it's, um yeah. So at, that's what I would say, like, factual's just like, a fact without any, without necessarily any background understanding to the fact, or how that fact relates to other facts. Um, and then a c– conceptual knowledge would be understanding more of the connections between disparate facts or generaliz– generalizing a trend, understanding the generalization of a trend. And then apply versus analyze, I'd say, yeah, kind of, how, how familiar is the context. Yeah."

In this excerpt, interviewee N considers her previous experience with learning taxonomies (lines 19-22; coded as Internalized Prior Experience) with the language of the BCT (lines 23-26; coded as Interaction with BCT. This leads her to remove factual knowledge from consideration (lines 22-23; coded as Elimination) based on pattern-matching (lines 26-28) her assumptions of how students would interact with the item (lines 26-27; coded as Assumptions of Student Reasoning or Learning Context,) with the language of the BCT. Because Interviewee N concluded that item Q7 required only conceptual knowledge, this excerpt was assigned the *a priori* code Conceptual Knowledge using directed content analysis.

Asked where she learned these lines of distinction, Interviewee N explained some of her previous experience of Marzano’s taxonomy (lines 36-34; coded as Internalized Prior Taxonomy Experience.)

> “I'm probably getting that from some of the taxonomy reading that I've done, I think. And I can tell you for sure, that my distinction between facts and concepts here, I'm remembering what I read in the Marzano’s book and, because they make a big, they make a clear distinction, between a fact being, just like a thing you remember and just straight up memorization without any understanding of what it really means, but you just know the answer to the thing, versus a concept is as a bit more connection-based.”

From these excerpts Interviewee N was double coded as sorting item Q7 as existing at the BCT intersections of Apply Conceptual Knowledge and Analyze Conceptual Knowledge, a visualization of which can be seen in Figure 4.8.

Interviewee O

Interviewee O started her sorting of item Q7 by using her preferred sorting dynamic (lines 7-27; coded as Internalized Prior Experience,) saying,

> “Okay, so to the following containers that contain each an equal amount of water, the same amount of sugar is added. I would just, that, that's a hard, that's a hard one to wrap my head around. I think I would, of me with my syntax, I would have to say, ‘so each of the following con–, containers, they've got an equal amount of water. The same amount of sugar is added to each one. In which container is distillation the fastest?’ Okay, so, I’m sort of having to envision a series of containers with same water, and same amount of sugar, and then, I'm looking at temperature, seems to be one of the uh key variables that were uh distinguishing. And then also the, the type of sugar, the form of sugar that we're putting in, um so when I'm looking at this, I'm thinking about um kind of connecting to conceptual

> knowledge about temperature and um how quickly uh molecules are moving and potentially they're colliding and um could result in dissolution. So I'm gonna have to connect to, connect to that knowledge, um I think I'm, mm, for me, I'm not sure what I'm connecting to you in terms of the, um, you can see like the, the the form of, the chemical form of the sugar, in terms of if it's in its powdered state versus a table already, in a–. So I think for me, I don't understand maybe the difference between the powdered and the table other than I think the powdered might have more surface area if it is in a more aerated state, or more 'fluffy' for a very scientific term versus a cube where you're compacted and um a lot, much less surface area, surface-to-volume ratio is is uh more, less surface for volume. So so I think I'm, I'm connecting to both of those of concepts and then I'm trying to make a prediction about um which of those is going to dissolve the fastest. Um, so would you want me to like, connect now with your, with your table?"

Interviewee O had extensive experience with Bloom's original taxonomy. In this excerpt, she walks through how she sees the item (lines 1-7; coded Item Forensics) and how she would attempt to answer it (lines 10-16 and 24-27; coded as Complexity.) Interestingly, Interviewee O does all this from a personal perspective instead of imagining what a student might do, or how it might be presented to a class, including her own difficulties (lines 1-7 and 16-24; coded as Difficulty) and uncertainties (lines 1-7, 16-24, and 27-28; coded as Unconfidence.) Once she had a sorting she was comfortable with, she switched to focus on the BCT.

Starting with the Knowledge Dimension, Interviewee O said,

> "Um, so, so I'm thinking in terms of…I'm definitely, in terms of the knowledge dimension, I think we've would, definitely at conceptual knowledge because we're thinking about those general principles of uh heat and surface area um, I guess I always find this a bit of a weakness for me like in terms of the procedural knowledge…I guess

34 I'm not sure whether there's any procedural knowledge required to
35 address this question, if there’s any practice, it seems like, um. Would
36 you consider the procedural knowledge also just like understanding
37 like, how you mix things together or how you um, like, sort of being
38 able to visualize this you know in a lab? It doesn't seem like that
39 would really be required for this?”

