8-1-2015

Development of a Stereochemistry Concept Inventory

Alexey Leontyev

UNIVERSITY OF NORTHERN COLORADO
Greeley, Colorado
The Graduate School

DEVELOPMENT OF A STEREOCHEMISTRY CONCEPT INVENTORY

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy

Alexey Leontyev

College of Natural and Health Sciences
Department of Chemistry and Biochemistry
Chemical Education Program

August, 2015

This Dissertation by: Alexey Leontyev

Entitled: *Development of a Stereochemistry Concept Inventory*

has been accepted as meeting the requirement for the Degree of Doctor of Philosophy in the College of Natural and Health Sciences in the Department of Chemistry and Biochemistry, Program of Chemical Education.

Accepted by the Doctoral Committee

____
Richard M. Hyslop, Ph.D., Research Advisor

____
Kimberly A. O. Pacheco, Ph.D., Committee Member

____
David L. Pringle, Ph.D., Committee Member

____
Youngjin Song, Ph.D., Committee Member

____
Steven Pulos, Ph.D., Faculty Representative

Date of Dissertation Defense ______________________________

Accepted by the Graduate School

____
Linda L. Black, Ed.D.
Associate Provost and Dean
Graduate School and International Admissions

ABSTRACT

Leontyev, Alexey. *Development of a Stereochemistry Concept Inventory*. Published Doctor of Philosophy dissertation, University of Northern Colorado, 2015.

The purpose of this study was to develop a concept inventory to assess students' knowledge of stereochemical concepts. Different rigorous methods were employed to ensure quality of the assessment instrument. Two national surveys were conducted to investigate which stereochemistry topics are important and to collect feedback on potential questions. Several methods were used to detect incorrect ideas about stereochemistry that were used to compose distracters for the Stereochemistry Concept Inventory items. Items were mapped onto a blueprint and corresponding content validity indices were measured to warrant suitability of the instrument for classroom assessment. Several pilot tests were conducted at different institutions and psychometric quality of items was investigated followed by revisions of problematic items. Overall, the newly developed Stereochemistry Concept Inventory is a useful tool that can provide practitioners with information about abundance of different incorrect ideas that students have developed or provide insights on relative efficiency of intervention methods.

ACKNOWLEDGMENTS

I would like to acknowledge several people that made my stay at the University of Northern Colorado valuable.

First, I shall thank my graduate advisor, Dr. Richard M. Hyslop for his help that he provided at all stages of this dissertation project. I am particularity grateful to him for reading multiple versions of my drafts and timely return of them. His honest feedback and helpful advice were constantly pushing me to produce better research and succinct writing.

I also would like to acknowledge Dr. Steven Pulos for all of the outstanding teaching in his measurement classes that I was fortunate to take. His ideas and views on measurement became extremely important for the development of my dissertation project. And I am also very grateful to him for all valuable conversations about my projects that we had. Most of knowledge about measurement and assessment I have leaned I owe to Steven.

I am thankful to Dr. Kimberly A. O. Pacheco, Dr. David L. Pringle, and Dr. Youngjin Song for taking their time to serve on my committee and all feedback that they provided at various stages of my dissertation research. It has been an honor to have such dedicated faculty members on my committee.

I would also like to acknowledge the help of the professors that have implemented the Stereochemistry Concept Inventory in their classrooms, and all students who agreed

to participate in this project. I would not be able to complete this project without numerous instructors opening their classrooms for data collection, and students letting me use their responses.

Last, but not least, I would like to thank Dr. Susan Hutchinson and Dr. Monica Geist for all knowledge about research methodology and statistics that I gained in their classes. After all, a research project is as strong, as the methodology that it is based on.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

CHAPTER I

INTRODUCTION

One of the reasons that chemistry courses are difficult is that students enter these courses having many incorrect ideas which are also called "alternative conceptions." Quite often those alternative conceptions remain unchanged even after the course ends (Nakhleh, 1992). The term "alternative conceptions" is used in the current chemistry education literature most often when the discussion deals with conceptions different from scientifically correct notions (McClary & Bretz, 2012; Rushton, Hardy, Gwaltney, & Lewis, 2008; Talanquer, 2006). The terms "misconceptions" (Cheung, Ma, & Yang, 2009), "alternate ideas" (Mulford & Robinson, 2002), "naïve ideas" (Stavy, 1990), and "incorrect ideas" (Villafañe, Loertscher, Minderhout, & Lewis, 2011) are often used interchangeably when describing the same phenomena. The debates concerning the use of certain terms began more than 25 years ago (Abimbola, 1988) and are still in progress. For this study, the term "incorrect ideas" is used in the instances covering conceptions that are different from scientifically accepted notions. The term "incorrect ideas" is used only for the students of the sample frame; the original terms from the literature are used in corresponding discussions.

