THESIS

CHARACTERIZING THE USE OF SPATIAL REASONING SKILLS IN INTRODUCTORY CHEMISTRY INSTRUCTOR'S MATERIALS

Submitted by

Lauren Gouldey

Department of Chemistry

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2026

Master's Committee:

Advisor: Brittney Morgan

Meena Balgopal
Nancy Levinger

ABSTRACT

CHARACTERIZING THE USE OF SPATIAL REASONING SKILLS IN INTRODUCTORY CHEMISTRY INSTRUCTOR'S MATERIALS

Undergraduate chemistry students require spatial reasoning skills to interpret static images of molecular and electronic structure, particularly when encountering new molecular concepts. Prior literature has established that spatial reasoning is an important indicator for success in chemistry, yet these skills are not always explicitly addressed in curricula. This study examined how spatial reasoning is represented and reinforced in curricular materials for a first-semester undergraduate General Chemistry course.

A full semester of instructional materials, including lecture slide decks and exams, from three instructors at an R1 institution in Colorado, as well as publisher-provided instructional materials, were analyzed. Materials were coded for instances of spatial reasoning using a deductive codebook adapted from a spatial reasoning framework originally developed for geosciences. In addition, instances where students were asked to use a spatial skill were coded for the cognitive level to characterize the demands placed on students. This work also served as a proof of concept to evaluate whether the adapted framework was appropriate for identifying and characterizing spatial reasoning skills in chemistry instructional materials.

Results indicate that certain topics within the course require more frequent use of spatial reasoning and certain spatial skills are relevant to the entire course while others are more frequent in specific units. There is also a higher level of spatial reasoning in instructor slides and exams when analyzed implicitly compared to explicitly. These findings highlight a potential misalignment between the spatial reasoning demands of instructional materials, and how these skills are assessed, and they suggest opportunities to more explicitly support the development of spatial reasoning skills in undergraduate chemistry curricula.

ACKNOWLEDGEMENTS

I would like to thank my thesis advisor Dr. Brittney Morgan and my committee members Dr. Meena Balgopal and Dr. Nancy Levinger for their guidance and feedback throughout this project. Their support throughout the data collection, analysis, and writing of this thesis was invaluable. I would also like to thank the CHEM111 instructional team for providing access to curriculum artifacts analyzed in this study.

TABLE OF CONTENTS

CHAPTER I: INTRODUCTION

Spatial reasoning plays an important role across science, technology, engineering, and mathematics (STEM) disciplines, including chemistry. Chemists rely on spatial reasoning to determine molecular geometry, symmetry, chirality, as well as to understand concepts such as group theory (Harle & Towns, 2010). Although there are many definitions of spatial reasoning and the components it encompasses, here I define spatial reasoning as the ability to recognize, interpret, and manipulate the spatial information of objects. Spatial information can include the shapes, locations, movement, and relations within and between objects (Newcombe & Shipley, 2015). Furthermore, spatial reasoning requires the ability to engage in mental rotation, spatial perception, and spatial visualization (Zhu et al., 2023). These three abilities enable individuals to perform tasks such as mentally rotating 3D objects and cutting and folding 3D objects (Yeaman et al., 2020).

Students with strong spatial skills are more likely to be successful in STEM courses and persist in STEM fields. Research has shown that students with more spatial reasoning skills in introductory STEM courses tend to have higher grade-point averages (GPAs) and are more likely to have a continued interest in STEM and pursue a STEM career than those who struggle with spatial reasoning (Sorby et al., 2018). Training spatial reasoning early in school also reduces students' struggle in later STEM courses (Lowrie et al., 2018). Although it is known that spatial reasoning is important for understanding STEM content, undergraduate curricula do not always explicitly address this competency, motivating this study.

Although demonstrating high spatial reasoning skills at a young age can indicate success in STEM courses, these skills are malleable and can be improved with intervention. Specific spatial training courses positively impact student performance in engineering programs, math courses, computer science courses, and chemistry courses (Lowrie et al., 2018; Yeaman et al., 2020). Spatial reasoning may be particularly critical in introductory STEM courses, where students have not yet developed the deep disciplinary knowledge that experts often rely on when solving problems. Buckley and colleagues (2018) suggest that in fields such as geoscience, physics, and chemistry, high content knowledge often alleviates the need to use spatial reasoning to problem solve. However, in introductory level courses, students do not yet have the content expertise necessary to compensate for low spatial reasoning (Buckley et al., 2018). Training spatial reasoning, with specific training tools such as video games or embedded curricular activities, can help students be more successful in STEM courses (Newcombe, 2017).

Studies have shown that spatial intervention training for students can improve spatial reasoning skills and lead to better STEM outcomes for students, such as higher STEM GPAs and more retention in STEM majors (Lowrie et al., 2018; Sorby et al., 2018). As outlined by Newcombe, there are two primary ways to train spatial reasoning: directly or through curriculum. Using a combination of both types of training is beneficial for student learning (Newcombe, 2017). A lack of explicitly supporting spatial skills in curriculum contributes to fewer students pursuing STEM fields (McLaughlin & Bailey, 2022). Including more spatial tools such as language, diagrams, and movement in STEM curriculum can help students develop the spatial reasoning necessary to be successful in subsequent STEM courses and careers (Newcombe, 2017). Using a combination of direct spatial training and spatial skills in specific academic contexts benefits students' understanding of spatial reasoning in many STEM contexts (Newcombe, 2017). Incorporating spatial reasoning in curriculum of STEM courses can improve student performance and retention in STEM majors.

Although spatial reasoning is clearly required to interpret many chemical representations, these skills are not always explicitly taught or identified within instructional materials. Students may therefore be expected to apply spatial reasoning implicitly while learning new chemical concepts. If spatial reasoning demands are embedded in instructional materials without being made explicit, students with weaker spatial skills may face additional barriers to learning chemistry content.

One way to better understand the role of spatial reasoning in chemistry instruction is to examine the spatial demands embedded within course materials. Lecture slides, problem sets, and exams shape how students encounter chemical representations and the types of reasoning expected of them. Systematically analyzing these materials can therefore provide insight into when spatial reasoning is required, what types of spatial reasoning skills are used, and whether these skills are explicitly addressed or implicitly assumed.

The study presented here investigates spatial reasoning in the curricular materials for CHEM 111: General Chemistry I, an introductory chemistry course at Colorado State University (CSU). CHEM 111 is the first course in a two-semester general chemistry sequence that most STEM majors complete early in their undergraduate studies. Spatial reasoning plays an important role in learning introductory chemistry concepts, as students must interpret and manipulate representations of molecular structure and behavior. Although subject-matter experts often use less spatial reasoning to complete tasks due to their extensive domain knowledge, research suggests that in chemistry novices and experts may use similar levels spatial skills to complete mental rotation tasks, amplifying the importance of spatial training in this course (Uttal & Cohen 2012). By identifying the spatial reasoning skills present in the curricular materials, instructors and students will have specific language to explicitly identify the skills necessary to be successful in the course.

To characterize these skills for this study, I use the framework developed by Newcombe and Shipley (2015). This framework identifies distinct categories and types of spatial reasoning skills and provides a structured way to describe spatial thinking, helping reduce ambiguity in how spatial reasoning is defined and applied. Newcombe and Shipley propose that there are two main dichotomies in spatial reasoning skills: intrinsic vs. extrinsic and static vs. dynamic. These dimensions combine to form four main categories: intrinsic-static, intrinsic-dynamic, extrinsic-static, and extrinsic-dynamic. Intrinsic-static includes identifying spatial features within an object, intrinsic-dynamic includes using and manipulating the spatial features of an object, extrinsic-static includes identifying the spatial location of an object in reference to other objects, and extrinsic-dynamic includes using and transforming the spatial relations between objects (Newcombe & Shipley, 2015).

The framework further subdivides these categories down into 11 subcategories that were originally developed to describe spatial reasoning in geoscience contexts. However, the authors do not examine how these skills manifest in other STEM disciplines (Newcombe & Shipley 2015). In this study, I applied this framework to instructional materials from an introductory chemistry course to identify the specific spatial reasoning skills required and identify whether this framework works effectively in the chemistry context.

To further characterize spatial reasoning within this general chemistry course, I adapted Bloom's taxonomy for visual literacy from Arneson and Offerdahl (2018). Bloom's taxonomy is a tool to hierarchically classify learning objectives and make instructional goals explicit (Arneson & Offerdahl, 2018). Traditionally, Bloom's taxonomy contains six levels, knowledge, comprehension, application, analysis, evaluation, and synthesis. For this study I adapted the model to include three levels of spatial reasoning, with one encompassing the lowest levels of Bloom's taxonomy of knowledge and comprehension, two including the middle levels, application and synthesis, and three including evaluation and synthesis, the highest levels of Bloom's taxonomy. Using this adapted codebook in this study provides a method for identifying explicit uses of different cognitive levels of spatial reasoning within this course.

The goal of this work is to provide a method to identify skills with curricular materials to gain insight into opportunities for undergraduate chemistry instructors to more explicitly support the development of students' spatial reasoning skills, ultimately helping students improve their spatial reasoning skills, setting them up for success in chemistry and other STEM courses.

CHAPTER II: LITERATURE REVIEW

Spatial Reasoning in STEM

The link between spatial reasoning skills and success in STEM has been well documented in the literature. Studies of success have measured performance, retention, or persistence. Performance in this context is indicated by grades earned in STEM subjects as well as problem solving and application of spatial skills. Retention refers to students staying in a STEM major and persistence refers to staying in a STEM career. A correlation between spatial reasoning and STEM performance has been found in children as early as preschool (Newcombe, 2017). Students with higher levels of spatial reasoning had improved mathematics performance in preschool through adulthood, with the level of spatial ability demonstrated by preschoolers correlated with arithmetic competence in early elementary school (Newcombe, 2017). Students in third through sixth grade showed improved math scores with spatial training (McLaughlin & Bailey, 2022).

Several standard tests are used to assess spatial reasoning in students. These include the Vandenberg-Kuse's mental rotation test (MRT), the Purdue Spatial Visualization Test: Rotations (PSVT:R), the Mental Cutting Test (MCT), and the Santa Barbara Solids Test (SBST). One study from Yeaman et al. (2020) coded the types of problem-solving skills students were using while completing the SBST. Identifying the types of problem-solving leads to an understanding of how to incorporate spatial skills into STEM education (Yeaman et al., 2020). Supporting post-secondary students in spatial reasoning has also yielded positive results. For example, in a study at Michigan

Technological University, students who initially scored low on spatial reasoning skills were required to take a spatial training intervention course that trained them in skills such as mental rotation, sketching 3D objects, and making 3D objects from 2D ones. The intervention course positively impacted students' STEM GPA. The intervention course was also a significant predictor of students' course grades in chemistry one among others (Sorby et al., 2018). Measuring spatial reasoning across different levels of schooling shows a correlation between spatial reasoning and STEM performance.

Although there is a significant *correlation* between spatial reasoning and success in STEM, it is more difficult to determine a *causal* relationship. Students' spatial reasoning skills influence the kinds of problem-solving students use to approach graphical and geometric problems. In a study from Buckley, Seery, and Canty (2018), students who had higher levels of spatial reasoning used different and more varied problem-solving techniques than those with lower spatial reasoning. Furthermore, students with lower levels of spatial reasoning were also more likely to reach an impasse where they could not figure out any techniques to further solve a problem (Buckley et al., 2018). Spatial skills identified for problem solving include generation of mental images and working spatial memory (Buckley et al., 2018). Although it is difficult for studies to conclude that high spatial reasoning causes higher STEM performance for students, spatial reasoning does provide better problem-solving skills which can transfer to STEM performance.

The National Science Board (NSB) also suggests that spatial reasoning is underutilized as an indicator for STEM success, missing potential candidates when recruiting for STEM majors and careers. According to the NSB, tests for verbal and mathematical talent missed 70% of students that were in the top 1% for spatial reasoning. The NSB calls for casting a wider net to recruit more of these adolescents who are talented in spatial reasoning. Opportunities such as summer programs for youth only recruit students in the top 1% for math or verbal ability, missing students in the top 1% for spatial reasoning but not for math or verbal ability (NSB, 2010). This can exclude them from talent searches despite spatial ability being an indicator for STEM success (Wai et al., 2009). By looking at spatial reasoning as an indicator for STEM performance, more students with potential for high STEM performance and persistence may end up in STEM majors and careers.

Once students choose STEM majors, continuing spatial reasoning in curriculum can increase performance and retention. Improving spatial training may reduce the number of STEM major dropouts, as students with low spatial reasoning in introductory STEM classes often become frustrated when they do not have the content knowledge to alleviate the need for spatial skills (NSB, 2010). For example, more than 40% of students in Ohio public universities dropped out of STEM majors before graduation; however spatial training interventions can lower dropout rates in spatial heavy STEM courses such as engineering (Uttal et al., 2012). Spatial reasoning ability can indicate retention in STEM and student likelihood to achieve STEM careers. One longitudinal study of 11+ years found that adolescents with high spatial ability went on to achieve STEM degrees and careers, indicating that spatial reasoning skill is an indicator of talent in fields such as physical sciences (Wai et al., 2009). The correlation between spatial reasoning and STEM success has no upper limit of expertise; however, the need for strong spatial skills is less among experts in their field. Experts in geoscience, physics, and chemistry can alleviate the need for spatial skills by using their content knowledge. Spatial ability is still beneficial for experts when coming up with new discoveries where existing knowledge is limited. Students in early STEM courses face extra challenges if their spatial ability and content knowledge is low (Buckley et al., 2018). Continuing to train spatial reasoning for STEM majors improves retention as students complete STEM majors and persistence as they continue to STEM careers.