In this excerpt, Interviewee O is drawn to the knowledge dimension categories of conceptual knowledge and procedural knowledge, based on how she assumes students might interface with the item (lines 30-32; coded as Assumptions of Student Reasoning or Learning Context,) but she is unsure (lines 32-39; coded as both Unconfidence and Difficulty) of the bounds of procedural knowledge (lines 33-39; coded as Interaction with BCT.) After her question was answered, interviewee O continued,

40 “Okay. Okay. Yeah. So I guess I don't think that this really would
41 require procedural knowledge to me, and in terms of our cognitive
42 processes, to me I would, like I would probably go with apply um as
43 I'm thinking about, I'm having to take those concepts and then make a
44 prediction about what would happen under these different conditions
45 and thinking about um how those, that interplay, um so they have to
46 already, they have to sort of remember some of the terminology be able
47 to understand these concepts of um how heat impacts um the solution
48 and form. But then they have to apply that to a situation that they
49 might not have seen before. So I think that's where I would probably
50 be somewhere in the apply and conceptual knowledge.”

Based on her new understanding, Interviewee O removed procedural knowledge from consideration (lines 40-41; coded as Elimination.) Because Interviewee O sorted item Q7 as existing at the BCT intersection of Apply Conceptual Knowledge this excerpt was deductively coded as containing the *in vivo* codes Apply and Conceptual Knowledge, using directed content analysis. Interviewee O came to this determination based on how she assumes students might interface with the item and how it might be presented to them (lines 43-49; coded as Assumptions of Student Reasoning or Learning Context) as well noting as all the things students have to keep in mind in order to answer the item (lines 45-49; coded as Complexity.) Her uncertain language was coded as Unconfidence (lines 40-42, 46, and 49-50.) Interestingly, Interviewee O's original sorting from the first excerpt matched her final sorting after she consulted the BCT.

From these excerpts Interviewee O was coded as sorting item Q7 as existing at the BCT intersection of Apply Conceptual Knowledge, a visualization of which can be seen in Figure 4.8.

Item Q8

Item Q8 was sorted by the 4 interviewees H, J, M, and N. Item Q was the last question sorted by interviewee H and a middle item for the remaining interviewees.

Interviewee H

Interviewee H began his sorting of item Q8 along the Knowledge Dimension (line 1; coded as Interaction with BCT,) saying,

> “If we look at the knowledge level, to me, this falls under the factual knowledge. Primarily because [the students are] trying, it's essentially asking you if it's an acid-base problem. It does float a little bit into the conceptual cause it's a buffer, which is, so they have to tie some stuff together. But I would put it in factual and then probably under, because they're trying, the basis of the question says that in this acid,

the buffer solution, you have an acid and a base, its conjugate, and
you're adding a base, so it's going to react with an acid. So I would, I
would put it under factual, remember, but again, depending on, uh,
situations, it could, it could expand into understand. But because it's,
again, a multiple choice question, I would probably stick it in,
remembering factual, factual, remember?"

Because Interviewee H concluded that item Q8 should be sorted as existing at the BCT intersection of Remember Factual Knowledge, this excerpt was assigned the *a priori* codes of Remember and Factual Knowledge using directed content analysis. Interviewee H made this determination based mainly the lines of reasoning with which he assumed students would engage when approaching this item (lines 2, 4, and 6-8; coded as Assumptions of Student Reasoning or Learning Context,) and the nature of the item itself (lines 2-4, 6-8, and 10-11 ; coded as Item Forensics.) Interviewee H suggested there were situations in which the item could expand into the cognitive process of understanding (lines 9-12; coded as Conditional,) and discussed the tension of complexity in the item, in that it is a multiple choice item, which should put item Q8 in the upper-leftmost intersection of the BCT, but the buffer subject matter of the item requires students to tie multiple things together (lines 4-5 and 10-11; coded as Complexity.) This tension led to some uncertainty in Interviewee H's sorting (lines 1-5 and 8-12; coded Unconfidence.)

From these excerpts Interviewee H was coded as sorting item Q8 as existing at the BCT intersection of Remember Factual Knowledge, a visualization of which can be seen in Figure 4.9.

Interviewee J

Interviewee J began her sorting of item Q8 along the Knowledge Dimension saying,

> 1 "So I think that once again, I think that this is a conceptual knowledge
> 2 question. Students, maybe you draw on procedural knowledge about
> 3 like, maybe, I think that my students would have done a calculation
> 4 like this before, um, or maybe we talk, we've sort of talked about this,
> 5 but, and so then that way it could be kind of factual knowledge, but I
> 6 think that the main, the strongest one for me is that the conceptual
> 7 knowledge, thinking about the relationship between NaOH, a strong
> 8 base and the two components and the buffer solution."