A large body of research has been performed to identify college students' conceptions in the area of general chemistry. Several researchers have reported that

students have misconceptions in numerous areas of general chemistry such as chemical bonding (Nakhleh, 1992), acid-base reactions (Nakhleh, 1994), and the particulate nature of matter (Yezierski & Birk, 2006). Several detailed reviews are available (Barke, Hazari, & Yitbarek, 2008; Taber, 2002) that cover more than two hundred misconceptions. Less research has been done on uncovering misconceptions for advanced chemistry courses. A few studies have addressed physical chemistry topics such as thermodynamics (Granville, 1985) and biochemistry topics (Bretz & Linenberger, 2012) such as enzyme-substrate interactions. Several studies suggest that students in these courses not only retain the prior formulated alternative conceptions but also develop misconceptions regarding the new topics taught. According to constructivism (Bodner, 1986), students attempt to use prior conceptions when they try to learn new material. Fourth year students (Rushton et al., 2008) and prospective teachers (Canpolat, Pinarbasi, & Sözbilir, 2006) also possess many misconceptions about chemistry.

Of the courses and corresponding chemistry areas that have received less attention in incorrect ideas research, organic chemistry is the one typically with the largest student enrollment. The most explored areas of organic chemistry are acid strength (McClary & Talanquer, 2011) and alkenes (Şendur, 2012). These account for only a small portion of the whole body of knowledge usually taught in a typical college-level organic chemistry course. Several topics from organic chemistry courses have been reported as being the most challenging to students (Duis, 2011) which include reaction mechanisms, acid-base chemistry, synthesis, and stereochemistry. Reaction mechanisms and synthesis are very broad topics, and they are foundational in the majority of topics covered in organic chemistry. Acid-base chemistry is a topic covered in general chemistry. Of the topics

reported as most challenging, stereochemistry is the only topic presented as a separate chapter in organic chemistry textbooks. Stereochemistry concepts are applied throughout the course, but the difference from other challenging topics is that stereochemistry is introduced as a separate chapter relatively early in the organic chemistry course.

Stereochemistry is often a source of confusion when students encounter it for the first time and even after multiple exposures (Bowen & Bodner, 1991). Failure to master stereochemical concepts can be a serious impediment to succeeding in an organic chemistry course (Beauchamp, 1984). Moreover, since stereochemical concepts are revisited in some of the upper-division courses, lack of conceptual understanding in stereochemistry can be a barrier for later courses such as inorganic chemistry, biochemistry, or spectroscopy.

Conceptual understanding of a certain area of knowledge can be assessed in various ways. One of the frequently used methods in science education is to administer a concept inventory. A concept inventory is a multiple-choice, diagnostic assessment test that probes the understanding of a single topic (Bailey, 2009) also called a construct (Wilson, 2005). Distracters for questions are composed typically from students' misconceptions.

Other methods to test conceptual understanding are also available, such as in-depth interviews or concept maps. However, data collection by means of in-depth interviews or concept maps requires much time and effort. Interviews yield massive qualitative data, while analysis of concept maps provides both qualitative and quantitative data. Direct interpretation and analysis of such pieces of data are not always possible, and perhaps conducting cognitive interviews is not a reasonable approach for the classroom

instructor. Thus, practitioners cannot always benefit from using interviews or concept maps analysis in their classrooms, especially when dealing with large student populations. On the contrary, concept inventories are the most suitable assessment method for classroom use, and they provide easily interpretable data.

Research Purpose and Rationale

The purpose of the current research study was to develop a concept inventory which assesses students' understanding of the stereochemistry concepts considered important by experts. An additional purpose of this research study was to investigate if this inventory produced reliable data and valid inferences. This inventory is referred as Stereochemistry Concept Inventory in this manuscript.