Spatial Reasoning in Chemistry

Like other sciences, chemists regularly rely on spatial reasoning. Spatial skills are used for many aspects of chemistry such as visualizing 3D structures from 2D models, classifying stereochemistry using rotations and reflections, and drawing organic structures, group theory, and more. Many skills learned by students in a foundational chemistry class employ spatial skills, such as identifying molecular geometry (Harle & Towns, 2010). On a broad scale, the ability to visualize and mentally manipulate 2D and 3D representations translates to many specializations of chemistry, including general chemistry, organic chemistry, and inorganic chemistry (Coleman & Gotch, 1998). Many of the topics in chemistry describe microscopic entities, so spatial reasoning is required to understand the models that represent objects and processes that cannot be seen in a macroscopic way. Models are essential to demonstrate chemistry concepts and spatial skills are needed to fully understand these models (Laricheva & Ilikchyan, 2023). Understanding chemistry requires the ability to recognize and mentally manipulate models. Spatial skills need to be taught in the context of chemistry models for students to be successful.

High spatial reasoning is correlated with using and producing effective chemical models. Students who test high on spatial reasoning assessments have been shown to perform better in interpreting crystal structure and are also more likely to draw molecular representations to solve structural problems (Harle &Towns, 2010). General chemistry students with high spatial ability significantly outperform students with low spatial reasoning in molecular geometry exams (Harle &Towns, 2010). Although spatial reasoning is important for chemistry and students in chemistry, it is often not taught directly (Harle &Towns, 2010). Developing spatial abilities in students has positive performance and retention rates in science and math courses, especially for women (Harle &Towns, 2010).

Teaching and learning chemistry rely on communication through visual models of objects and processes like molecules and reactions. In general chemistry courses, students are taught concepts like valence shell electron pair repulsion (VSEPR) theory and interpreting diagrams that require spatial skills to be integrated with content knowledge to gain meaningful understanding. In organic chemistry, students must use spatial skills to understand concepts like chirality and mirror images (Harle & Towns, 2010). Incorporating spatial exercises in chemistry classes can help students develop their own mental images to aid their understanding of the content (Coleman & Gotch 1998).

Many chemistry classes rely on 2D representations or 3D model kits to incorporate the spatial nature of chemistry. These give a static representation of the spatial nature of chemistry. Computer models can be used for a more dynamic representation, but most of these give a 2D view of 3D structures (Laricheva &

Ilikchyan, 2023). In an effort to directly train spatial reasoning in chemistry, researchers at Utah Valley University provided students with virtual reality (VR) simulations to allow them to access more 3D dynamic models. Results showed the VR training led to improvement in students' spatial reasoning; however, they were not able to show statistical significance so further research is needed to draw more conclusions (Laricheva & Ilikchyan, 2023).

Chemists use spatial skills at all levels of education and career. Students use mental rotation skills on symmetrical and asymmetrical objects as they do not yet have analytical skills. Although mental rotation ability is not necessary for success in organic chemistry, it is useful for students before they have developed analytical skills (Harle & Towns, 2010). Using several tools and manipulatives also gives students the opportunity to reinforce how spatial representations are used in chemistry and how they can be used for problem solving (Laricheva & Ilikchyan, 2023). At a higher level, chemists use visual representations to support claims and communicate their findings. Spatial reasoning skills are used to generate and understand diagrams in papers and presentations, which aid communication between chemists (Harle &Towns, 2010). Experts often have strategies that circumvent the need for using spatial skills as often as students need them. For example, experts use analytical strategies when dealing with symmetrical objects, while using mental rotation for asymmetrical objects (Harle &Towns, 2010). Although the uses for spatial skills change with increased expertise, for chemists the need for spatial reasoning never disappears.

Spatial Reasoning is Malleable

Spatial reasoning is a malleable skill that can be improved with practice and intervention. For spatial training to be most useful, the skills must have transferability to problems and tasks that were not specifically trained. A meta-analysis examined how three types of spatial training, i.e., video games, courses, and spatial task training, affected spatial reasoning (Uttal et al., 2012). This study revealed that all the types of training studied improved the spatial reasoning abilities of participants. Uttal and colleagues (2012) also found that the results could be durable for weeks after the training. For those who received spatial training, there were no significant differences in scores when posttests were administered immediately after spatial training, a week after training, and a month after training. More research is needed to show which types of spatial training led to the most lasting effects (Uttal et al., 2012). Additionally, Terlecki and colleagues (2007) tested participants before spatial training, then gave a posttest immediately after the training and a retake test several months later. Participants scored highest on the immediate posttest, but the retake scores were still significantly better than the pretest scores, showing durability of training up to four months. The authors concluded that this supported the long-term impact of repeated testing and spatial practice.

Furthermore, Uttal and colleagues (2012) reported that training was transferable as there was evidence for improved transferability to novel tasks within the same type of spatial reasoning and transferability to a different type of spatial reasoning than what was trained. In addition to spatial training transferring to other spatial skills, the training transfers to STEM subjects. A meta-analysis by Hawes et al. (2022) showed that improvements in spatial reasoning also improved mathematics performance. Mathematics skills more closely related to the spatial training showed greater transferability, but there was some improvement in skills less related to the training. The type of training also affected the improvement in mathematics, with training using manipulatives leading to more improvement (Hawes et al., 2022). Training has been shown to improve spatial reasoning in participants and many types of training lead to durable and transferable skills.

The malleability of spatial reasoning has been found for students with different personal characteristics such as gender and age. Men generally have higher spatial ability than women, but both can improve spatial skills with training (Uttal et al., 2013). Although training did not close the gap in spatial skills between men and women, both have malleability of spatial skills. Women typically demonstrate lower initial spatial ability and may be slower to improve but still show improvement with training. It may be possible to close the gender gap in spatial reasoning, but whether this is necessary for STEM achievement is uncertain (Uttal et al., 2013; Terlecki et al., 2007). In addition to gender, studies have also focused on how spatial training results vary across age groups. There is generally not a significant difference in the malleability of spatial reasoning for different age groups. The difference in effectiveness of spatial training between children, adolescents, and adults was not statistically significant (Uttal et al., 2013). The transfer of spatial training to mathematics increased with age from three to 20 years old (Hawes et al., 2022). They concluded that spatial training could improve spatial ability at any age and can be more effective with increased age.

The existing research on spatial research is clear – it is essential for undergraduate students in introductory chemistry courses to build their spatial reasoning for future success.

Frameworks for Analyzing Spatial Reasoning

Although spatial reasoning is widely recognized as an important cognitive skill in STEM learning, the term encompasses a broad range of abilities and is often defined differently across research contexts. Some definitions include: "encoding and mentally manipulating the shapes and locations of objects, their relations to each other, the paths they take as they move, and paths we take as we move in, around, and through larger objects" (Uttal et al., 2024), a collection of cognitive skills using a combination of concepts of space, tools of representation, and processes of reasoning (McLaughlin & Bailey, 2022), and "representing and processing the location of objects, their shape, their relation to each other, and the orbits they take as they move" (Zhu et al., 2023). However, the definition of spatial reasoning ability has historically been a matter of convention (Uttal et al., 2013). This variability can make it difficult to consistently identify and characterize spatial reasoning within educational settings. To address this challenge, researchers have developed frameworks that categorize spatial reasoning into specific types of skills and processes. These frameworks provide structured ways to analyze when and how spatial reasoning is required in a given task or learning environment. In the context of STEM education, such frameworks can help clarify the spatial demands embedded within instructional materials and learning activities. For example, Buckley and colleges (2018) developed a framework of spatial ability based on the Cattell-Horn-Carroll theory for cognitive factors. In expanding this framework, they identify 25 factors that impact spatial ability, including a distinction between spatial ability in STEM and non-STEM disciplines.

One widely used framework was developed by Newcombe and Shipley (2015), which organizes spatial reasoning according to two key dimensions, intrinsic vs. extrinsic and static vs. dynamic, and further subdivides these categories into specific spatial skills. This framework has been applied in several STEM disciplines to describe spatial reasoning demands and provides a useful lens for examining the spatial thinking required in introductory chemistry instruction. One gap in this current framework is the lack of distinction between small scale and large-scale spatial reasoning. (Uttal et al., 2024) To develop a more coherent theoretical framework for spatial reasoning, more information is needed on spatial skills. The tests used to gather this information are often dated, inaccessible, and it is not clear what types of spatial reasoning are or are not measured in these tests. To develop more specific and applicable frameworks, first tools for collecting data need to be improved and implemented to fill the current gap in knowledge (Uttal et al., 2024).

The framework proposed by Newcombe and Shipley (2015) categorizes spatial reasoning along two primary dimensions. The first dimension distinguishes between intrinsic and extrinsic spatial information. Intrinsic spatial reasoning involves processing information about the internal properties of a single object, such as its shape, orientation, or structural features. In contrast, extrinsic spatial reasoning involves understanding the spatial relationships between multiple objects or between an object and its surrounding environment. The second dimension distinguishes between static and dynamic spatial reasoning. Static spatial reasoning refers to interpreting spatial information without imagining changes or movement, whereas dynamic spatial reasoning requires mentally transforming objects or spatial relationships through processes such as rotation, folding, or movement. Together, these two dimensions form four categories of spatial reasoning: intrinsic-static, intrinsic-dynamic, extrinsic-static, and extrinsic-dynamic.

Within these four categories, Newcombe and Shipley (2015) further identify eleven specific spatial reasoning skills that capture the range of spatial thinking processes used in STEM tasks. These skills were developed through research examining spatial reasoning in geoscience contexts and provide a detailed way to characterize how individuals interact with spatial information. By distinguishing between specific spatial reasoning processes, the framework allows researchers to systematically identify when and how spatial reasoning is required in educational materials or problem-solving tasks. Although this framework has been applied primarily in geoscience research, it provides a useful structure for examining spatial reasoning in other STEM disciplines. In this study, the framework is applied to introductory chemistry instructional materials to identify the spatial reasoning skills embedded in course content and to explore how these skills appear in a chemistry learning context.

To examine how spatial reasoning appears in introductory chemistry instruction, the framework developed by Newcombe and Shipley (2015) was applied to curricular materials from CHEM 111: General Chemistry I at Colorado State University. Lecture slide decks and exam questions from three course instructors, along with instructional materials provided by the course textbook publisher, were systematically analyzed for instances of spatial reasoning. Each instance was coded according to the spatial reasoning categories and subskills described in the framework. This approach allows for a structured analysis of the spatial reasoning demands embedded within course materials and provides insight into how these skills are represented in introductory chemistry instruction.

Another important aspect of science education is making the skills necessary for the desired outcome explicit. Bloom's taxonomy is a hierarchical organization of thinking that has been used historically to make the goals of instruction explicit (Arneson & Offerdahl, 2018). One application of this framework comes from Arneson and Offerdahl (2018) where they develop the Visualization Blooming Tool (VBT) to explicitly articulate the skills that make up visual literacy. They take the six categories of Bloom's taxonomy (knowledge, comprehension, application, analysis, evaluation, and synthesis) and define characteristics of each category and provide examples within each category related to visual literacy. A lower-level skill would involve labeling components of an image or summarizing what is represented in an image. A middle level visual literacy skill would involve comparing representations or sketching a graph from given data. A high-level visual literacy task includes examples like determining the best method to solve a problem or critiquing an existing representation.

While the VBT was developed for biochemistry, it was also applied to general chemistry, introductory biology, and cell biology (Arneson & Offerdahl, 2018). The researchers intend for this to be a tool to assess the cognitive level of visual based tasks by making them explicit. They also propose VBT as a metacognitive tool for students as they move towards expert type visualization skills (Arneson & Offerdahl, 2018). For this study, this framework for classifying visual literacy using Bloom's taxonomy was adapted to classify the cognitive level of spatial reasoning present in a question in instructional materials and exams in a general chemistry course. Similar tasks from visual literacy apply to spatial reasoning, where labeling components of a spatial image would be a low cognitive level while generating a spatial representation would be more mid-level to high level depending on how that representation is further used or manipulated. The following section describes how these frameworks work together to examine the presence spatial reasoning skills in curricular materials in a general chemistry course.

CHAPTER III: METHODS

Research Context

This study was conducted at Colorado State University, a public land grant institution in northern Colorado. The goal of this study was to identify and classify spatial reasoning demands within the curricular materials of an introductory chemistry course using thematic content analysis (Braun & Clarke, 2006).

Setting

CHEM 111: General Chemistry I is the first course in a sequence of introductory chemistry courses typically taken by students pursuing STEM majors. At CSU, most students who complete CHEM 111 continue to enroll in CHEM 113, the second part of the sequence, with the rare exception of students taking the course to fulfill core curriculum or elective credit requirements. Progression through higher level chemistry courses, such as organic chemistry, depends on the students' major and program requirements.

The instructors for CHEM 111 are primarily nontenured, teaching-focused faculty. The course includes a linked recitation and a separate lab course, CHEM 112. Artifacts for study were collected from the Fall 2023 semester, during which three instructors taught the lecture course of CHEM 111. Artifacts for this study were made available directly from the instructors, such as publisher slides and the final exam, or were available online within a shared learning management system (LMS) for the main lecture sections of the course.

Data Collection

Data collected for analysis included curricular materials from CHEM 111, including learning objectives, instructor lecture materials, and exams. Materials were accessed through the LMS Canvas or provided directly by the instructors. The analysis focused largely on instructional materials available to students through the Canvas page for the lecture portion of the course, excluding the electronic textbook (eBook) and any homework from electronic learning platforms. Slides provided by the book publisher that accompanied the eBook sections covered in the course were also analyzed to compare to instructor slide sets. Each instructor had their own individualized lecture slides, three slide sets for each week, that were organized according to the shared lecture schedule. The learning objectives for the 15-week course were primarily based on 11 chapters from the textbook.

This thesis expands my honors thesis which primarily focused on weeks 9 and 10 of the course, which covered learning objectives 7.1 through 8.1, including topics of molecular geometry, polarity, intermolecular forces, orbital hybridization, and chemical reactions. This specific material was selected after an initial coding of the learning objectives because it had a high amount of spatial reasoning skills compared to the rest of the course. To get a more well-rounded understanding of what types are spatial skills are represented within the CHEM 111 curriculum, the same analysis was expanded to investigate the full 15-week course.