Because Interviewee J concluded that item Q8 required only conceptual knowledge, this excerpt was assigned the *a priori* code Conceptual Knowledge using directed content analysis. Interviewee J came to this conclusion based on a number of situations in which she could see students engaging with the item (lines 2-7; coded as both Conditional and Assumptions of Student Reasoning or Learning Context.) Interviewee J does admit some uncertainty in her sorting (lines 1-6; coded as Unconfidence.)

Moving on to sort item Q8 along the Cognitive Process Dimension (lines 9 and 13-14; coded as Interaction with BCT,) Interviewee J said,

> 9 "Um, on the cognitive processes, um, I would say that this is an apply
> 10 or an analyze question. Um, and so again, I think that's, I think the

student is asked to, has to analyze the solution that is described in this question, um, in order to figure out like, what, what are these components? So that's where we're, um, ‘breaking the material into constituent parts to see how those connect to each other,’ and then, and distinguishing between the acid and the base. And the NaOH is a base, and they have to identify that. And then, um, and use that to make the answer, but then again, it could also be an apply question if we've talked about similar examples like this, and now this is again a new spin on something that we've talked about more generally with buffers.”

Because Interviewee J sorted item Q8 as requiring either apply or analyze cognitive processes depending on the context in which the item was presented to students (lines 17-20; coded as Conditional,) this excerpt was deductively coded as both Apply and Analyze. Interviewee J started her sorting of item Q8 by examining the how the question was phrased (lines 10-12; coded as Item Forensics) and then based her decision on the actions students would need to take to answer the item (lines 11-20; coded as Assumptions of Student Reasoning or Learning Context.) Interviewee J thought that if a student was familiar with buffers, that item Q8 might require the cognitive process of apply, but that otherwise, it would be a analyze process (lines 9-10 and 17-18; coded as Unconfidence.)

From these excerpts interviewee J was double coded as sorting item Q8 at both the BCT intersection of Apply Conceptual Knowledge and of Analyze Conceptual Knowledge, a visualization of which can be seen in Figure 4.9.

Interviewee M

Interviewee M once again began her sorting by considering the nature of the question (lines 1-8; coded as Item Forensics,) saying,

> “Now I have to read it again to myself to understand it. Okay. Okay. So, um, I have some questions about the validity of the response. Like one of these is not like the others, so C doesn't really fit in the blank, but regardless, um, and this is going to be based almost entirely on my experience teaching this content. Okay. Um, there's a lot of things this requires learners to think through. Um, and i- it's like, not trivial for students. This is a difficult thing for, for students to reason through th- in my experience.”

In this excerpt, Interviewee M discusses the atmosphere surrounding teaching buffers to students (lines 5-8 coded as Assumptions of Student Reasoning or Learning Context, Complexity, and Difficulty.)

Moving on to consider the Knowledge Dimension, Interviewee M said,

> “Um, so I would like for sure say that this is conceptual knowledge. Um, because I have seen students try to use only algorithmic reasoning or just like throw ICE tables at things, indiscriminately to try and reason about this. And you really cannot, you have to think about what's happening in the system. So to that end, I think it's like, very safely in conceptual knowledge. Was there another one that I missed? Meta- Oh yeah. I don't think there's evidence of metacognitive, um, knowledge targeting here.”

Because Interviewee M concluded that item Q8 required only conceptual knowledge, this excerpt was assigned the *a priori* code Conceptual Knowledge using directed content analysis. Interviewee M made this decision based on the multi-faceted nature of learning buffers (lines 10-13; coded as both Complexity and Difficulty.) At the end of this excerpt she consults the BCT to make sure she considered the full spread of the Knowledge Dimension (lines 14-15; coded as Interaction with BCT) and then removed metacognitive knowledge from consideration (lines 15-16; coded as Elimination.)

Then Interviewee M considered the Cognitive Process Dimension, saying,

> “Um, okay. So let me think about, and I would feel pretty comfortable putting this in the analyze section or category, particularly because of the language, like ‘breaks material into its constituent parts to figure out how they're related.’ And I think that's really what, because you have some multiple sort of reactions happening that you have to think about. And, um, yeah, I think that that's what you have to do to solve a problem like this. So I would say cogn- conceptual knowledge and analyze.”

Because Interviewee M sorted item Q8 as requiring only the analyze cognitive process, this excerpt was assigned the *a priori* code Analyze using directed content analysis. Interviewee M sorted item Q8 as requiring the cognitive process of analyze by citing both the way she assumes students would solve the item (lines 20-23; coded as Assumptions of Student Reasoning or Learning Context) and the headings of the Cognitive Process Dimension (lines 17-20; coded as Interaction with BCT,) as well as re-emphasizing the number of things of which students must keep track (lines 20-22; coded as Complexity.) This excerpt was also coded as Unconfidence to reflect Interviewee M’s uncertain language in sorting this item (lines 17 and 22-23.)