Diagnostic tools are necessary to assess students' conceptual understanding. When used for instructional improvement or educational research, these diagnostic tools must produce reliable and valid data. In science education, especially in physics and biology, much effort has been invested to develop concept inventories. Some concept inventories are available for selected topics in general chemistry: Lewis structures (Cooper, Underwood, & Hilley, 2012), moles (Krishnan & Howe, 1994), Le Châtelier's principle (Voska & Heikkinen, 2000), and the particulate nature of matter (Nyachwaya et al., 2011). Comprehensive diagnostic tools that assess the content of an entire general chemistry course are available (Mulford & Robinson, 2002; Pavelich, Jenkins, Birk, Bauer, & Krause, 2013). The only instrument reported for organic chemistry identifies alternative conceptions related to acid strength (McClary & Bretz, 2012).

A critical need for the development of new tools for an assessment of concepts has been established by the National Research Council (2012) in their recent report

"*Discipline-Based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering*." Having a wide range of assessment instruments improves a curriculum change process by providing data-based evidence. The need for curriculum reform and call for diagnostic instruments beyond general chemistry has also been emphasized in chemistry education literature (Holme et al., 2010).

After an exhaustive search of the literature, no concept inventory that assesses students' understanding of stereochemistry was found. However, from a survey of organic chemistry instructors, Duis (2011) reported that instructors identified stereochemistry as one of the most challenging and foundational areas of chemistry. For research and instructional purposes, an assessment tool that measures the level of understanding of the concepts that relate to stereochemistry. To this researcher's knowledge, there is no assessment tool reported in the chemistry education research literature that measures the understanding of stereochemistry would be beneficial.

Research Questions

Over the course of this study, specific research questions were addressed. To develop an appropriate diagnostic test, the content of a specific field of knowledge should be clearly defined (Treagust, 1988). The content of the Stereochemistry Concept Inventory was defined in research question 1.

Q1 What stereochemistry topics do organic chemistry instructors consider important?

A survey of organic chemistry instructors was conducted to answer Q1. The survey probed the content covered in stereochemistry instruction within the organic

chemistry curriculum and the learning expectations. Stereochemistry topics and learning objectives that instructors consider important were confirmed for the appropriate coverage in the most commonly used organic chemistry textbooks.

The second area considered in the development process involved obtaining information about students' incorrect ideas. A few studies (Krylova, 1997; Mdachi, 2012) report students' incorrect ideas of stereochemistry, relating mostly to a visualization of organic molecules and *R/S* nomenclature rules. The number and range of misconceptions reported were insufficient to develop an instrument assessing the broad spectrum of concepts taught in stereochemistry. Therefore, Q2 was developed to address students' incorrect ideas.

Q2 What incorrect ideas do organic chemistry students hold regarding stereochemistry?

Incorrect ideas were identified in clinical interviews with students or open-ended questions. In the interviews, students were asked conceptual questions about stereochemistry and to provide their reasoning for their answers. Also, several conceptual questions were given to organic chemistry students on lecture and laboratory quizzes. An analysis of interviews and responses for common themes was done to uncover incorrect ideas.

Using incorrect ideas both from this study (Q2) and appropriate literature, a set of multiple-choice questions covering topics that instructors consider relevant (Q1) was generated. This set of questions constituted a preliminary draft of the Stereochemistry Concept Inventory and was subjected to multiple validation trials both with students and instructors. Reliability and validity were addressed in research question 3.

Q3 Can the Stereochemistry Concept Inventory produce reliable data and valid inferences for the assessment of important concepts of stereochemistry?

Benefits of the Study

In this study, incorrect ideas of stereochemistry were identified and a concept inventory was developed. The instrument, measuring the presence and abundance of incorrect ideas, can potentially be useful not only for organic chemistry instructors, but also for instructors whose courses require stereochemistry concepts as a prerequisite.

Effectiveness of Instruction

A number of studies reported improvement of stereochemistry instruction by various means including programmed instruction (Kurbanoglu, Taskesenligil, & Sözbilir, 2006), card games (Costa, 2007), and combining laboratory experiments with molecular modeling (Clausen, 2011). A common drawback of these studies is a lack of an assessment tool that can be used to measure an improvement in students' conceptual understanding of stereochemistry as a result of targeted instruction intervention.

Cooper (2007) reported those instructors who develop educational strategies often try to promote them even if there is only anecdotal evidence of their effectiveness. Quite often they do not engage in collecting evidence to support the effectiveness of their method of instruction. Use of tests and surveys constructed by researchers themselves without evidence of validity is quite common in educational research.