Materials evaluated for the 15-week course included lecture presentations from the three instructors, slides from the textbook publisher, and five exams including the final. There were a total of 42 lecture presentations per instructor throughout the semester. Exams from past semesters and the current semester were available to students on the course Canvas page, so I analyzed these exams from five past semesters: fall 2021, spring 2022, fall 2022, spring 2023, and fall 2023. Exams consisted of 33-40 multiple choice questions depending on the unit and semester. Only one final exam was available because instructors for the course used the same final exam semester to semester to serve as a baseline for student performance over time. Unlike the other exams, the final exam is not made available to students in Canvas, and was therefore acquired directly from the instructors. Each learning objective, slide, practice question, and exam question was treated as a single unit of analysis for coding. Multiple spatial reasoning codes could be assigned to a single unit when more than one spatial reasoning skill was required to complete the task.

Data Analysis

To determine the spatial reasoning demands present in the course, learning objectives for the full course were first examined. Subsequently, I analyzed the spatial reasoning across the learning objectives, instructor in-class materials, and exams for the 15 week course. Two instruments, Bloom's taxonomy for visual literacy and Newcombe and Shipley's 2x2 framework, informed the data analytic process. Because this study explores the application of a spatial reasoning framework developed in geoscience to a chemistry context, the analysis focused on identifying patterns in spatial reasoning demands rather than testing statistical hypotheses.

Bloom's Taxonomy for Visual Literacy

I coded for the level of spatial reasoning using a codebook adapted from Arneson and Offerdahls' (2018), which draws from Bloom's taxonomy for visual literacy. I coded each learning objective, practice question, and exam question on a four-level scale ranging from zero to three, where zero indicated no spatial reasoning and three indicated a high level of spatial reasoning. Following Arneson and Offerdahls' (2018) VBT, tasks that involve low spatial reasoning include identifying features of a provided spatial image and a high level of spatial reasoning involves generating and manipulating a spatial representation. Additionally, low-level spatial reasoning includes reasoning in two dimensions, mid-level spatial reasoning includes reasoning in two or three dimensions, and high-level spatial reasoning includes a reasoning in three dimensions. The questions were coded based on the skills that were necessary to correctly answer the question. For example, if a question required students to pick a correct definition, the level of spatial reasoning would be zero even if the definition referred to spatial concepts, because the task required only recall rather than spatial reasoning. Each example was coded twice: once at a semantic level with only what was explicitly stated in the text and once at a latent level with what is implied that students needed to answer a question (Braun & Clarke, 2006).

Newcombe and Shipley's 2x2 Framework

The codebook I used to identify the type of spatial reasoning was adapted from the 2x2 spatial reasoning framework developed by Newcombe and Shipley (2015). This framework categorizes the types of spatial reasoning according to two dimensions: intrinsic or extrinsic and static or dynamic. Combining these dimensions results in the four distinct categories: intrinsic-static, intrinsic-dynamic, extrinsic-static, or extrinsic-dynamic. Exemplars for this context are presented in Figure 1.

Figure 1

2x2 Spatial Reasoning Framework with General Chemistry I Examples

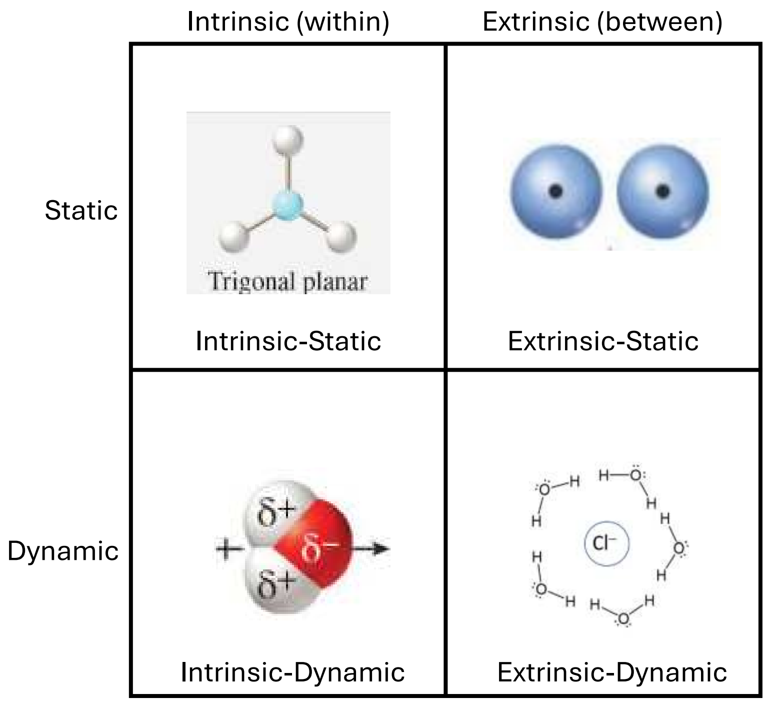


Figure adapted from 2X2 spatial reasoning framework depicted by Uttal (2012) with examples of images associated with spatial reasoning from General Chemistry I (CHEM 111) content. The examples show a mix of 2D and 3D representations, as each type of spatial reasoning can be used in 2D or 3D space. The intrinsic-static example uses identifying electron domains within a molecule to categorize molecular geometry. The extrinsic-static image identifies the distance between molecules being relatively far apart, so they do not interact. The intrinsic-dynamic example shows using the interactions within the molecule to show symmetry and polarity. The extrinsic-dynamic image shows how the interactions between molecules affect orientation in space in solutions.

As mentioned, Newcombe and Shipley (2015) identified 11 skills within the framework based on a geoscience context: disembedding, categorization, visualizing 3D from 2D, penetrative thinking, mental transformation, sequential thinking, locating self and other objects, alignment, perspective taking, relations among objects, and updating movement through space. All 11 spatial reasoning skills were initially included in the coding process to determine which skills transfer to chemistry. During coding, one of the 11 spatial skills (sequential thinking) was not found in the course materials, so it was removed from the codebook. Following this codebook, I used deductive qualitative coding for each learning objective, lecture slide (including practice questions), and exam questions in the available artifacts to identify where the 10 remaining types of reasoning were present. Practice and exam questions were coded on a latent level with only the type of spatial reasoning that was necessary to solve a question. Each of the 10 spatial reasoning types were added in their larger categories to find the frequencies and percentages of intrinsic, extrinsic, static, and dynamic spatial reasoning in the CHEM 111 materials.

Data Analysis Software

MaxQDA and Microsoft Excel were used to support the qualitative coding and quantitative summarization of the data. MaxQDA allowed for effective management of the qualitative coding process. Each document artifact, which included learning objectives, slides, practice examples, and exam questions, were uploaded into MaxQDA (2024 edition). Following an initial round of deductive coding using the two analytic frameworks mentioned above, the codebook was refined to better reflect the spatial reasoning patterns observed in the chemistry course materials. Each artifact was coded individually, and the counts for each coded segment were exported from MaxQDA into Excel for further analysis.

The total number of occurrences of each code across the artifacts was aggregated in Excel across all artifacts. Each instance coded of either level or type of spatial reasoning from learning targets, exam questions, practice questions, etc., was added into their larger unit categories based on the course schedule to obtain the frequencies of each code. The number of each level was divided by the total number of questions within the sets of slides or exams to calculate the percentage of each level. Each type of spatial reasoning skill for each set of slides and exams was divided by the total number of instances of any type of spatial reasoning coded within that set for the unit to obtain the percentage. The frequencies and percentage of each type of spatial reasoning skill within each set of slides and exams for the unit were added into the four larger categories in the 2x2 framework, then further added into larger intrinsic and extrinsic categories and static and dynamic categories for comparison purposes.

Trustworthiness

Several strategies were used to support the trustworthiness of this study. First, the codebook used in this analysis was adapted from established frameworks developed by experts in the field of spatial reasoning research. The 2x2 spatial reasoning typology developed by Newcomb and Shipley (2015) has been used to categorize spatial reasoning in other scientific disciplines. This framework has also been successfully applied in the Uttal et al study of the malleability of spatial reasoning (Uttal et al., 2013). The 11 subcategories were developed by Newcombe and Shipley based on interviews with geoscientists, so they are based on the skills professionals use often in their career (Newcombe and Shipley 2015). The Visual literacy Bloom's Taxonomy was developed using feedback from students and experts and had a high interrater reliability (Arneson & Offerdahl, 2018). This study used a simplified version for the codebook, but it is based on a reliable tool from previous studies.

To further strengthen the credibility of the coding process, several examples from lecture materials were coded collaboratively with other experts in science education research from my committee. Selected examples from lectures, student practice, and exam questions were coded for both types of spatial reasoning and cognitive level students needed to answer a question. The collaborative coding discussion helped refine the interpretation of the codebook. All coders agreed on the value of the codebook and the way examples were coded, establishing trustworthiness of the analysis.

Together, these analytic procedures enabled the systematic identification and quantification of spatial reasoning types and levels across the CHEM 111 instructional materials. The research questions guiding this study are descriptive in nature and focus on identifying patterns in the presence and types of spatial reasoning within instructional materials. The purpose of these questions is not to test statistical differences but rather to characterize spatial reasoning demands in the curriculum and demonstrate how established spatial reasoning frameworks can be applied in a chemistry education context. Specifically, this study identifies both the cognitive level of spatial reasoning required by course materials and the types of spatial reasoning skills present by applying established spatial reasoning frameworks to a chemistry education context.

CHAPTER IV: FINDINGS AND DISCUSSION

This section presents the results of applying a spatial reasoning framework originally developed for geoscience contexts by Newcombe and Shipley (2015) to instructional materials in a general chemistry course to explore the extent to which this framework can be used to identify and characterize spatial reasoning within chemistry instruction, positioning this work as an initial proof of concept for its use in a chemistry education context. Because the goal of this work is exploratory, the statistics presented in this section are descriptive and are intended to illustrate patterns in the data rather than to support claims of statistical significance.

Learning objectives from the 15-week course were first analyzed to provide an overview of the course. Each learning objective was coded for the level of spatial reasoning required based on factors such as if spatial reasoning was 2D or 3D and if students were using a spatial representation already given or generating their own, adapted from the Bloom's taxonomy for visual literacy (Arneson & Offerdahl, 2018). The types of spatial reasoning skills as described in the framework developed by Newcombe and Shipley (2015) were then coded based on what type of spatial reasoning was involved in the learning objectives that were identified to require some level of spatial reasoning.

Within the 231 learning objectives, the majority did not require any spatial reasoning. When coding both semantically and latently, the percentage of learning objectives receiving a zero-level code was 69% and 68%, respectively suggesting that when spatial reasoning was involved, it was generally expressed explicitly in the learning objective. The frequency and percentage of the level of spatial reasoning, coded both semantically and latently, are presented in Table A1 in Appendix. In this coding scheme, a zero code indicates no spatial reasoning was required for a task, whereas the highest level of spatial reasoning required was coded as three.

The frequency and percentage of each type of the 10 types of spatial reasoning found within the learning objectives for the course are presented in Table 1. Latent coding captured spatial reasoning implied by the task even when not explicitly stated, whereas semantic coding reflected spatial reasoning that was directly described in the text of the objective. All coding of spatial reasoning types in this study was conducted at the latent level to represent all types of spatial reasoning required for a task even if not explicitly stated.

Table 1

*Frequency (n) and percentage (%) in all learning objectives of each of the 10 spatial reasoning skills from Newcombe and Shipley found in the course.**

Category of Spatial Reasoning	Typology Skill	n	%
Intrinsic-static	Disembedding	14	14
	Categorization	7	7
Intrinsic-dynamic	Visualizing 3d from 2d	6	6
	Penetrative Thinking	37	37
	Mental Transformations	8	8
Extrinsic-static	Locating Objects	0	0
	Alignment	0	0
Extrinsic-dynamic	Perspective Taking	4	4
	Relations Among Objects	18	18
	Updating Movement	7	7

*The sequential thinking skill is omitted in reporting because it did not occur in the data.

The distribution of spatial reasoning skills across the learning objectives reflects a strong emphasis on certain typologies within the Newcombe and Shipley framework. Among intrinsic-static skills, which involve understanding properties of objects independent of movement, disembedding, the ability to isolate and attend to one aspect of a complex display or scene, appears in 14% of objectives, while categorization, or identifying categories based on shared spatial relations, accounts for 7% of skills identified. Within intrinsic-dynamic skills, which require transforming the spatial relations of objects mentally, penetrative thinking, visualizing spatial relations inside an object, dominate with 37%, followed by mental transformations (8%) and visualizing 3D from 2D representations (6%). In contrast, extrinsic-static skills, which focus on identifying the spatial location of objects relative to other objects or to a reference frame without movement, such as locating objects and alignment, are entirely absent from the learning goals. Finally, extrinsic-dynamic skills, involving transforming the inter-relations of objects as one or more of them moves, show moderate representation: relations among objects (18%), updating movement (7%), and perspective taking (4%). Overall, the data from the learning goals suggest a heavy emphasis on intrinsic-dynamic skills and we could expect the content itself should reflect these trends.

After analyzing the learning objectives to establish an overview of the course, instructional materials from each instructor and from the textbook publisher were analyzed. Instructional materials were categorized into five groups based on the topic covered on an exam as outlined in the course schedule. The total number of instances of a type of spatial reasoning from Newcombe and Shipley’s (2015) framework was added for each unit, as well as the total number of slides for each instructor. To allow for comparison across instructors and materials, the total number of coded instances were divided by the number of slides to produce a normalized value of spatial reasoning codes per slide for each instructor and for the publisher materials, as shown in Table 2. Slides were used as the unit of analysis because they represent the primary instructional structure used across instructors in this course.