Asked for any final thoughts on item Q8, interviewee M once more considered the multi-faceted nature of the item's subject matter (lines 38-42; coded as both Complexity and Difficulty) and reflected on her recent teaching of it (lines 25-36; coded as Assumptions of Student Reasoning or Learning Context,) saying,

> "I actually, I actually kinda like, it. I didn't really, I guess I'm reflecting now on, I mean, it's like, been a second since I was asked some of this, but we, I like, I have this very vivid memory of last spring teaching this and we had to, um, we had to like spend one full day where we just talked through how to think about buffer solutions. And I remember being very conscious that my assessment did not consider what, um, adding like an outside species was actually doing with the buff-, like how it was interacting with the buffer solution. It only considered sort of the outcome. And I, I remember thinking like the only way to get to the outcome is to reason about what in the system this outside species is reacting with. Um, but I realized while teaching it that my students were like, not, not thinking about that. I mean, they just, they, it's like something, I mean, no, it's probably just that acid-base is hard and you want to, like, you want to maybe like, maybe it just makes it more cognitively accessible to only consider one acid-base reaction. Like if you only think about it in one way and with buffers, you're really, you're really asking them to think about it a little bit more complexly. So I don't know, this is, it's kind of cool to target this, I'm kind of here for it."

From these excerpts Interviewee M was coded as sorting item Q8 as existing at the BCT intersection of Analyze Conceptual Knowledge, a visualization of which can be seen in Figure 4.9.

Interviewee N

Interviewee N began her sorting of item Q8 along the Knowledge Dimension, saying,

> “Okay, so let me just make sure I think about the answer, the NaOH will react with the hydrazoic acid that's present, um thereby decreasing hydrazoic acid, increasing the azide, becoming slightly more basic.
>
> “Okay. Alright. So to me, this seems factual. And I think that they're just remembering factual knowledge here. Uh, or, well. Okay, it depends, on whether–so it could maybe even be an apply question because ‘students are asked within a given context to employ, calculate, carry out, change, conduct, execute, implement, infer, uh predict.’ Uh, let’s see, ‘understand: students are asked to clarify, conclude, describe, explain’...mmm. Maybe it's understand factual knowledge. I would say between understand and apply. Or remember. So anywhere in those, that first three.”

Because Interviewee N concluded that item Q8 should be sorted as existing at the three BCT intersections of Remember Factual Knowledge, Understand Factual Knowledge and Apply Factual Knowledge, this excerpt was assigned the *a priori* codes of Remember, Understand, Apply, and Factual Knowledge using directed content analysis. After answering this item for herself (lines 1-4; coded as Item Forensics,) Interviewee N made her sorting by comparing her assumptions of how students would approach the item (lines 6-7; coded as Assumptions of Student Reasoning or Learning Context) with the Cognitive Process Dimension headings (lines 9-12; coded as Interaction with BCT) but she does not seem to be able to narrow her sorting any more than that (lines 6-8 and 12-14; coded as Unconfidence.)

When asked about how she was interacting with the BCT to make these determinations, Interviewee N said, “Good question. You know, I'm trying to defer to the tool, and it's, it's been a while since I've looked at Marzano's. I had to kind of put that project on hold this quarter. So it's not as much in my mind as, as it would have been two months ago. So but yeah, I'm trying to make sure that I'm employing this tool that I see in front of me.”

In this excerpt, Interviewee N discusses the tension between her previous experience (lines; coded as Internalized Prior Taxonomy Experience) and her use of the BCT (lines; coded as Interaction with BCT.)

Asked if there was anything else she would like to share about this item, interviewee N compared it to item Q9 (lines 22-35; coded as Item Forensics,) mentioning that item Q8 was easier to sort than item Q9 (line 20; coded as Difficulty.)

“Um, this one was easier [than Item Q9.]

“I would certainly, at this point, the way that I'm teaching the class, I’d feel more comfortable asking [item Q8,] because [item Q9,] I don't feel like the course I teach right now scaffolds them to answer that appropriately at this point. Though, and that's not y– good or bad, it's just the way it is. But yes, this…question, I feel like I'm in a bit of a transition between my old style, pre-’awakening’ uh knowledge of taxonomy and learning, learning science. [Item Q8] is more, like, what I would typically ask in the past because I’d have primarily, well, not primarily, all multiple choice questions, all the time. But I'm reall– after reading all this stuff over the past year, or so, like, I'm really trying to move towards much more of a students have to produce something instead of picking things out of a list. So this sec– the questions I would, am asking at this point are kind of in between the two extremes.”