Previously developed concept inventories have been instrumental in measuring the effectiveness of alternative instruction methods. For example, Hake (1997) reported using the Force Concept Inventory to measure the effectiveness of active-engagement versus traditional methods of instruction in introductory physics courses. Likewise, the

Stereochemistry Concept Inventory could be used to measure the effectiveness of various instructional methods.

Formative Assessment

The Stereochemistry Concept Inventory may be used as a brief formative assessment tool at the end of the lecture module on stereochemistry. Instructors may use this inventory to test their students' understanding of stereochemistry concepts and quickly decide if any additional clarification on certain concepts is needed.

Pre-assessment or Placement Test

The instrument may be given as a pre-test for students who are starting advanced courses such as biochemistry or advanced organic chemistry, which use stereochemistry concepts as a foundation. Based on the results, instructors teaching these courses can decide if they need to revisit specific stereochemistry concepts.

Curriculum Design

The data obtained by administering the instrument to students could provide information about the presence and relative abundance of incorrect ideas. The data pertaining to the prevalence of students' alternative conceptions about stereochemistry may guide chemistry instructors to better design curriculum to facilitate learning stereochemistry.

Clickers

Teaching in a large classroom is a challenging experience because it is difficult for instructors to interact with students directly (Asirvatham, 2009). This limits the amount of feedback an instructor can provide to students. Clickers, which have become widely used in college chemistry classrooms (MacArthur & Jones, 2008), can partially

solve the interaction problem. However, one of the problems often reported by instructors who use clickers is the limited number of questions available. The questions from the Stereochemistry Concept Inventory could be used with clickers.

Limitation of the Study

The study has several limitations that are listed below along with possible ways to minimize their impact on the findings.

Self-selection Bias

Students who volunteer for interviews may be different from the rest of the population with regard to their communication ability or reasoning level.

Hawthorne Effect

When people know they are participating in a research study, they may act differently as opposed to a normal setting. This behavioral pattern is known as the Hawthorne effect. Educational research is not free of the Hawthorne effect (Cook, 1962). An attempt to minimize this effect was made by interviewing volunteers or by giving a low stakes test; however, it is impossible to eliminate this effect completely.

Non-response Bias

The response rate on surveys of instructors rarely reaches 30% (Emenike & Holme, 2012). The interpretation of the survey results may be biased if people who did not respond are different from those who complete the survey.

Definition of Terms

Alternate-form reliability is an approach to estimating test reliability in which individuals' score on one version of a test are correlated with their scores on a different version of the test (Gall, Gall, & Borg, 2003).

Concept inventory is a multiple-choice instrument that focuses on a single topic or small subset of closely related topics, containing numerous questions on each idea in order to gauge a student's understanding of the content (Bailey, 2009).

Constant comparison is a process for analyzing qualitative data to identify categories, to create sharp distinctions between categories, and to decide which categories are theoretically significant (Gall et al., 2003).

Construct is a concept that is inferred from commonalities among observed phenomena and that can be used to explain those phenomena. In theory development, a concept that refers to a structure or process that is hypothesized to underlie particular observable phenomena (Gall et al., 2003).

Construct validity is the extent to which inferences from a test's scores accurately reflect the construct that the test is claimed to measure (Gall et al., 2003)

Correlational coefficient is a mathematical expression of the direction and magnitude of the relationship between two measured variables (Gall et al., 2003).

Effect size is an estimate of the magnitude of difference in the population represented by a sample (Gall et al., 2003).

Expert is an instructor who has taught an organic chemistry course at a tertiary educational institution within last three years.

Incorrect ideas are conceptions that are different from scientifically correct notions.

Internal consistency is an approach to estimating test reliability that examines the extent to which individuals who respond one way to a test item tend to respond the same way to other items on the test (Gall et al., 2003).

Percentile is a type of a rank score that represents a given raw score on a measure as the percentage of individuals in the norming group whose score falls below that score (Gall et al., 2003).

Process validity is a judgment about the credibility of an action research project based on the adequacy of the process used in different phases of the project (Gall et al., 2003).

Stereochemistry is an organic chemistry topic in the first semester of a two-semester college-level organic chemistry topic, focusing on foundational topics of spatial arrangements of the atoms in molecules.