Table 2

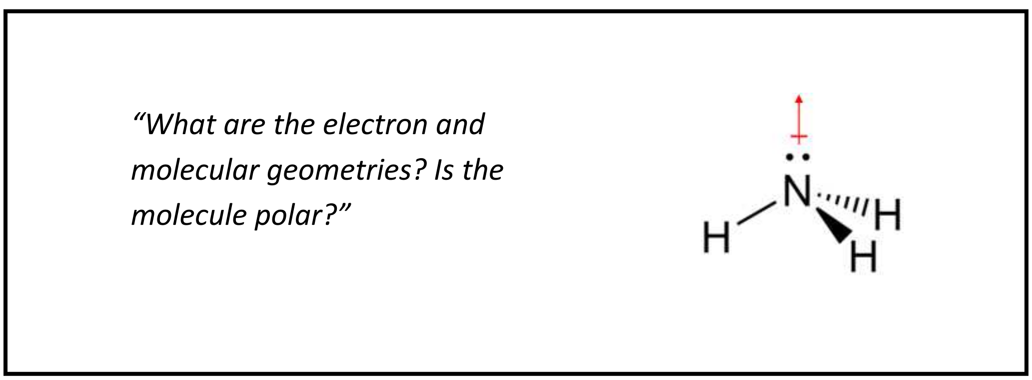
Total Instances of Spatial Reasoning, Number of Slides, and Codes per Slide presented by instructors or provided by textbook

		Instructor 1	Instructor 2	Instructor 3	Book
	Code	239	269	123	94
Unit 1	Slide	228	316	223	127
	Code/Slide	1.05	0.85	0.55	0.74
	Code	200	270	97	98
Unit 2	Slide	244	385	265	170
	Code/Slide	0.82	0.7	0.37	0.58
	Code	661	1131	573	452
Unit 3	Slide	262	422	269	202
	Code/Slide	2.52	2.68	2.13	2.24
	Code	203	235	157	134
Unit 4	Slide	251	323	235	146
	Code/Slide	0.81	0.73	0.67	0.92
	Code	138	192	87	65
Final	Slide	178	258	167	97
	Code/Slide	0.78	0.74	0.52	0.87

The total number of slides was similar between each instructor and the publisher materials across units, with the final unit having the least slides as there were fewer classes compared to the other units. Unit three contained the greatest number of spatial reasoning codes with more than double the number of codes for each group of slides compared to the other units with an average of more than two instances of spatial reasoning per slide. Because multiple spatial reasoning skills could be coded within a single slide, slides from this unit often contained more than one coded instance. This finding is consistent with the initial pilot of this project which coded only a subset of these materials in Unit 3 that were identified as having a high spatial reasoning demand using the initial coding of learning objectives. The higher frequency of spatial reasoning instances in unit three suggests that some chemistry topics require more spatial reasoning than others. Unit three covered topics including drawing Lewis structures, molecular geometry and polarity, and intermolecular forces. These topics require multiple types of spatial reasoning to fully understand and correctly answer related questions. Each type of spatial reasoning was coded individually, leading to an increased frequency in instances of spatial reasoning coded in unit three. Figure 2 gives an example from a section of an instructor slide from unit three that involves multiple types of spatial reasoning.

Figure 2

Example of questions using disembedding, categorization, penetrative thinking, and visualizing 3D from 2D spatial reasoning skills to describe molecular polarity.

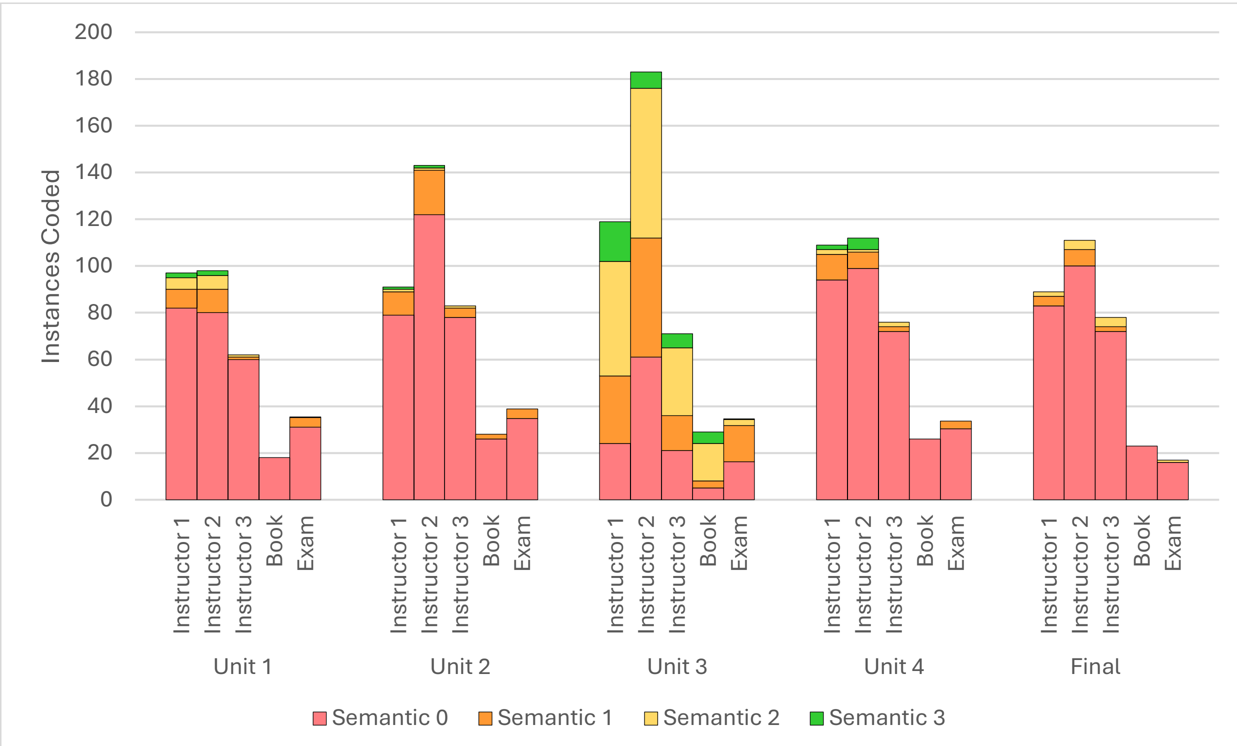


Addressing the questions in Figure 2 requires several types of spatial reasoning. Students must use intrinsic-static skills including disembedding and categorization. Disembedding is required to identify important factors within the representation, such as lone pairs and bonded electron domains. Students must identify the appropriate spatial category by naming the electron and molecular geometries. Additionally, students use intrinsic-dynamic skills including visualizing a three-dimensional (3D) molecular arrangement from the two-dimensional (2D) to understand this wedge and dash structure and visualize the geometry. Finally, penetrative thinking is required to evaluate symmetry to identify the polarity of the whole molecule. These types of questions within unit three help explain why the frequency of spatial reasoning is so much higher, as many of the topics within this unit use multiple types of spatial reasoning to interpret a single molecular representation.

The frequencies and percentages of each code were averaged together for comparison to the lecture materials except for the final, where only one document was available on Canvas for the final exam. This only covers content new to the final exam, not cumulative questions from past units, which is why there are fewer questions than for exams 1-4. Figure 3 presents the frequency of each level of spatial reasoning when coding semantically for each set of slides and exam for each unit. Table A2 in Appendix contains the frequency as well as the percentage for each level on the semantic and latent level for the slides and exams.

Figure 3

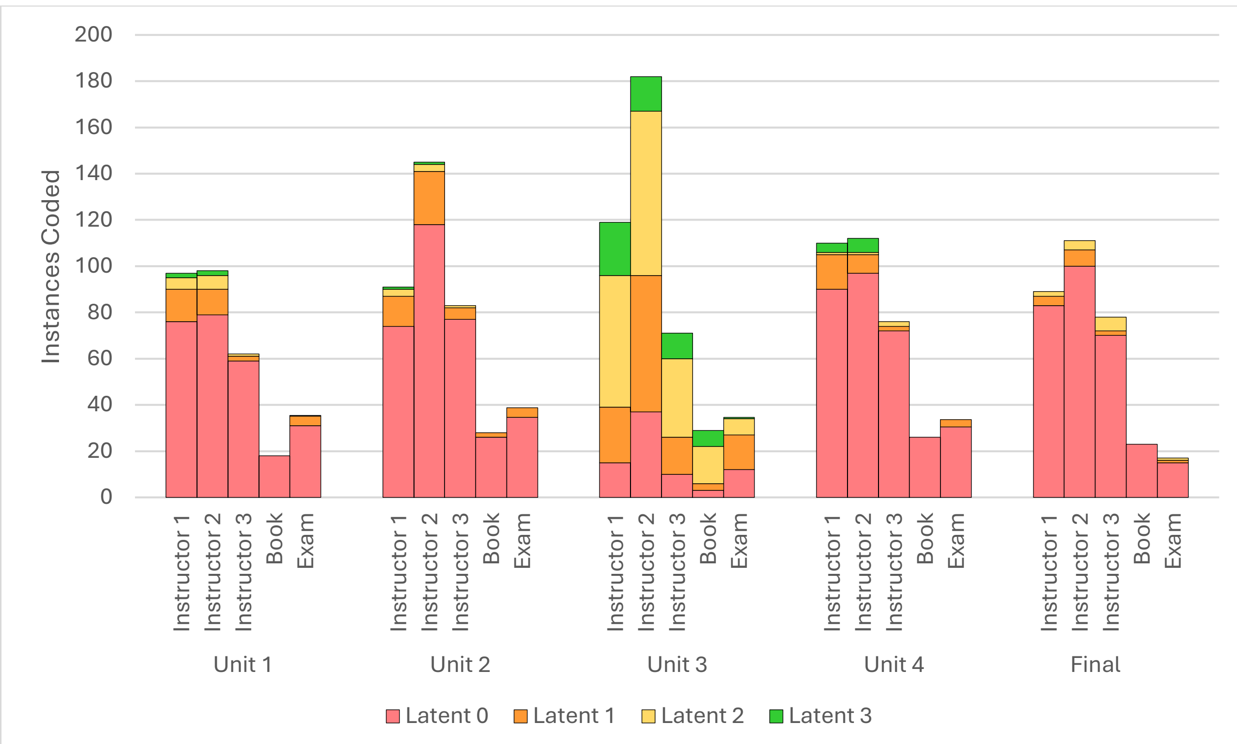
Frequency of levels 0, 1, 2, and 3 spatial reasoning coded semantically for instructional materials and exams within five units



The level of spatial reasoning is not evenly distributed throughout the course. Unit three has the highest frequency of levels one, two, and three spatial reasoning. When students get to unit three, they could struggle more if they do not have adequate spatial skills. Since the use of spatial skills is not evenly built throughout the course, it is important for instructors to be aware of when higher level spatial skills are required in the course so they can integrate practice into the class with explicit spatial strategies such as using spatial language, sketching, models, and gestures (Newcombe, 2017). Intentionally incorporating these strategies into curriculum can help students make meaning out of content specific representations and improve their spatial reasoning without requiring extensive training outside of class (Harle & Towns, 2010)(Newcombe, 2017). Figure 4 shows the frequency of each level of spatial reasoning when coding latently for each set of slides and exams.

Figure 4

Frequency of levels 0, 1, 2, and 3 spatial reasoning coded latently for instructional materials and exams within five units



Taken together, the results in Figures 3 and 4 suggest that spatial reasoning plays a particularly important role in topics that rely heavily on molecular representations. In unit three, students must interpret and translate between multiple representations of molecules, including Lewis structures, wedge and dash drawings, and 3D molecular geometries. Chemistry requires translations between macroscopic, microscopic, and symbolic representations, and while experts easily translate between representations, students often struggle with microscopic and symbolic representations (Oliver-Hoyo & Babilonia-Rosa, 2017). These tasks require students to use several spatial reasoning processes simultaneously in order to interpret structural information and predict molecular properties such as polarity and intermolecular interactions. The high frequency of spatial reasoning instances in this unit therefore highlights the central role that spatial thinking plays in developing representational competence in chemistry.

Unit three had the highest frequency of levels two and three reasoning by both frequency and percentages for slides and exams. The instructor slides had more questions and activities for students to practice spatial reasoning in the content compared to the number of example problems from the book slides. When coding semantically and latently, all three instructors had a higher frequency of instances of levels one, two, and three spatial reasoning when compared to the book, as seen in Figures 3 and 4. This could suggest that instructors draw from their own pedagogical content knowledge and practitioner experience to have a better sense of what topics students need the most support with, in comparison to textbook publishers and editors. For units one, four, and the final, the book slides had the lowest percentage of problems that used spatial reasoning (See Table A2 in Appendix for percentages). For unit two the book slides had a similar percentage of questions that used spatial reasoning as instructor three but was lower than the other two instructors. While instructor three consistently had a lower frequency of all levels of spatial reasoning than the other two instructors, it is possible more practice was given in class than is present on the slides. However, it is still important for explicit practice with high level spatial reasoning to appear in the slides as those materials are often most accessible to students where they can see examples of manipulations of spatial representations, rather than having to seek out or generate all representations on their own.

The unit three exams also had the highest frequency of levels one, two and three reasoning of all the unit exams, being the only unit where the exams contained any level three reasoning. This again supports that certain topics in chemistry require more spatial reasoning or a higher cognitive level of spatial reasoning, as unit three had a higher frequency of levels two and three reasoning than the other units.

Materials from unit three also displayed the largest differences when coded semantically versus latently, as seen when comparing Figures 3 and 4. When coded semantically, 53% of the questions on exam three were found to require spatial reasoning, whereas latent coding increased this percentage to 65%. A similar trend was observed in the instructor slides for unit three, where more instances of levels two and three spatial reasoning were identified when coding latently. Although the publisher slides for unit three showed this trend to a lesser extent, the percentage of level three spatial reasoning also increased when coding latently versus semantically. There were fewer examples coded for the book slides than for the instructors, which likely contributed to the smaller changes in percentage. For the other four units, the levels of reasoning did not change for the publisher slides when comparing latent and semantic coding. For instructor slides in these units, the number of questions identified as requiring spatial reasoning increased slightly when coding latently but was not as pronounced as in unit three.