Asked if the three intersections into which she sorted item Q8 corresponded to differing levels of subject-matter expertise, Interviewee N explained, “Yeah yeah yeah. I think um as a subject matter expert, I would hope that [the students] would do the higher-order, obviously, um and when I'm reading a question like this, I'm bringing all of that experience to bear on it. I'm like ‘obviously the NaOH, because it's a strong base so it’s gonna react with the weak acid and then blah blah blah blah blah.’ But students probably are not doing all of that. So they're probably much more at the remember, or perhaps, understand phase, um. And this is something that, you know, in reading about the Marzano’s taxonomy, this, this tension between, what we, what we see in the question, versus what we want our students to do, versus what they're really doing, like, that has been, you know, whether it's, you know, Marzano's taxonomy has helped me see that in sharp relief itself, or just as kind of a triangulating between Bloom’s and Marzano’s has helped me kind of get at that tension, but that's really helped me, like I mentioned before, like, feel more comfortable asking more cognitively difficult questions, because I don't want them to just cough up a bunch of facts. I want them to be able to analyze and apply and think higher-level and explain their reasoning. So that's, yeah, it's given me more confidence to be able to do that.”

In this excerpt, Interviewee N discusses some of the assumptions she is making about the way students might interact with item Q8 (lines 37 and 41-42; coded as Assumptions of Student Reasoning and Learning Context) and how they apply to each Cognitive Process Dimension category (lines 41-42; coded as Conditional.) She bases these assumptions on the way she interacts with the item (lines 36-40 and 49-54; coded as Item Forensics.) This discussion doesn’t help her narrow her sorting any further (lines 36, 41-42: coded as Unconfidence,) but she does explain more of the inter-relationship between her Bloom’s taxonomy experience and her Marzano’s Taxonomy experience (lines 42-54; coded as Internalized Prior Taxonomy Experience.)

Asked if she thought, as a subject-matter expert, that she answered item Q8 by applying factual knowledge, Interviewee N said.

> “Oh yeah, gosh, um I'll say that I just got done teaching the course in which we talk about buffers. So I’m thinking about buffers a lot and I had a buffer question on my final at the beginning of last week. So I don't know. I mean, it's hard to say, because like I've done a lot of work to like, dig in and understand how buffers work, but I'm not like probably actively doing that analysis now, when I read a buffer question just because I know how to do it. So, at this point, and this is something that I've been thinking about a lot, in terms of, as, as reading all these taxonomies, is like, at what point does knowledge just become factual knowledge even if it's super high-level the first time you see it? Like, it's not, it's, it's a moving target, like, people learn things, they absorb them, and then it's just becomes automatic and then once it becomes automatic, is it still, like, high cognitive order, or is it now like, ‘oh, it's just factual knowledge for me.’ Like, I don't know.”

In this excerpt, Interviewee N expands on her experience with item Q8 (lines 57-61; coded as Item Forensics,) discussing her recent experience teaching it (lines 55-57; coded as Assumptions of Student Reasoning or Learning Context,) and articulating a question about the trajectory of knowledge and cognitive process as one becomes an expert (lines 61-69; coded as Internalized Prior Taxonomy Experience.)

Asked if she thinks items require less cognitive work as one gains more expertise, Interviewee N discussed the tension (lines 70-76; coded as Unconfidence) between her previous experience with learning taxonomies (lines 70-71; coded as Internalized Prior Taxonomy Experience) and her assumptions of how students interact with given items (lines 72-76; coded as Assumptions of Student Reasoning or Learning Context.)

70 “Probably! I mean, it's got to be like that, right? Otherwise, like,
71 learning would never like, happen. Like, it would always be a slog.
72 You would never like, get to the point where you've mastered it
73 because you're always trying to like, actively analyze and reason and
74 create and evaluate and you can–you can't just draw on like, your
75 lived experience, and all of the, just the content that you now
76 remember. I don't know.”

From these excerpts Interviewee N was coded as sorting item Q8 as existing at the three BCT intersections of Remember Factual Knowledge, Understand Factual Knowledge, and Apply Factual Knowledge, a visualization of which can be seen in Figure 4.9.

Item Q9

Item Q9 was likewise sorted by 4 interviewees, interviewees A, F, K, and N. Item Q was the first question sorted by interviewee F, and the last question sorted by the remaining interviewees.

Interviewee A

Interviewee A began his sorting of item Q9 by discussing the context-dependent nature of sorting items by cognitive level, saying

1 “Whew. Okay. Well, It's, this is, so this is so context-dependent, right,
2 about what the students learned. And so this literally could have been
3 a textbook, like, paragraph, and so they're just remembering it, right?”