Survey is an attempt to collect data from members of a population in order to determine the current status of that population with respect to one or more variables (Gay, 1992).

Test reliability is the extent to which there is measurement error present in the scores yielded by a test (Gall et al., 2003).

Test-retest reliability is an approach to estimating test reliability in which individuals' score on a test administered at one point in time are correlated with their scores on the same test administered at another point in time (Gall et al., 2003).

Theoretical framework is the underlying structure that scaffolds the study (Merriam, 2009).

Validity is the appropriateness, meaningfulness, and usefulness of specific inferences made from the scores (Gall et al., 2003).

CHAPTER II

LITERATURE REVIEW

The knowledge of stereochemistry is widely believed to be fundamental to understanding the molecular level of chemistry. However, stereochemistry is often a source of confusion for students, even after multiple exposures (Bhattacharyya & Bodner, 2005). The most serious impediment in learning stereochemistry is visualization of three-dimensional aspects of the molecule (Brand & Fisher, 1987). Upper-division courses, such as biochemistry or advanced organic chemistry, use stereochemistry knowledge as foundational.

Incorrect Ideas in Science

Among others, incorrect ideas are one of the most important factors that prevent students' meaningful and permanent learning (Köse, 2008). Incorrect ideas are what students themselves develop erroneously and differ from scientifically accepted concepts. The origins of students' incorrect ideas may vary and depend on educational level, complexity and level of abstraction of the concept, cultural-specific content, and other variables. The most cited sources of misconceptions are preconceived ideas about the natural world (Talanquer, 2006), misapplication of principles (Wenning, 2008), and word usage and analogies (Taber, 1998). Mass media can be a source of scientific alternative

conceptions as well (Stamm, Clark, & Eblacas, 2000). Determining incorrect ideas is one of the major focuses in educational research at all levels.

Research has concluded that misconceptions are not specific to any age, ability, gender, or race (Steif & Dantzler, 2005). The frequencies of some misconceptions appear to change very little with time, and some incorrect ideas in particular domains appear to be more resistant to change to correct ones than others (Nakhleh, 1992). Misconceptions and incorrect ideas may be overcome if classroom instruction involves strategies to facilitate conceptual change (e.g. strategies based on cognitive conflict). A review of misconceptions by Wandersee, Mintzes and Novak in the Handbook of Research on Science Teaching and Learning (1994) presents eight claims which are drawn from multiple studies with high convergence of findings. Expressed succinctly, these claims are:

1. Learners come to formal science instruction with a diverse set of misconceptions about natural objects and events.
2. Misconceptions are robust with respect to age, ability, gender, and culture.
3. Misconceptions are persistent and resistant to change by conventional instruction strategies.
4. Misconceptions often resemble previous explanations of natural phenomena offered by an earlier generation of scientists.
5. Misconceptions are rooted to learners' background, language, as well as teachers' explanations and instructional materials.
6. Teachers often possess alternative conceptions similar to their students.
7. Formal instruction interacts with prior knowledge leading to unintended learning outcomes.
8. Instructional approaches leading to conceptual change can be effective classroom tools.

For this study, the term "incorrect ideas" is used when discussing students' conceptions that are different from scientifically accepted notions. The original terms (misconceptions or difficulties) are used in instances addressing the original literature.

Detection of Incorrect Ideas

There is no single method for detecting incorrect ideas. Depending on the population, the nature of topics, and the discipline, many different methods can be employed. Open-ended questions (Eisen & Stavy, 1988), two-tier diagnostic tests (Haslam & Treagust, 1987), concept mapping (Hazel & Prosser, 1994), prediction-observation-explanation (Liew & Treagust, 1995), and interviews about instances and events (Osborne & Cosgrove, 1983) allow one to detect a broad range of misconceptions in students of various levels.

Incorrect Ideas in Organic Chemistry

A large body of research has been performed to identify college students' conceptions in the area of general chemistry and introductory chemistry courses. Several detailed reviews are available (Barke et al., 2008; Taber, 2002) and cover more than two hundred misconceptions; however, studies of misconceptions in chemistry classes beyond general chemistry are scarce. A limited number of studies have reported misconceptions in organic chemistry, which is one of the chemistry courses with the largest student enrollment.