Across the course, certain examples showed a required level three spatial reasoning when coded semantically and latently, such as the example below:

> *"Draw the Lewis structure of SF_4 and then determine its electron domain and molecular geometries. Draw a 3D sketch and label the bond angles."*

This example uses explicit spatial language such as "draw" and "3D sketch" and gives a sequential order of spatial tasks. Based on the working of the question alone, students are clearly prompted to use spatial reasoning skills to solve the question. In contrast, some questions may have students use spatial reasoning, but do not make it explicit in the question. For example:

> *"What intermolecular forces occur between water and ammonia (NH_3) molecules?"*

Although this question does not include explicit spatial language as an indicator for the individual to use spatial skills, students are more likely to answer it successfully by employing spatial reasoning skills such as mental visualization or sketching molecular structures. The types of intermolecular forces between the two molecules are related to the polarity, based on the symmetry within the molecule and the orientation of the molecules related to each other. These skills would be categorized as both intrinsic- and extrinsic-dynamic skills under the Newcomb and Shipley (2015) framework. Students would be required to identify which spatial skills to use based on the content language. Students need to be able to understand spatial language specific to chemistry, such as wedge and dash structures, but they should receive direct instruction using spatial language and chemistry language on how to use spatial representations (Harle & Towns, 2010).

The higher percentages of level two and three spatial reasoning identified through latent coding suggest that the spatial reasoning demands are often implicit rather than explicitly stated. Making spatial reasoning more explicit within instructional materials and exams with clearer cues and spatial language could give students a better chance of success for tasks that require spatial reasoning. Students in CHEM 111 are still at a beginner level where they are taking concepts they learned in class and applying them to novel molecules or situations on exams. Spatial reasoning opens more problem-solving techniques for students (Buckley et al., 2018).

On exams, I found a higher percentage of level one and two spatial reasoning when coding latently compared to semantically. While information on instructor slides could be verbally clarified in class with instructor input to make spatial reasoning more explicit, exams are completed independently by students. As a result, if required spatial reasoning is not made explicit in exam questions, it is up to the students to correctly infer when to use it.

Exams had a lower proportion of level two and three spatial reasoning when compared to learning objectives or instructor slides. This finding may suggest that students are not assessed on spatial reasoning used in the course. However, this study only coded the spatial reasoning required by the wording of the exam questions themselves. Even when a question on an exam does not explicitly require a high level of spatial reasoning, students may use spatial skills to solve problems. Students often use more spatial reasoning than experts while they are learning content (Harle & Towns, 2010). More research is needed to determine if chemistry students use spatial reasoning to solve problems on exams when they are not explicitly written or if they are using other techniques or a combination of the two.

The slides from instructors and the book and the exams were coded latently for each type of spatial reasoning identified by Newcombe and Shipley (2015). These codes were entered into the larger categories within the 2x2 typology to find the frequency of intrinsic-static, intrinsic-dynamic, extrinsic-static, and extrinsic-dynamic for each set of slides and exams in each unit. These were further added to the total frequency of intrinsic and extrinsic codes, which can be seen in Table A3 in Appendix. Intrinsic spatial reasoning involves the identification and manipulation of spatial information within an object and extrinsic spatial reasoning includes the identification and manipulation of spatial relationships between objects. The frequency of intrinsic and extrinsic codes for each set is shown as well as the percentage of all the types of spatial reasoning codes that were intrinsic versus extrinsic.

For all five units, there were far more instances of intrinsic than extrinsic spatial reasoning. For exams one, two, four, and the final, out of around 35 questions on each exam, there were few instances of any type of spatial reasoning, with the unit two and final exams having no instances of extrinsic spatial reasoning. As previously identified, unit three had the highest frequency of both intrinsic and extrinsic spatial reasoning across both the slides and exams. There was a higher percentage of intrinsic spatial reasoning compared to extrinsic for units two and three, where of all the instances of a type of spatial reasoning in each set of slides and exams, at least 80% were intrinsic. The other units had some sets where around 70% of all the types of spatial reasoning was intrinsic, still showing a majority spatial reasoning within the units are intrinsic.

There is a variation across instructors and book slides and the exams of the percentage of spatial reasoning. For example, in the final unit, the exam had 100% of the instances of spatial reasoning were intrinsic where for instructor two it was 71% and for instructor three it was 80%. Unit three had overall a higher percentage of intrinsic spatial reasoning but there was also less variation in the percentage of intrinsic and extrinsic spatial reasoning across the sets. While the percentages of the categories of spatial reasoning are not widely different between the instructor slides and the book slides, the frequency of spatial reasoning is varied. The book publisher slides had the lowest frequency of any type of spatial reasoning across all units. Instructor three had the most similar frequency of spatial reasoning to the book publisher slides as their slides most closely followed what was provided by the book. However, they still had a higher frequency of intrinsic and extrinsic spatial reasoning skills than the book publisher slides. Instructors one and two had a much higher frequency of spatial reasoning than the book publisher slides. This indicates that instructors are adding more examples for students to see examples of and use spatial reasoning within the course than what the book provides.

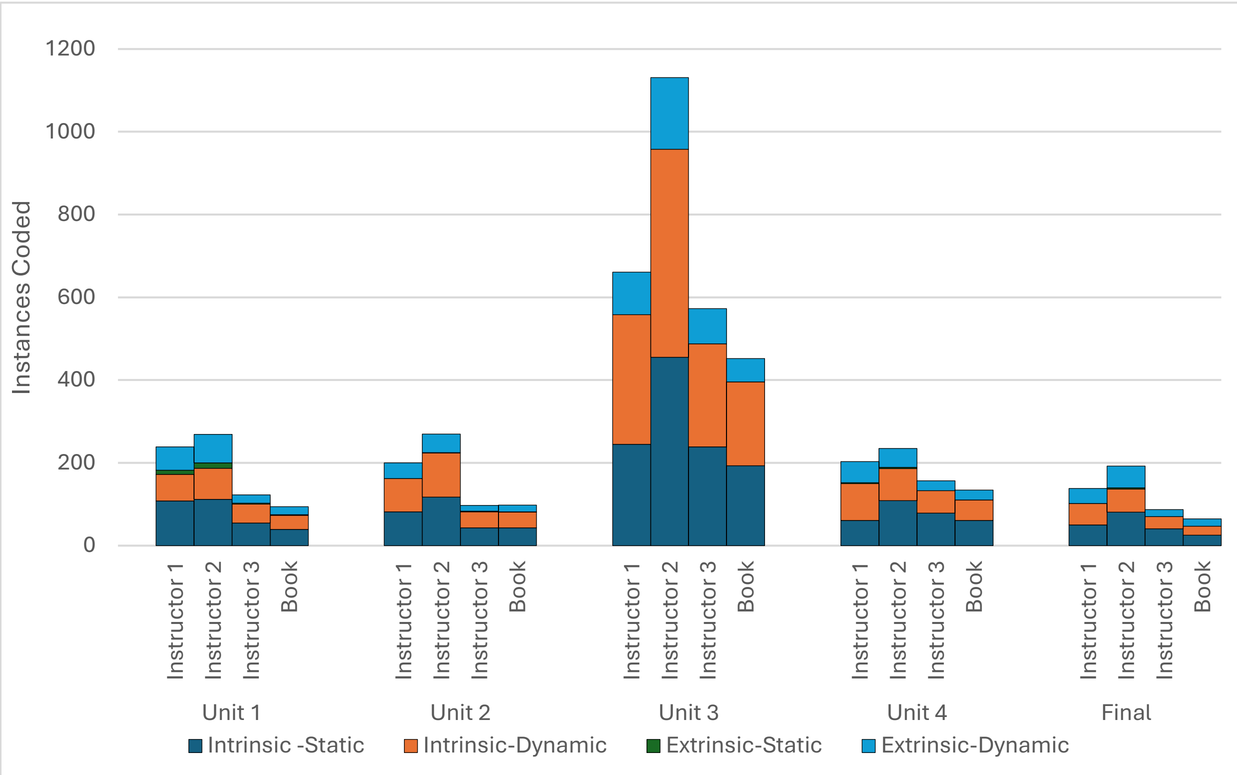
Static spatial reasoning involves the identification of spatial information and dynamic spatial reasoning involves interpreting and manipulating spatial information. The four categories of intrinsic-static, intrinsic-dynamic, extrinsic-static, and extrinsic-dynamic were also added to find the total number of static and dynamic for each set of slides and exams for each unit. The percentage of all instances of a type of spatial reasoning that was static and dynamic was also calculated and is shown in Table A4 in Appendix.

For each unit and across the slides, all the instances of a type of spatial reasoning, around 50% were static and 50% were dynamic. The instances of dynamic spatial reasoning in unit three was closer to 60%. Of the 10 types of spatial reasoning present in the course, six were dynamic types and four were static, as outlined in Table 1. Unit three had the highest frequency of total codes with a higher percentage of them being dynamic compared to other units. Of the sets of slides, instructor one had the highest percentage of dynamic skills most often. The exams for each unit had a higher percentage of static spatial reasoning compared to the slides from instructors and the textbook for each unit. Exam one had the lowest percentage of dynamic spatial reasoning at 19%. For exam three, 51% of the spatial reasoning was dynamic compared to around 60% for the slides for that unit. For the final exam, 60% of the spatial reasoning was dynamic, which was more than some of the sets of slides for that unit. The lower percentage of dynamic spatial reasoning on exams indicates that in class, instructors are including examples of identifying spatial relationships and manipulating them, while spatial reasoning on exams more often only requires identifying spatial relationships.

The frequency of each of the categories of spatial reasoning from the 2x2 typology (Newcombe and Shipley, 2015) for each set of slides within each unit is presented in Figure 5.

Figure 5

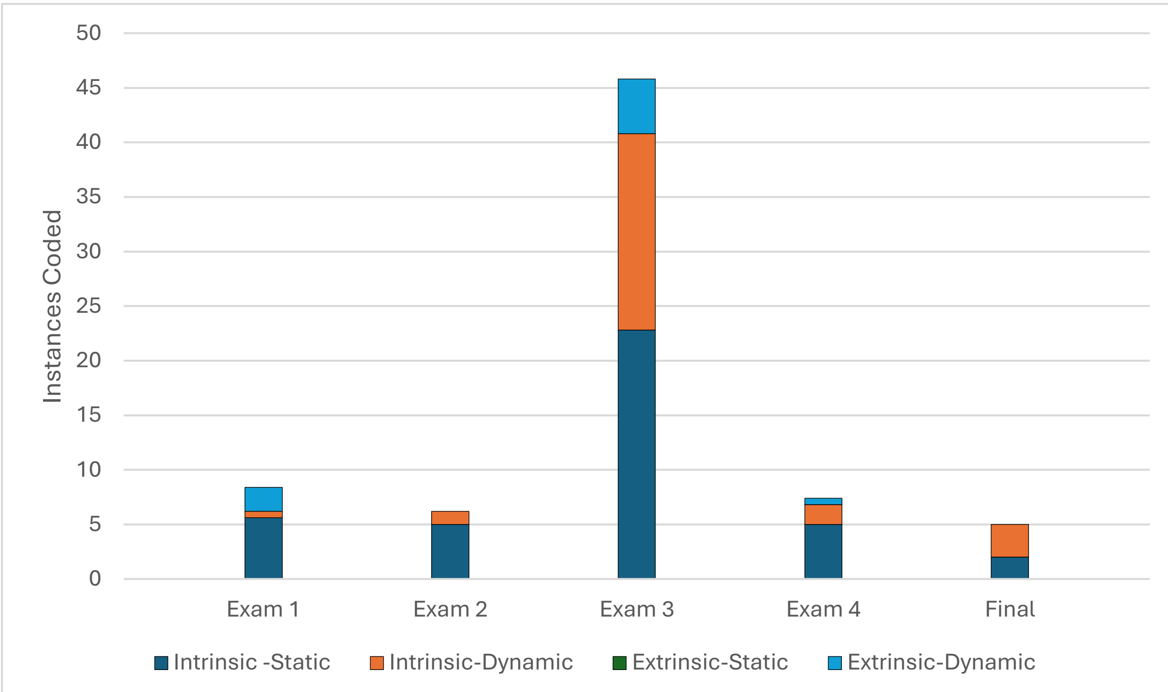
Frequency of Intrinsic-Static, Intrinsic-Dynamic, Extrinsic-Static, and Extrinsic-Dynamic spatial reasoning across instructional materials in five units of CHEM 111



Of the four categories, extrinsic-static had the least instances in the slides and exams for all the units. Unit one had the most instances of extrinsic-static spatial reasoning of the five units, but it still made up the smallest percentage of the four categories within unit one (See Table A5 in Appendix for percentages of all types of spatial reasoning made up by each of the four categories). Intrinsic-static and intrinsic-dynamic made up the highest percentage of codes across the units. Unit one had a higher percentage of intrinsic-static in the slides for the instructors and book. Units two and three have a similar percentage of intrinsic-static and intrinsic-dynamic across slide sets, with instructor one having a higher amount of intrinsic-dynamic. For unit four and the final, instructor one had a higher percentage of intrinsic-dynamic while the other sets of slides had a higher percentage of intrinsic-static. For each exam except the final, there was the highest percentage of intrinsic-static spatial reasoning of the instances of spatial reasoning, as shown in Figure 6. Extrinsic-dynamic spatial reasoning was more common than extrinsic-static but not as frequent as either of the intrinsic categories for each unit. The exams for each unit had a higher percentage of intrinsic-static spatial reasoning than the slides. The percentage of intrinsic-static was most similar from the slides to the exam in unit three, but the exam did have a slightly higher percentage. This again supports that the exams more often stop at spatial reasoning that requires identifying spatial relationships, where in class instructors are adding opportunities to practice manipulating the relationships.