Assuming the students were not repeating information verbatim, Interviewee A went on to look at the Knowledge Dimension, saying,

> "Now in the absence of that, there is factual knowledge here. They have to sort of remember what a reaction flask is, what a manometer is, what partial pressure is, maybe even envision that plot of partial pressure versus time, right, or something like that. Right, and so there is factual knowledge going on here. I would hope that if this is coming out-of-the-blue, that students do have some experience in setting up apparatuses to measure concentration or partial pressure. And so if they do have that, that's some procedural, isn't it? Well, in a sense, I think there is a, there is a, you know, there's a procedure here with regard to setting up an experiment. Now whether they have prior knowledge in this particular type of experiment or just general being in the lab, sort of makes me think some, some conceptual knowledge might have to come in here as well, based off of what the student, you know, has to go on here. And so that, that conceptual knowledge might be something about a rate or measuring a rate or, or something like that. Or it might not be, right? You don't necessarily have to understand anything about a rate conceptually to set up an apparatus to measure partial pressure over time. Um...Hmm. And so this student just might unlock, 'I, I know how to measure H2 concentration. And I have these tools, so let me, let me build something.' And that might be some procedural knowledge that might be that, you know, depending on the situation that might be metacognitive, because they might have to go back and think about when they failed in a lab and what went wrong and whatnot, and just depending on the situation, right.
> So this could literally be anything depending on what, what, not only the student knows, but what is the purpose of this question? What, what, what is this designed to measure? What is this designed to elicit?"

Interviewee A was not able to narrow his sorting on the knowledge dimension, so instead, he moved on to sort item Q9 along what he termed the "upper level" (line 32; coded as Interaction with BCT,) saying,

> "And along the upper level, depending on again, what, what it is, they could just be remembering something that they were told, they could be creating. They could, depending on again, what they know and

what the expectation here is, I don't, I don't think they are, and they uh, I don't think they are analyzing or evaluating. Unless, well no they're not, they're literally not doing this, they're designing it. They're not carrying it out. And so there could be sort of like a cyclical analyze and evaluate your, you know, your successes with that evaluation. So so yeah, create I think."

Because Interviewee A concluded that item Q9 required only the create cognitive process, this excerpt was assigned the *a priori* code Create using directed content analysis. Interviewee A made this determination based primarily on both his assumptions of the line of reasoning that students would employ when tasked with this item and his assumptions of how this item would be presented to students (lines 32-39; coded as Assumptions of Student Reasoning or Learning Context.) He did re-examine what the question was asking (line 37; coded as Item Forensics) in order to eliminate the cognitive processes of analyze and evaluate (lines 35-37; coded as Elimination,) however his determination was uncertain throughout (lines 32-40; coded as Unconfidence.)

When asked if the item could exist at the BCT intersection of Create Factual Knowledge, interviewee A said,

"No, no no I think it would be creation using, creating something new, using some knowledge pieces. So, so when I say create, so I'm going with create, if a student has literally never read this procedure or done it, or really done anything like it, but we have an understanding or a knowledge of, sorry, those are two different things. They do remember what all these components are. They do. They do understand what, what sorry. Not, what partial pressure is, but if they don't know how to measure it automatically, then they really have to understand what partial pressure is to come up with a way to measure it. So that is sort of my dividing line between remember and create but not create factual. Mm...Hmm.

“They are, they're remembering a lot of factual knowledge. They might be understanding some conceptual knowledge might be creating some procedural knowledge. Again, just depending on, on what they know.”

In this excerpt, Interviewee A refutes the intersection of Create Factual Knowledge by citing the context in which this item would be presented to students (lines 41-55; coded as Assumptions of Student Reasoning or Learning Context.) Interviewee A does conflate the terms understanding and knowledge, but corrects himself (lines 41-42, 44-45, and 49-55; coded as Unconfidence.) Interviewee A affirms his sorting of the cognitive process of Create, but he also identifies Remember Factual Knowledge, Understand Conceptual Knowledge, and Create Procedural Knowledge, as other BCT intersections at which he thinks item Q9 could exist (lines 53-55; coded as Conditional.)

From these excerpts Interviewee A was coded as sorting item Q9 as existing at the three BCT intersections of Remember Factual Knowledge, Understand Conceptual Knowledge, and Create Procedural Knowledge, a visualization of which can be seen in Figure 4.10.

Interviewee F

Interviewee F began her sorting of item Q9 along the Knowledge Dimension, saying,

> "I guess, because you're asking about designing an experiment, I would land more on procedural."

Because Interviewee F concluded that item Q9 required only procedural knowledge, this excerpt was assigned the *a priori* code Procedural Knowledge using directed content analysis, but her indecision was coded as Unconfidence (lines 1-2.)