Organic chemistry students often possess misconceptions of general chemistry topics such as structure-properties relationships (Cooper, Corley, & Underwood, 2013), Lewis structures (Cooper, Grove, Underwood, & Klymkowsky, 2010), and hydrogen bonding (Henderleiter, Smart, Anderson, & Elian, 2001; Taagepera & Noori, 2000). Several topics from organic chemistry have been reported as being the most challenging to students (Duis, 2011), which include reaction mechanisms, acid-base chemistry, synthesis, and stereochemistry. The most explored misconceptions of organic chemistry

are organic acid strength (Bhattacharyya, 2006; Cartrette & Mayo, 2011; McClary & Talanquer, 2011) and organic reaction mechanisms (Bhattacharyya & Bodner, 2005; Grove, Cooper, & Cox, 2012; Grove, Cooper, & Rush, 2012). Students often rely on their eidetic memory when they try to reproduce organic mechanisms, which does not promote deep learning, especially for students with a limited working memory.

A meta analysis conducted by Bhattacharyya (2014) revealed two of the most prevalent students' misconceptions in a body of research literature conducted with organic chemistry students. According to this meta analysis, students across different studies thought that organic reactions always yield the lowest energy product. Another problem often encountered by students in the aggregated sample was a failure to recognize multiple reaction sites.

A study by Schmidt (1992) of German high school students reported conceptual difficulties with isomerism. Students were inclined to restrict the concept of isomerism to compounds within the same class. Students also believed that isomeric molecules must have branched structures. A potential explanation for these beliefs lies in the discrepancy between the scientific definition of isomerism and instructional practices that overuse iso-nomenclature.

Chemistry misconceptions do not disappear after students complete an organic chemistry course. A study (Rushton et al., 2008) done with fourth-year chemistry students reported a wide range of alternative conceptions that students hold regarding reactivity and stability of organic compounds, as well as an understanding of "curved arrow" notation in reaction mechanisms. Even chemistry graduate students have misconceptions regarding foundational ideas in organic chemistry (Bhattacharyya, 2006).

Turkish authors reported misconceptions held by prospective science teachers about alkenes (Şendur, 2012) and by university and high school students about aromaticity (Topal, Oral, & Özden, 2007). A study by Taagepera & Noori (2000) revealed common misconceptions among organic chemistry students. Some misconceptions relate to general chemistry topics, for example, beliefs that bond polarities depend solely on absolute electronegativity of atoms, regardless of whether they are connected or not.

Practitioners (Mdachi, 2012; Zoller, 1990) have reflected on their own classroom experience with teaching stereochemistry. According to their empirical observations, students often have difficulties identifying a plane of symmetry. Some stereochemistry misconceptions have been uncovered in two chemistry education research dissertations (Krylova, 1997; Lyon, 1999). While the primary purpose of these dissertations was not to uncover misconceptions, the dissertations contain the interview transcripts from which can be concluded that students in these studies have difficulties with a visualization of organic molecules and *R/S* nomenclature rules. A textbook analysis by Kumi et al., (2013) suggests that the origin of students' difficulties with visualization of organic molecules, especially with translation from 2D to 3D representations, are potentially reinforced by the diagrams presented in organic chemistry textbooks. Several studies reiterate students' difficulties with representations of organic molecules (Koutalas, Antonoglou, Charistos, & Sigalas, 2014; Olimpo, Kumi, Wroblewski, & Dixon, 2015) emphasizing preference to static images and students' inflexibility to perform rotational tasks. A comprehensive review of research on students' learning (Graulich, 2015) emphasized the lack of generalizable findings and an insufficient amount of research on students' understanding of organic chemistry.

Visualization in Chemistry

From van't Hoff's early tetrahedral models to modern-day protein imaging, visualization plays an important role in chemistry. To understand ideas of chemistry, such as the shape of molecules, reactivity, intermolecular forces, and polarity, chemists must understand and use a variety of atomic and molecular representations. Students also need to construct and mentally manipulate three-dimensional images from these drawings (Pribyl & Bodner, 1987).

Instructors and textbooks heavily convey concepts to chemistry students using different types of visualization. Organic textbooks contain many representations of different types such as skeletal representations, geometric isomers, enantiomers, etc. Visualization is extremely important in understanding a reaction mechanism. Students need to be able to interconvert between different aspects of two- and three-dimensional representations to understand the content (Wu & Shah, 2004).