Figure 6

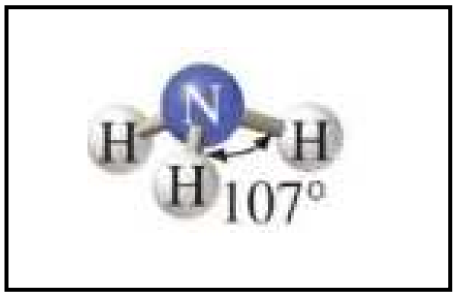
Frequency of Intrinsic-Static, Intrinsic-Dynamic, Extrinsic-Static, and Extrinsic-Dynamic spatial reasoning across exams in five units of CHEM 111



None of the exams had any instances of extrinsic-static spatial reasoning and unit two and the final exams had no instances of extrinsic-dynamic spatial reasoning. While there is an overall trend of intrinsic-static and intrinsic-dynamic skills being the most frequent, each unit having a different percentage of each of the four categories supports that different topics in chemistry require different types of spatial reasoning. Within this course, unit one required the most extrinsic-static skills and unit three generally required the most intrinsic-dynamic skills. Many instances of extrinsic spatial reasoning were similar content to intrinsic spatial reasoning but had different definitions or how something was being manipulated to change what was defined as within object versus between objects. For instance, in Figure 7 shown below, a question about the bond angle would fall under extrinsic-dynamic as Newcombe and Shipley define relations among objects, an extrinsic-dynamic skill, to include angles formed by three points. The same image could also be used to ask about molecular polarity, which would fall under intrinsic-dynamic as it requires penetrative thinking to think about the symmetry using spatial relations within the molecule.

Figure 7

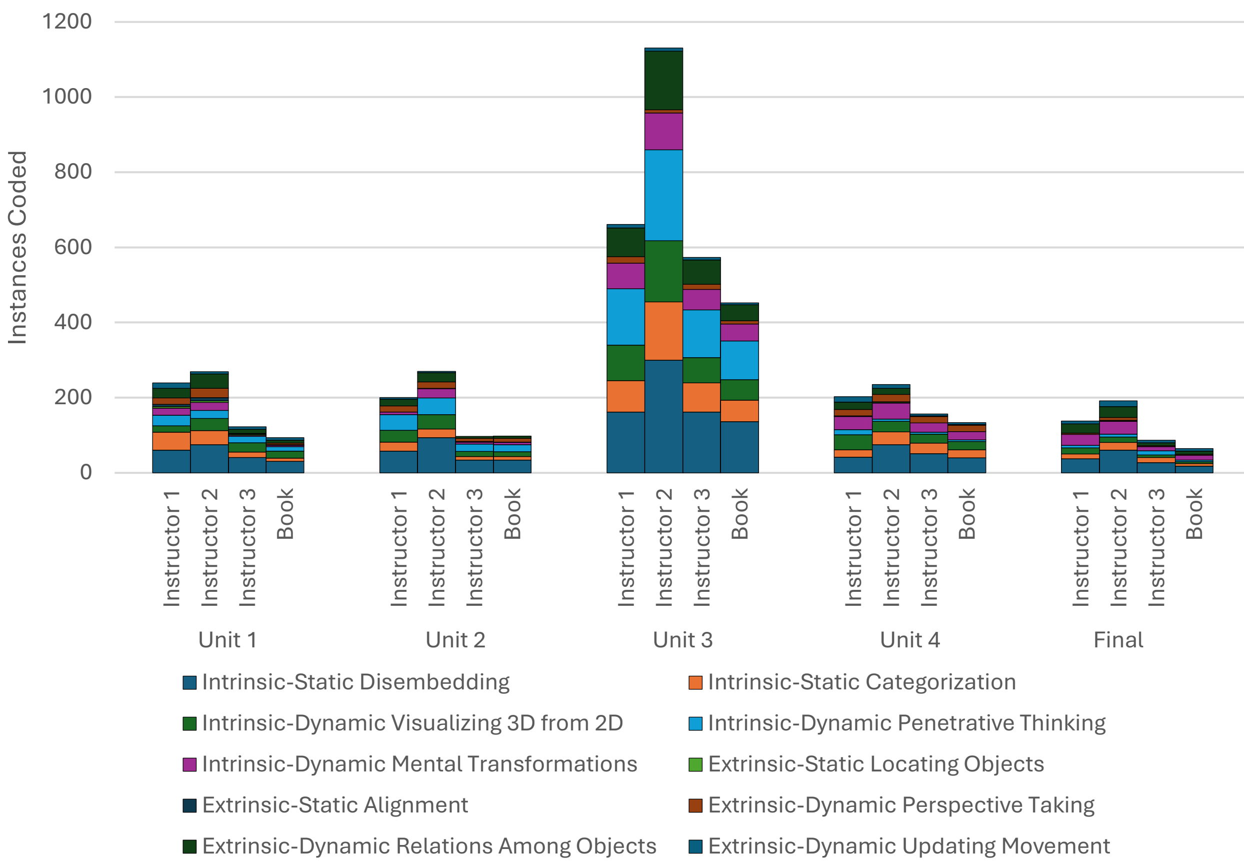
Example of 3D Image Requiring Spatial Reasoning from CHEM 111 Instructor Slides



The 10 types of spatial reasoning skills within the 2x2 framework identified from geoscience (Newcombe & Shipley, 2015) found in this chemistry curriculum are presented by unit in Figure 7. The numbers of instances of each of the 10 skills and the percentage of total instances of each skill is shown for each set of slides and exams separated by unit in Tables A6, A7, A8, A9, and A10 in Appendix.

Figure 8

Frequency of 10 spatial reasoning skills across instructional materials in five units of CHEM 111



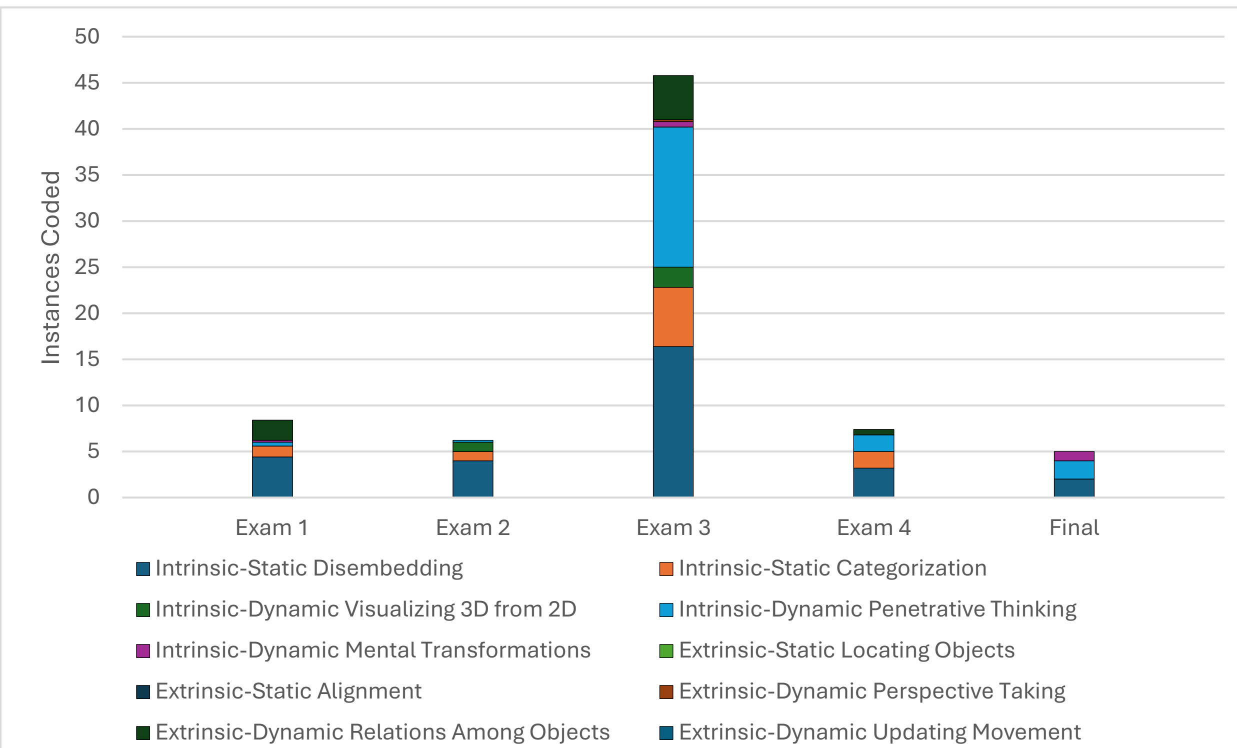
Disembedding was the most frequent spatial reasoning skill across all units, instructors, and exams. Disembedding involves identifying and isolating relevant parts of a complex display, which is often a prerequisite for all other spatial reasoning skills. If something requires another type of spatial reasoning, it likely requires disembedding as well because to manipulate something students must identify what needs to be manipulated first. While disembedding was the most frequent across all units, the other more frequent spatial reasoning skills were dependent on the unit.

Relations among objects is the between objects equivalent of penetrative thinking within objects, as seen in the example in Figure 7. The most frequent skills within the intrinsic-dynamic and extrinsic-dynamic categories being penetrative thinking and relations among objects makes sense as they employ similar thought processes. Penetrative thinking being more common indicates that in this chemistry course, the focus was on employing these skills for within object thinking and manipulation.

The exams for each unit had less variety of spatial reasoning skills within the questions, as seen in Figure 9.

Figure 9

Frequency of 10 spatial reasoning skills across exams in five units of CHEM 111



Of all the units, one had the most variety of spatial reasoning skills, with all 10 showing up at least once in all sets of slides. The exams had less variety with only disembedding, categorization, penetrative thinking, mental transformations, and relations among objects being present, meaning half of the spatial reasoning skills used in the slides were not required on assessment. For unit two, disembedding and penetrative thinking were the most frequent skills in the instructional slides. This pattern suggests that the instructional materials primarily required students to identify relevant features within representations and mentally reason about internal structures. However, for the exam categorization and visualizing 3D from 2D were more frequent than penetrative thinking, showing a potential misalignment between which skills were focused on in instructional materials and exams. Students may have encountered a somewhat different set of spatial reasoning demands during assessment than those most frequently practiced in class, potentially creating a barrier to their success. Unit three had the most variety of spatial reasoning skills in the exam, with seven being present. However, in the instructional materials eight of the spatial reasoning skills were present, indicating the exam is still not assessing all the types of spatial reasoning used in class. In the unit four exams, only four of the skills were present, disembedding, categorization, penetrative thinking, and relations among objects. In the slides, visualizing 3D from 2D, penetrative thinking, and perspective taking all showed up frequently but were not present on the exams, highlighting again the lack of variety of spatial skills in the exams compared to instructional materials.

When applying the typology developed by Newcombe and Shipley to identify the types of spatial reasoning to chemistry learning, the only skill of the 11 not found was sequential thinking. The framework with the 11 skills was originally developed for geoscience. Within this course, there was a focus on small scale spatial reasoning over a short time frame. Within geoscience, there is often more of a focus on large scale objects which change over a longer time frame, requiring sequential thinking as things change slowly in a series of events over time. The short time frame and small scale focused on in this course means things change more at once rather than in a series of transformations over time. This class also focuses more on the start and end states of a display, rather than the mechanisms of the change, leading to a focus on mental transformation rather than sequential thinking, which is made up of a series of mental transformations. Different STEM disciplines require different types of spatial reasoning. Chemistry may not require sequential thinking but still involves skills such as mental rotation (Newcombe & Shipley, 2015). Even though not all 11 of the skills applied to this chemistry course, the 2x2 framework still provided a solid way to identify types of spatial reasoning. Despite the spatial reasoning skills being developed for geoscience, there is enough overlap between the types of skills in chemistry where the same language can be used to describe spatial reasoning in both fields.

The curriculum analyzed in this study was focused on a molecular scale. The course analyzed was a general chemistry course, but different types of spatial reasoning may be required more for other chemistry courses. For example, mental transformations are often required for rotating chiral molecules in organic chemistry (Harle &Towns, 2010). The content covered in CHEM 111 did not focus on extrinsic-static skills as there was more focus on interactions within and between objects, rather than just identifying location. Other STEM disciplines such a geoscience more often use locating objects on a larger reference frame like a map (Newcombe & Shipley, 2015). There was less focus on locating objects on a larger scale in the course because of the general small scale of molecules. Chemistry is affected by the interactions within and between molecules, featuring intrinsic-static, intrinsic-dynamic, and extrinsic-dynamic, rather than extrinsic-static reasoning. It is possible all 11 skills are used in some areas of chemistry, despite not being used for CHEM 111. Even lacking one of the 11 skills from geoscience were present in the material I analyzed for CHEM 111, this framework is a useful tool for categorizing specific spatial reasoning skills used in the course.

One historic problem with teaching spatial reasoning has been a lack of clear definition (McLaughlin & Bailey, 2022). Using the four larger categories of the 2x2 framework as well as the skills within those categories clearly defines spatial reasoning. The identification can be used to clearly communicate when and how spatial reasoning is used in general chemistry. While the specific types of spatial skills as well as the 2x2 framework worked to describe the spatial reasoning within the instructor slides and exams, one suggestion to make this framework more specific includes dividing extrinsic into large scale and small scale (Uttal et al., 2024). This could be applied in future studies to get a more specific description of spatial reasoning, which could be useful if comparing across science disciplines.

CHAPTER V: SUMMARY, IMPLICATIONS, AND LIMITATIONS

By examining the presence and types of spatial reasoning embedded in CHEM 111 instructional materials and assessments, this study provides insight into how spatial reasoning is represented within an introductory chemistry curriculum. The findings highlight patterns in the cognitive level and types of spatial reasoning present across course units and materials, while also illustrating how the spatial reasoning framework developed by Newcombe and Shipley (2015) can be applied in a chemistry education context. Here I discuss the implications of these findings for teaching practice, spatial skill development, and future research, followed by the limitations of the current study.

Summary

This study identified the level of spatial reasoning across instructional materials in CHEM 111 based on Bloom's taxonomy (Arnson and Offerdahl, 2018). Unit three in the course displayed the highest frequency of high-level spatial reasoning, indicating certain topics in general chemistry employ more spatial reasoning than others. There was also different levels of spatial reasoning within instructor materials and exams when from latent and semantic coding, especially within unit three. This finding supports the idea that spatial reasoning could be made more explicit in materials to benefit students when using spatial reasoning. There was also a lower level of spatial reasoning required for exams than present in other course materials. Although additional research could reveal how students use spatial reasoning in the course and on exams, from the analysis there is a low level of spatial reasoning being assessed.