Moving on to sort item Q9 along the Cognitive Process Dimension, Interviewee F said,

> "And then, so I guess then this is where I hedge and so it really depends on the context. So if they've already learned exactly how to do this, so if it's saying 'design an experiment,' but really they've actually seen the design, then I would put this at remember I guess I would wonder, well, maybe not. So I wonder how much you could do this by eye tracking too, to see how long I spent looking at each one. So I'm not really popping over to create. So I may, I can say that out loud to you. Which is interesting because even though it's 'design an experiment'--so if I'd really never done this before I could put it at create, I would think about what I'd already done to measure gasses or what I knew about measuring gasses. I wouldn't put it on evaluate. So I guess most likely I would put it somewhere, either in apply, so to modify an existing procedure, maybe that I knew or maybe even jump into create if it's really something I didn't know anything about. Yeah. And I would say remember only if they had already like really seen it point for point."

Despite initially noting that she was not instinctively considering the cognitive process of create (lines 8-10 and 13; coded as Elimination,) Interviewee F eventually sorted item Q9 as requiring either apply or create cognitive processes depending on the requirements of the question (lines 3-6 and 11-18; coded as Conditional,) and thus, this excerpt was deductively coded as both Apply and Create, even though Interviewee F felt the need to further explicitly hedge her answer (lines 3, 6-9, 11, and 14-16; coded as Unconfidence.) Interviewee F made this determination based on her assumptions of how this item would be presented to students (lines 3-6, 11-12, and 14-18; coded as Assumptions of Student Reasoning or Learning Context,) augmented by her assumptions of the lines of reasoning students might use in either scenario, and her understanding of what the item was trying to elicit (lines 5 and 10-11; coded as Item Forensics.) Interestingly, Interviewee F also spent much of this excerpt considering the tool itself and how she interacted with it (lines 7-11; coded as Interaction with BCT.)

From these excerpts Interviewee F was double coded as sorting item Q9 as existing at both the BCT intersections of Apply Procedural Knowledge and Create Procedural Knowledge, a visualization of which can be seen in Figure 4.10.

It is worth noting that even though interviewee F mentioned that if the students had explicitly seen the example point-for-point, that she could also see an argument for the BCT intersection of Remember Factual Knowledge, it was assumed that any item could exist at this intersection if it were explicitly provided for a student, and so interviewees mentioning this exception were not coded as sorting their items at the uppermost-left intersection.

Interviewee K

Interviewee K began his sorting of item Q9 along the Cognitive Process Dimension, starting with an explanation of the reasoning he thought the item was trying to elicit (lines 1-11 and 16-18; coded as Item Forensics,) saying,

> “Okay. So, um, uh, for sure, I would say that's create, um, because you're designing the apparatus. And so you're creating that, that lab set-up, uh, to monitor partial pressure of H2 produced as a function of time. Yeah. So I would say create and, um. Hmm. Um, so you want to monitor it so you don't actually want to do any calculations with it. You just want to see how much is coming off. Uh, so let's say create and, um, creating conceptual knowledge. Yes. Because it's not saying, you know, like to, to figure out, you know, like, like to do any stoichiometric calculations or anything like that. It's saying design the apparatus. Um, or des- yeah. Design the experimental apparatus to basically collect the data, so to monitor the partial pressure. Right. And so it would be create because you have to design and put together, or I guess draw the schematic for the experimental apparatus and you need to have the theories and the models and an understanding of what's happening in the reaction in order to successfully put that all together. Yeah. And it's not asking for calculations, like, it's not saying after you've done this, you know, like, what's your, what's the stoichiometry here or anything. And so, so I would say yes. Create and conceptual knowledge.”

Because Interviewee K concluded that item Q9 should be sorted as existing at the BCT intersection of create conceptual knowledge, this excerpt was assigned the *a priori* codes of Create and Conceptual Knowledge using directed content analysis. Interviewee K made this determination based on both the aforementioned Item Forensics, but also the reasoning with which he assumed students would engage when tasked with this item (lines 12-16; coded as Assumptions of Student Reasoning or Learning Context.) Interestingly, the interviewee K noted that not having to engage in calculations beyond the creation of a procedure (lines 5-9 and 16-18; coded as Complexity) still left this item sorted in the rightmost cognitive process dimension category of create. This excerpt was also coded as Unconfidence to reflect Interviewee K's uncertain language in sorting this item (lines 13-18.)

Following up on that sorting, Interviewee K was asked how students having to conduct stoichiometric calculations as part of the task would conditionally change his sorting. In response, interviewee K said,

20 "Yeah. So if it were asking, you know, like to ultimately, um, do some
21 sort of psychometric calculation or something like that, then I would
22 probably say create and both conceptual and procedural."

As the question is written, students are not required to conduct these calculations, so interviewee K was not coded as double sorting item Q9.

Moving on to discuss how she was interacting with the BCT to make these determinations, Interviewee K said,

23 "Oh, no. I was just, just entirely focusing on the tool here. Yes."