Stereochemistry as a Scientific Discipline

Stereochemistry is a sub-discipline of chemistry that refers to the study of molecules in three dimensions. Since most molecules are three-dimensional, stereochemistry is foundational for most chemistry knowledge. An important branch of stereochemistry is the study of chiral molecules. A chiral molecule is a type of molecule that has a non-superimposable mirror image. Eliel, Wilen, and Mander (1994) factored stereochemistry into static and dynamic domains. Static stereochemistry deals with the stereochemistry of molecules, as well as with their energy, spatial arrangement, physical properties, and most of their spectral properties. Dynamic stereochemistry deals with

stereochemistry of reactions, as well as with their stereochemistry requirements and outcomes.

Historically, stereochemistry originates from the discovery of plane-polarized light by the French physicist Malus in 1809 (Eliel et al., 1994). In 1812, Biot, another French scientist, discovered rotation of light by quartz plates. Later he discovered that the phenomenon of optical rotation extends to organic molecules such as sucrose, camphor, and tartaric acid. In 1848, Pasteur separated crystals of the sodium ammonium salts of tartaric acid by means of a pair of tweezers. Twelve years later, in 1860, Pasteur arrived at a molecular explanation of the phenomenon of optical rotation and postulated that molecules of the salts of tartaric acid in the crystals that he managed to separate are mirror images of each other. In 1874, Le Bel and van’t Hoff independently arrived at the idea of a tetrahedral arrangement around a carbon atom. To describe the relationship of molecules that are mirror images, Lord Kelvin introduced the term “chirality” in 1894.

Stereochemistry evolved extensively in the 20th century. Many scientists have contributed to the development of stereochemistry as a scientific discipline. Some of them were awarded Nobel prizes: Hassel and Barton in 1969 for the development of conformation analysis; Cornforth and Prelog in 1975 for their work on the stereochemistry of organic molecules and enzyme-catalyzed reactions; Knowles, Noyori, and Sharpless in 2001 for their work on asymmetric catalysis. Stereochemistry has played a preeminent role in the development of many drugs, including birth control pills and the tragic example of thalidomide. Since most drugs are required to be enantiomerically pure, this need stimulated the development of asymmetric catalysis.

Stereochemistry is central to the understanding of many chemistry concepts. For example, knowledge of configuration and structure is required to describe the architecture of a complex molecule; molecules of identical composition but different configuration often have vastly different biological functions. In addition, differences in reactivity can result from differences in stereochemistry as well as differences in functionality.

Concept Inventories in the Science Classroom

A typical college student in the U.S., whose degree program requires organic chemistry, enrolls in an organic chemistry class in the second year of attending college. Some students have received formal instruction about functional groups prior to beginning an organic chemistry course. A general chemistry course, which is usually a prerequisite for organic chemistry, also contains information about molecular geometry of simple molecules such as H_2O, NH_3, and CH_4. Thus, most organic chemistry students' prior knowledge as well as their beliefs about chemistry are formed in introductory classes (Duis, 2011).

Few research studies exist in the literature that have explored conceptions that undergraduate organic chemistry students have about stereochemistry (Krylova, 1997). Previous studies regarding incorrect ideas in stereochemistry have reported findings from think-aloud protocols from a small sample of students (Krylova, 1997).. The small sample size in these studies impedes generalization of the findings to larger populations. The development of a diagnostic tool that can be administered to a larger sample produces data that can be generalized.

A concept inventory (CI) is a test designed to measure if a student has an accurate knowledge of a specific concept or several concepts. CIs usually consist of multiple-

choice questions, but there are rare examples where the instrument is composed of questions with several correct answers and open-ended questions. Several questions can cover one concept. Questions for CIs are developed based on students' incorrect ideas. Administering a CI and scoring the results provide information about the presence and abundance of certain incorrect ideas in a student population. Instructors may use concept inventories as a way of evaluating their own effectiveness as an instructor and diagnosing common student problems.

Concept inventories differ from other types of assessment used by instructors. Exams, created by instructors who teach a course, often include a wider variety of content with fewer items testing a single concept. Distracters may be based upon the instructors' experience with some incorrect ideas prevalent in a given population but rarely utilize research that have identified alternative conceptions. Unlike the concept inventory developers, instructors who compose final exams also typically do not assess them rigorously to ensure clarity.