Additionally, this study serves as a proof of concept for using the spatial reasoning framework developed by Newcombe and Shipley in a chemistry context. Applying the 2x2 spatial reasoning framework, I identified intrinsic-static, intrinsic-dynamic, extrinsic-static, and extrinsic-dynamic skills used in CHEM 111. I was also able to apply the framework developed for geoscience from Newcombe and Shipley (2015) to identify specific spatial reasoning skills within the instructional materials and exams. Overall, intrinsic spatial reasoning made up the majority of spatial reasoning present in CHEM 111. Although certain skills, like disembedding, were prevalent throughout the entire course, different units showed varying frequencies of the 10 spatial reasoning skills identified, supporting the concept that different topics within this chemistry course used different types of spatial reasoning skills. Exams also showed less variety in the types of spatial reasoning skills required when compared to the instructional materials for that unit, indicating students are not required to use all the spatial reasoning skills on assessment that are presented in class materials.

Implications

Spatial reasoning is an important indicator of performance, retention, and persistence in STEM fields. However, spatial reasoning has historically been underutilized as an indicator for STEM success. Spatial reasoning is also malleable so focusing on spatial training in STEM courses can help students be successful (Uttal et al., 2013). Chemistry requires spatial reasoning in many aspects. Making spatial reasoning more explicit in CHEM 111 may help to increase performance, retention, and persistence in STEM for the students. High spatial reasoning ability in students has been shown to improve performance in STEM (Sorby et al., 2018). Although all students can benefit from spatial reasoning skills, historically women have had lower spatial reasoning than men, so promoting spatial reasoning may be especially beneficial for female students. The gender differences in spatial reasoning skills is one of the most researched and well documented pieces of spatial reasoning (Sorby et al. 2018). The gap in spatial reasoning skills between men and women disadvantages women. Between the ages of nine and 23, the gap increases with age (Sorby et al., 2018). One study tracked the spatial reasoning skills of men and women in chemistry from the 1980s to the 1990s. They concluded that the gap between men and women decreased overtime. However, their results show that it did not decrease because women's spatial reasoning was getting better, but because men's spatial reasoning got worse over the time frame (Coleman & Gotch 1998). Men have consistently scored higher than women on measure of spatial skills. Both men and women respond positively to training in spatial reasoning. Although training has not closed the gender gap, both men and women show improved performance in spatial reasoning with training (Uttal et al., 2013).

Although fully closing the gender gap in spatial reasoning has not been achieved, it may not be necessary. The main goal of decreasing the gender gap in spatial reasoning has been to decrease the gender gap in stem success (Uttal et al., 2013). Women have been historically underrepresented in STEM fields and there is a goal to bring more representation (NSB, 2010). Although spatial skills are helpful in STEM fields, there are other factors that also contribute to a likelihood of success in STEM. It is possible that there is a lower limit of spatial skills that determines likelihood of entering a STEM field, and any level of spatial skills above the threshold are helpful but not necessarily. Using training in spatial reasoning to close the gender gap in the lower limit would likely help close the gap in entering a STEM field (Uttal et al., 2013). Supporting spatial reasoning explicitly in courses such as CHEM 111 can help bring all students to a baseline of spatial reasoning that is needed for success in STEM education and careers.

Spatial reasoning skills are correlated with students' grades in STEM courses, and intervention in spatial reasoning can lead to improvement in STEM GPA (Sorby et al., 2018). Making spatial reasoning more explicit in courses can help students recognize when to use it to help with problem solving. Using spatial skills for problem solving can help students stay in a STEM major by alleviating the need for expertise to answer questions (NSB, 2010). There is also no upper limit in the correlation between spatial reasoning and STEM success, even as students become experts. Supporting spatial reasoning skills in general chemistry can help with future career prospects as chemists use spatial reasoning at all levels of expertise (Buckley et al., 2019). This study identified the cognitive level of spatial reasoning in parts of the CHEM 111 course both latently and semantically. This can be used for instructors to reflect on where spatial reasoning can be made more explicit in the curriculum to lessen the difference between semantic level and latent level from Bloom's taxonomy. The CHEM 111 learning objectives already have nearly the same proportion of each level of spatial reasoning when coded semantically and latently. If other instructional materials such as lecture slides and exams are made more explicit, students may recognize when spatial reasoning could improve comprehension for certain concepts.

Additionally, the types of spatial reasoning identified in this study can be used to include more specific spatial language in the curriculum to clearly outline problem-solving techniques that enlist spatial reasoning methods. Having a specific framework for identifying types of spatial reasoning can help make spatial training more explicit through identifying certain skills for chemistry concepts. Explicitly writing when to sketch, visualize in 3D, mentally transform, etc. can help students recognize certain problem-solving techniques in a chemistry context. Spatial reasoning has been extensively shown to help students be successful in STEM, so explicitly supporting spatial reasoning in CHEM 111 could benefit students taking general chemistry or higher-level chemistry courses. Using the types of spatial reasoning from the Newcombe and Shipley (2015) framework applied in this study can give instructors specific language to explicitly train types of spatial reasoning students need to be successful in chemistry. It is imperative for instructors to be aware of types of spatial reasoning students need so they can incorporate them in the course. However, experts have content knowledge that alleviates the need for them to use all the spatial reasoning skills that benefit students. This study lays a foundation for identifying specific types of spatial reasoning within a general chemistry course so instructors can explicitly identify and train students in spatial reasoning that will help them be successful.

This study also identified a misalignment between the level and types of spatial reasoning between instructional materials and exams. Exams had less level two and three spatial reasoning, less variety of spatial skills, and focused on different spatial skills than instructional materials. This could disincentivize students to practice spatial skills that are not required on exams, even if those skills could benefit student performance. Many students in CHEM 111 continue to take organic chemistry, where spatial ability is a predictor of success (Stieff et al., 2012). Spatial training is known to be durable, so practicing a variety of spatial skills in CHEM 111 could set students up for success in future chemistry classes (Uttal et al., 2013). While students may rely less on spatial problem-solving skills after instruction in organic chemistry, spatial problem-solving skills are still the most common strategies employed by students (Stieff et al., 2012). While experts in organic chemistry have more analytical strategies, they often still use spatial reasoning to support those analytical strategies when solving problems (Stieff & Raje, 2010). Course materials and exams in CHEM 111 should support students in using spatial reasoning as it could lead to increased performance not only in this course, but future courses and careers.

One difficulty in spatial reasoning studies has been creating a consistent framework to apply when studying spatial reasoning skills. While this study shows how the 2x2 framework of intrinsic-static, intrinsic-dynamic, extrinsic-static, and extrinsic-dynamic spatial reasoning as well as the specific skills within each category can be used to describe spatial reasoning in a chemistry context, there is potential for this framework to be made more specific by distinguishing between large-scale and small-scale within extrinsic (Uttal et al., 2024). Future studies could expand on this framework by further specifying types of spatial reasoning using a 2x3 model, including intrinsic-static, intrinsic-dynamic, small-scale extrinsic-static, small-scale extrinsic-dynamic, large-scale extrinsic-static, and large-scale extrinsic-dynamic, which may be helpful in distinguishing when each type of spatial reasoning is relevant, especially if comparing across STEM disciplines, where some areas focus more on small-scale and some focus on large-scale spatial reasoning. These results demonstrate the usefulness of applying the spatial reasoning framework developed by Newcombe and Shipley (2015) to chemistry contexts while also revealing that spatial reasoning demands in introductory chemistry materials are frequently implicit rather than explicitly communicated to students.

Limitations

This study only considered instructional materials and exams available on the Canvas LMS for CHEM 111 and slides provided by the book publisher, excluding any material from external homework platforms or instructor delivery. Although spatial reasoning requirements may not have been made explicit in the written materials, instructors could have given more detail during the class session that provided additional support for spatial reasoning. Although some information about instructor input was gathered by comparing instructor materials to the book publisher slides, this information is limited. Future work could address this limitation through observational studies that examine classroom instruction directly, particularly focusing on the spatial language instructors use and the verbal clarifications they provide alongside visual materials. Such work could reveal additional ways in which instructors support spatial reasoning that are not captured in written course materials.

This study also only applied the spatial reasoning framework to one course at one institution to show how the framework can be applied in a chemistry context, but does not provide any conclusive evidence for how spatial reasoning is represented in chemistry at large. Additionally, all comparisons within this study are descriptive rather than statistically significant. This study aims to provide a basis of how this framework could be used in future work where further statistical analysis could be applied. Future studies could apply this framework across multiple courses, institutions, or instructors where statistical analyses could be conducted.

Additional research could also investigate how chemistry students use spatial reasoning in problem solving, as they could be using spatial reasoning even when not explicitly instructed. The two codebooks used were also not integrated for this study. Any comparison between the cognitive level of spatial reasoning and the type of spatial reasoning was not investigated. Further analysis could explore the correlation between the cognitive level of spatial reasoning and type.

Overall, this study contributes to the growing body of research examining the role of spatial reasoning in STEM education by demonstrating how spatial reasoning frameworks can be systematically applied to a chemistry curriculum. By identifying both the cognitive level and types of spatial reasoning embedded within CHEM 111 instructional materials and assessments, this work provides insight into how spatial reasoning demands are distributed across course topics and how often those demands remain implicit rather than explicitly communicated to students. Methodologically, this study establishes a proof of concept for using the spatial reasoning frameworks developed by Newcombe and Shipley (2015) and Arnson and Offerdahl (2018) to analyze chemistry instructional materials, offering a structured approach that future researchers can expand across courses, institutions, and STEM disciplines. Although the findings are descriptive and limited to a single course context, they highlight important opportunities for making spatial reasoning more explicit in chemistry instruction and for supporting students in developing spatial skills that contribute to success in STEM. Future research can build on this foundation by incorporating classroom observations, examining student problem-solving processes, and exploring how explicit spatial reasoning instruction influences learning outcomes and persistence in chemistry.

REFERENCES

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

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Buckley, J., Seery, N., & Canty, D. (2018). A Heuristic Framework of Spatial Ability: a Review and Synthesis of Spatial Factor Literature to Support its Translation into STEM Education. *Educational Psychology Review*, *30*(3), 947–972. https://doi.org/10.1007/s10648-018-9432-z

Buckley, J., Seery, N., & Canty, D. (2018). Investigating the use of spatial reasoning strategies in geometric problem solving. *International Journal of Technology and Design Education*, *29*(2), 341–362. https://doi.org/10.1007/s10798-018-9446-3

Coleman, S. L., & Gotch, A. J. (1998). Spatial Perception Skills of Chemistry Students. *Journal of Chemical Education*, *75*(2), 206. https://doi.org/10.1021/ed075p206

Harle, M., & Towns, M. (2010). A Review of Spatial Ability Literature, Its Connection to Chemistry, and Implications for Instruction. *Journal of Chemical Education*, *88*(3), 351–360. https://doi.org/10.1021/ed900003n

Hawes, Z. C. K., Gilligan-Lee, K. A., & Mix, K. S. (2022). Effects of spatial training on mathematics performance: A meta-analysis. *Developmental Psychology, 58*(1), 112–137. https://doi.org/10.1037/dev0001281

Laricheva, E. N., & Ilikchyan, A. (2023). Exploring the Effect of Virtual Reality on Learning in General Chemistry Students with Low Visual-Spatial Skills. *Journal of Chemical Education*, *100*(2), 589–596. https://doi.org/10.1021/acs.jchemed.2c00732

Lowrie, T., Logan, T., Harris, D., & Hegarty, M. (2018). The impact of an intervention program on students' spatial reasoning: student engagement through mathematics-enhanced learning activities. *Cognitive Research: Principles and Implications*, *3*(1). https://doi.org/10.1186/s41235-018-0147-y

McLaughlin, J. A., & Bailey, J. M. (2022). Students need more practice with spatial thinking in geoscience education: a systematic review of the literature. *Studies in Science Education*, 1–58. https://doi.org/10.1080/03057267.2022.2029305

Newcombe , N. (2017), "Harnessing Spatial Thinking to Support Stem Learning ", *OECD Education Working Papers*, No. 161, OECD Publishing, Paris, https://doi.org/10.1787/7d5dcae6-en.

Newcombe , N. S., & Shipley, T. F. (2015). Thinking About Spatial Thinking: New Typology, New Assessments. *Studying Visual and Spatial Reasoning for Design Creativity*, 179–192. https://doi.org/10.1007/978-94-017-9297-4_10

Oliver-Hoyo, M., & Babilonia-Rosa, M. A. (2017). Promotion of Spatial Skills in Chemistry and Biochemistry Education at the College Level. *Journal of Chemical Education*, *94*(8), 996–1006. https://doi.org/10.1021/acs.jchemed.7b00094

PREPARING THE NEXT GENERATION OF STEM INNOVATORS: Identifying and Developing our Nation's Human Capital. (2010). https://nsf-gov-resources.nsf.gov/nsb/publications/2010/nsb1033.pdf

Sorby, S., Veurink, N., & Streiner, S. (2018). Does spatial skills instruction improve STEM outcomes? The answer is "yes." *Learning and Individual Differences*, *67*, 209–222. https://doi.org/10.1016/j.lindif.2018.09.001

Stieff, M., & Raje, S. (2010). Expert algorithmic and imagistic problem solving strategies in advanced chemistry. *Spatial Cognition and Computation, 10*(1), 53–81. https://doi.org/10.1080/13875860903453332

Stieff, M., Ryu, M., Dixon, B., & Hegarty, M. (2012). The Role of Spatial Ability and Strategy Preference for Spatial Problem Solving in Organic Chemistry. *Journal of Chemical Education*, *89*(7), 854–859. https://doi.org/10.1021/ed200071d

Terlecki, M. S., Newcombe , N. S., & Little, M. (2007). Durable and generalized effects of spatial experience on mental rotation: gender differences in growth patterns. *Applied Cognitive Psychology*, *22*(7), 996–1013. https://doi.org/10.1002/acp.1420

Uttal, D. H., & Cohen, C. A. (2012). Spatial Thinking and STEM Education: When, Why, and How (p. 57). Amsterdam: Elsevier Inc. https://doi.org/10.1016/B978-0-12-394293-7.00004-2

Uttal, D. H., McKee, K., Simms, N., Hegarty, M., & Newcombe , N. S. (2024). How Can We Best Assess Spatial Skills? Practical and Conceptual Challenges. *Journal of Intelligence*, *12*(1), 8. https://doi.org/10.3390/jintelligence12010008

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe , N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, *139*(2), 352–402.