When asked about how his sortings may have depended on each of the taxonomic tools with which he was familiar, Interviewee K went on to discuss the overlap between the BCT and his preferred method of sorting items by cognitive level (lines; coded as Internalize Prior Experience.)

> “Um, I, my initial response would be no. No, I don't think so. No. I think, I think they probably give pretty consistent, um, pretty consistent answers, I think, among, among the different tools. I think so.
>
> “And so, yeah. So the, so again, it depends on like the, for [Item Q1,] especially depends on the, the problem versus exercise, right? And so I think if it's, if it's something that's given early on, when students are first, you know, being exposed to trends in electronegativity and everything, and you're explaining the whole concepts and everything, then I would say for sure, conceptual, but then if a student has done 10 of these questions on a homework assignment, then I think they're going to stop thinking about, you know, like, like all they're going to do is look at the periodic table. And, and so then I would just, like, I think once it becomes an exercise, then I think it kind of maybe can cross over into the factual knowledge category. And so to me that would be the only sort of, I guess, a little asterisk that I would put there, yes.”

Once again, Interviewee K asserted his belief that the sorting of items is dependent primarily on the context in which they are presented to students (lines; coded as both Assumptions of Student Reasoning or Learning Context and Conditional.)

From these excerpts interviewee K was coded as sorting item Q9 at the BCT intersection of Create Conceptual Knowledge, a visualization of which can be seen in Figure 4.10.

Interviewee N

Interviewee N began her sorting of item Q9 along the Knowledge Dimension, saying,

> "Okay. So, starting with the knowledge dimension, there's not much, to me, it doesn't seem like there's metacognition required there. Um, procedural knowledge, I think for sure, since we're actually like manipulating stuff and like creating something new. So, I'm using language from the top axis but I'll come back to that in a second. Conceptual knowledge? I think that requires conceptual knowledge as well, since you have to be aware of how partial pressure, what it relies on. So, I guess I would put it between procedural and conceptual, probably, well, I one issue is I'm trying to, I'm mixing the two axes in my head is another thing that's going on. So let me talk about the, the cognitive processes here for a second."

Interviewee N narrowed her potential sorting to between conceptual and procedural knowledge (lines 2-11; coded as Unconfidence.) First she removed metacognitive knowledge from consideration (line 2; coded as Elimination) and then gravitated towards procedural knowledge based on her assumptions of how students might interact with the item (lines 3-4 and 7-8; coded as Assumptions of Student Reasoning or Learning Context) but because she found herself using language from the cognitive process Dimension to justify it (lines 1-11; coded as Interaction with the BCT) Interviewee N decided to table the Knowledge Dimension for a moment.

Switching gears to the Cognitive Process Dimension, Interviewee N said,

> "So create? 'Students are asked, in order to form a coherent or functional whole, to construct, design, hypothesize, plan, produce, recogniz–reorganize,' etc. So, it seems like this is a create type of question, for me. Um, 'Evaluate: students are asked to, based on criteria and standards, judge, assess, assign value, critique, detect, examine, test.' Maybe in terms of that 'monitor?' But I feel like, yeah thinking about like a physical thing that they're doing, manipulating physical objects in three-dimensional space, does that count as evaluate from the point of view of a cognitive process? I'm not sure. And this is where my kind of three-out-of-five expertise in Bloom's might kind of get in my way here. But I would say it's probably a create thing since they are actually having to make a thing in physical space, and then somewhere between procedural and conceptual knowledge, and maybe both playing together."

In this excerpt, Interviewee N compares her assumptions of how students might interact with item Q9 (lines 17-19 and 23-24; coded as Assumptions of Student Reasoning and Learning Context) against the language of the Cognitive Process Dimension (lines 13-17; coded as Interaction with BCT) to tentatively (lines 13-26; coded as Unconfidence) sort the item at the interface of Create Conceptual Knowledge and Create Procedural Knowledge.

After clarifying the relationship between the two axes, Interviewee N went on to say,

> "Yeah. Okay. That's very helpful. Um, yeah, I think that does. Yes. I would say that they are, this question is asking them to create a procedural knowledge. Procedural Knowledge."

Because Interviewee N concluded that item Q9 should be sorted as existing at the BCT intersection of Create Procedural Knowledge, this excerpt was assigned the *a priori* codes of Create and Procedural Knowledge using directed content analysis. Interviewee N emphasized that this intersection is what the item is requiring of students (lines 28-29; coded as Item Forensics) though she still displays some uncertainty (lines 27-28; coded as Unconfidence.)

Asked if she had anything else she wanted to share about this item, interviewee N quipped,

30 "Is that the right answer?"

From these excerpts Interviewee N was coded as sorting item Q9 as existing at the BCT intersections of Create Procedural Knowledge, a visualization of which can be seen in Figure 4.10.