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, *101*(4), 817–835. https://doi.org/10.1037/a0016127

Yeaman, A., Bairaktarova, D., & Knott, T. (2020). *Developing a Coding Rubric for Students' Spatial Visualization Strategies*. https://doi.org/10.18260/1-2--30289

Zhu, C., Chloe Oi-Ying Leung, Lagoudaki, E., Velho, M., Segura-Caballero, N., Jolles, D., Duffy, G., Maresch, G., Pagkratidou, M., & Remke Klapwijk. (2023). Fostering spatial ability development in and for authentic STEM learning. *Frontiers in Education*, *8*. https://doi.org/10.3389/feduc.2023.1138607

APPENDIX

Frequency and Percentage of Levels and Types of Spatial Reasoning within CHEM 111

Table A1

Frequency (n) and percentage (%) of level of spatial reasoning in all learning objectives when coded semantically and latently, with levels adapted from Bloom's taxonomy with 0 indicating no spatial reasoning, 1 indicating low level spatial reasoning, 2 indicating mid level spatial reasoning, and 3 indicating high levels of spatial reasoning.

	Level of Spatial Reasoning	n	%
Semantic	0	159	69
	1	27	12
	2	31	13
	3	14	6
Latent	0	156	68
	1	28	12
	2	32	14
	3	15	6

Table A2

Frequency (n) and Percentage (%) of each Level of Spatial Reasoning in Slides and exams for each Unit of CHEM 111 Coded Semantically and Latently

		Instructor 1		Instructor 2		Instructor 3		Book		Unit Exam	
Level of Spatial Reasoning		n	%	n	%	n	%	n	%	n	%
Unit 1											
Semantic	0	82	85	80	82	60	97	18	100	31	88
	1	8	8	10	10	1	2	0	0	4.2	12
	2	5	5	6	6	1	2	0	0	0.2	1
	3	2	2	2	2	0	0	0	0	0	0
Latent	0	76	78	79	81	59	95	18	100	31	88
	1	14	14	11	11	2	3	0	0	4.2	12
	2	5	5	6	6	1	2	0	0	0.2	1
	3	2	2	2	2	0	0	0	0	0	0
Unit 2											
Semantic	0	79	87	122	87	78	94	26	93	34.8	90
	1	10	11	19	14	4	5	2	7	4	10
	2	1	1	1	1	1	1	0	0	0	0
	3	1	1	1	1	0	0	0	0	0	0
Latent	0	74	81	118	84	77	93	26	93	34.6	89
	1	13	14	23	16	5	6	2	7	4.2	11
	2	3	3	3	2	1	1	0	0	0	0
	3	1	1	1	1	0	0	0	0	0	0
Unit 3											
Semantic	0	24	20	61	34	21	30	5	17	16.2	47
	1	29	24	51	28	15	21	3	10	15.6	45
	2	49	41	64	35	29	41	16	55	2.6	7
	3	17	14	7	4	6	8	5	17	0.2	1
Latent	0	15	13	37	20	10	14	3	10	12.0	35
	1	24	20	59	32	16	23	3	10	15.0	43
	2	57	48	71	39	34	48	16	55	7.0	20
	3	23	19	15	8	11	15	7	24	0.6	2
Unit 4											
Semantic	0	94	86	99	88	72	95	26	100	30.4	90
	1	11	10	7	6	2	3	0	0	3.2	10
	2	2	2	1	1	2	3	0	0	0	0
	3	2	2	5	4	0	0	0	0	0	0
Latent	0	90	83	97	87	72	95	26	100	30.4	90
	1	15	14	8	7	2	3	0	0	3.2	10
	2	1	1	1	1	2	3	0	0	0	0
	3	4	4	6	5	0	0	0	0	0	0
Final											
Semantic	0	83	93	100	90	72	92	23	100	16	94
	1	4	4	7	6	2	3	0	0	0	0
	2	2	2	4	4	4	5	0	0	1	6
	3	0	0	0	0	0	0	0	0	0	0
Latent	0	83	93	100	90	70	90	23	100	15	88
	1	4	4	7	6	2	3	0	0	1	6
	2	2	2	4	4	6	8	0	0	1	6
	3	0	0	0	0	0	0	0	0	0	0

Table A3

Frequency (n) and Percentage (%) of Intrinsic and Extrinsic Spatial Reasoning in Materials for each Unit of CHEM 111

Category of Spatial Reasoning	Unit 1									
	Instructor 1		Instructor 2		Instructor 3		Book		Unit Exam	
	n	%	n	%	n	%	n	%	n	%
Intrinsic	172	72	187	70	100	81	73	78	6.2	74
Extrinsic	67	28	82	30	23	19	21	22	2.2	26
	Unit 2									
Intrinsic	162	81	224	83	82	85	81	83	6.2	100
Extrinsic	38	19	46	17	15	15	17	17	0	0
	Unit 3									
Intrinsic	558	84	958	85	448	85	396	88	40.8	89
Extrinsic	103	16	173	15	85	15	56	12	5	11
	Unit 4									
Intrinsic	150	74	186	79	133	85	110	82	6.8	92
Extrinsic	53	26	49	21	24	15	24	18	0.6	8
	Final									
Intrinsic	102	74	137	71	70	80	47	72	5	100
Extrinsic	36	26	55	29	17	20	18	28	0	0

Table A4

Frequency (n) and Percentage (%) of Static and Dynamic Spatial Reasoning in Materials for each Unit of CHEM 111

Category of Spatial Reasoning	Unit 1									
	Instructor 1		Instructor 2		Instructor 3		Book		Unit Exam	
	n	%	n	%	n	%	n	%	n	%
Static	118	49	125	46	58	47	41	44	5.6	68
Dynamic	121	51	144	54	65	53	53	56	2.8	32
	Unit 2									
Static	82	41	118	44	44	45	44	45	5	81
Dynamic	118	59	152	56	53	55	54	55	1.2	19
	Unit 3									
Static	245	37	455	40	239	42	193	43	22.8	49
Dynamic	416	63	676	60	334	58	259	57	23	51
	Unit 4									
Static	63	31	112	48	79	50	61	46	5	68
Dynamic	140	69	123	52	78	50	73	54	2.4	32
	Final									
Static	50	36	84	44	41	47	25	38	2	40
Dynamic	88	64	108	56	46	53	40	62	3	60

Table A5

Frequency (n) and Percentage (%) of Spatial Reasoning from 2x2 Typology in Materials for each Unit of CHEM 111

Category of Spatial Reasoning	Instructor 1		Instructor 2		Instructor 3		Book		Unit Exam	
	n	%	n	%	n	%	n	%	n	%
Unit 1										
Intrinsic-static	108	45	112	42	55	45	39	41	5.6	68
Intrinsic-dynamic	64	27	75	28	45	37	34	36	0.6	7
Extrinsic-static	10	4	13	5	3	2	2	2	0	0
Extrinsic-dynamic	57	24	69	26	20	16	19	20	2.2	26
Unit 2										
Intrinsic-static	82	41	117	43	43	44	43	44	5	81
Intrinsic-dynamic	80	40	107	40	39	40	38	39	1.2	19
Extrinsic-static	0	0	1	0	1	1	1	1	0	0
Extrinsic-dynamic	38	19	45	17	14	14	16	16	0	0
Unit 3										
Intrinsic-static	245	37	455	40	239	42	193	43	22.8	49
Intrinsic-dynamic	313	47	503	44	249	43	203	45	18	40
Extrinsic-static	0	0	0	0	0	0	0	0	0	0
Extrinsic-dynamic	103	16	173	15	85	15	56	12	5	11
Unit 4										
Intrinsic-static	61	30	109	46	79	50	61	46	5	68
Intrinsic-dynamic	89	44	77	33	54	34	49	37	1.8	25
Extrinsic-static	2	1	3	1	0	0	0	0	0	0
Extrinsic-dynamic	51	25	46	20	24	15	24	18	0.6	8
Final										
Intrinsic-static	50	36	81	42	41	47	25	38	2	40
Intrinsic-dynamic	52	38	56	29	29	33	22	34	3	60
Extrinsic-static	0	0	3	2	0	0	0	0	0	0
Extrinsic-dynamic	36	26	52	27	17	20	18	28	0	0

Table A6

Frequency and Percentage of 10 types of spatial reasoning skills within Instructional Materials in Unit 1

		Unit 1									
		Instructor 1		Instructor 2		Instructor 3		Book		Unit Exam	
Category of Spatial Reasoning	Typology Skill	n	%	n	%	n	%	n	%	n	%
Intrinsic-static	Disembedding	60	25	75	28	41	33	31	33	4.4	53
	Categorization	48	20	37	14	14	11	8	9	1.2	15
Intrinsic-dynamic	Visualizing 3d from 2d	17	7	33	12	25	20	19	20	0	0
	Penetrative Thinking	28	12	21	8	17	14	12	13	0.4	5
	Mental Transformations	19	8	21	8	3	2	3	3	0.2	2
Extrinsic-static	Locating Objects	4	2	5	2	2	2	1	1	0	0
	Alignment	6	3	8	3	1	1	1	1	0	0
Extrinsic-dynamic	Perspective Taking	17	7	25	9	2	2	4	4	0	0
	Relations Among Objects	26	11	38	14	11	9	8	9	2.2	26
	Updating Movement	14	6	6	2	7	6	7	7	0	0

Table A7

Frequency and Percentage of 10 types of spatial reasoning skills within Instructional Materials in Unit 2

		Unit 2									
		Instructor 1		Instructor 2		Instructor 3		Book		Unit Exam	
Category of Spatial Reasoning	Typology Skill	n	%	n	%	n	%	n	%	n	%
Intrinsic-static	Disembedding	58	29	94	35	34	35	34	35	4	65
	Categorization	24	12	23	9	9	9	9	9	1	17
Intrinsic-dynamic	Visualizing 3d from 2d	31	16	38	14	14	14	13	13	1	17
	Penetrative Thinking	42	21	44	16	20	21	19	19	0.2	3
	Mental Transformations	7	4	25	9	5	5	6	6	0	0
Extrinsic-static	Locating objects	0	0	0	0	0	0	0	0	0	0
	Alignment	0	0	1	0	1	1	1	1	0	0
Extrinsic-dynamic	Perspective Taking	16	8	17	6	9	9	10	10	0	0
	Relations Among Objects	18	9	25	9	5	5	6	6	0	0
	Updating Movement	4	2	3	1	0	0	0	0	0	0

Table A8

Frequency and Percentage of 10 types of spatial reasoning skills within Instructional Materials in Unit 3

		Unit 3									
		Instructor 1		Instructor 2		Instructor 3		Book		Unit Exam	
Category of Spatial Reasoning	Typology Skill	n	%	n	%	n	%	n	%	n	%
Intrinsic-static	Disembedding	162	25	300	27	162	28	136	30	16.4	36
	Categorization	83	13	155	14	77	13	57	13	6.4	14
Intrinsic-dynamic	Visualizing 3d from 2d	95	14	163	14	68	12	55	12	2.2	5
	Penetrative Thinking	150	23	242	21	127	22	103	23	15.2	34
	Mental Transformations	68	10	98	9	54	9	45	10	0.6	1
Extrinsic-static	Locating Objects	0	0	0	0	0	0	0	0	0	0
	Alignment	0	0	0	0	0	0	0	0	0	0
Extrinsic-dynamic	Perspective Taking	17	3	8	1	14	2	9	2	0.2	0
	Relations Among Objects	77	12	156	14	65	11	42	9	4.8	11
	Updating Movement	9	1	9	1	6	1	5	1	0	0

Table A9

Frequency and Percentage of 10 types of spatial reasoning skills within Instructional Materials in Unit 3

		Unit 4									
		Instructor 1		Instructor 2		Instructor 3		Book		Unit Exam	
Category of Spatial Reasoning	Typology Skill	n	%	n	%	n	%	n	%	n	%
Intrinsic-static	Disembedding	42	21	75	32	51	32	40	30	3.2	44
	Categorization	19	9	34	14	28	18	21	16	1.8	24
Intrinsic-dynamic	Visualizing 3d from 2d	40	20	28	12	24	15	22	16	0	0
	Penetrative Thinking	14	7	6	3	5	3	5	4	1.8	25
	Mental Transformations	35	17	43	18	25	16	22	16	0	0
Extrinsic-static	Locating objects	1	0	2	1	0	0	0	0	0	0
	Alignment	1	0	1	0	0	0	0	0	0	0
Extrinsic-dynamic	Perspective Taking	17	8	20	9	17	11	17	13	0	0
	Relations Among Objects	19	9	16	7	2	1	2	1	0.6	8
	Updating Movement	15	7	10	4	5	3	5	4	0	0

Table A10

Frequency and Percentage of 10 types of spatial reasoning skills within Instructional Materials in the Final Unit

		Final									
		Instructor 1		Instructor 2		Instructor 3		Book		Unit Exam	
Category of Spatial Reasoning	Typology Skill	n	%	n	%	n	%	n	%	n	%
Intrinsic-static	Disembedding	37	27	60	31	27	31	18	28	2	40
	Categorization	13	9	21	11	14	16	7	11	0	0
Intrinsic-dynamic	Visualizing 3d from 2d	16	12	14	7	7	8	5	8	0	0
	Penetrative Thinking	7	5	8	4	11	13	4	6	2	40
	Mental Transformations	29	21	34	18	11	13	13	20	1	20
Extrinsic-static	Locating Objects	0	0	1	1	0	0	0	0	0	0
	Alignment	0	0	2	1	0	0	0	0	0	0
Extrinsic-dynamic	Perspective Taking	4	3	7	4	2	2	2	3	0	0
	Relations Among Objects	24	17	29	15	8	9	9	14	0	0
	Updating Movement	8	6	16	8	7	8	7	11	0	0