Concept inventories typically do not attempt to measure any affective variables, such as interest, motivation, or attitudes. However, tools to assess components of the affective domain with regard to chemistry knowledge are available. For example, attitudes towards chemistry can be measured using a semantic differential scale (Bauer, 2005).

Concept inventories exist for a wide variety of topics and are in great demand for both educational researchers and practitioners. Examples of different concept inventories from various fields in science are presented in Table 2.1. There are also diagnostic tests aimed to assess students' knowledge when they are entering or leaving a certain course.

Even though they are often called concept inventories, they are not testing a specific narrow concept, but are aimed more to test the set of concepts taught in a specific course, for example the Statistics Concept Inventory (Stone et al., 2003).

Table 2.1.

Examples of Concept Inventories Grouped by Fields.

Field	Examples of Concept Inventories	References
Biology	Concept Inventory of Natural Selection	(Anderson, Fisher, & Norman, 2002)
	Osmosis and Diffusion Concept Inventory	(Fisher, Williams, & Lineback, 2011)
	Genetics Concept Assessment	(Smith, Wood, & Knight, 2008)
	Genetics Literacy Assessment Instrument	(Bowling et al., 2008)
	Host-Pathogen Interactions	(Marbach-Ad et al., 2010)
Physics and Astronomy	Star Properties Concept Inventory	(Bailey, Johnson, Prather, & Slater, 2012)
	Force Concept Inventory	(Hestenes, Wells, & Swackhamer, 1992)
	Statics Concept Inventory	(Steif & Dantzler, 2005)
	Lunar Phases Concept Inventory	(Lindell & Olsen, 2002)
Chemistry	Test to Identify Student Conceptualization	(Voska & Heikkinen, 2000)
	Mole Concept Test	(Krishnan & Howe, 1994)
	Chemistry Concepts Inventory	(Mulford & Robinson, 2002)
	ACID I Concept Inventory	(McClary & Bretz, 2012)
	Enzyme-Substrate Interactions Concept Inventory	(Bretz & Linenberger, 2012)
	Chemistry Concept Reasoning Test	(Cloonan & Hutchinson, 2011)
	Understanding of Acid and Bases	(Cetin-Dindar & Geban, 2011)

Concept inventories are subject to extensive research and multiple validation trials. Development of a concept inventory often starts as a dissertation project or as an initiative of a group of researchers and practitioners from multiple institutions. Currently, the development process of about 25 concept inventories is adequately described with a

sufficient level of details in the science education research literature. Some concept inventories are either still under development or the studies reporting on them are lacking details.

Treagust (1988) in his seminal paper described a developmental process for diagnostic tests, which has been implemented to create a number of assessment instruments. He presented three broad areas required for creating these tests: “defining the content,” “obtaining information about students’ misconceptions,” and “developing a diagnostic test.” These areas can be subdivided into ten individual steps:

1. Identifying propositional knowledge statements.
2. Developing a concept map.
3. Relating propositional knowledge to a concept map.
4. Validating the content.
5. Examining related literature.
6. Conducting unstructured student interviews.
7. Developing multiple-choice content items with free response.
8. Developing the two-tier diagnostic tests.
9. Designing a specification grid.
10. Continuing refinements.

Many researchers who have developed concept inventories in different fields modified or even eliminated steps proposed by Treagust. A review by Lindell, Peak, and Foster (2007) summarized the design and validation methodologies for 12 concept inventories in the field of physics and astronomy. The authors report a lack of consistent methodology, differences in detecting the concept domain of interest, and discrepancies in detecting reliability and validity. Published reviews of chemistry assessment tools (Barbera & VandenPlas, 2011; Bretz, 2014) highlight and elaborate on discrepancies that occur in the design process of assessment tools. To summarize these reviews, it can be

said that no two instruments are alike, neither in the development process nor in the implication for practice.

In general, the testing process follows an iterative design in which a version of the instrument is created, administered to a sample of participants, and analyzed, with revisions made based upon the outcome of the analysis. The subsequent version is then administered to another sample of participants and the process reiterates.

Several examples of concept inventories from various fields are presented in the following sections along with information about the development process, test format, statistics reported by authors, and evidences of validity. Development processes, described below, vary significantly.

Concept Inventory of Natural Selection

Anderson et al. (2002) presented a diagnostic test to assess students' understanding of natural selection. The alternative conceptions about natural selection were identified by open-ended questions, from two research articles, and by interviewing seven upper-division ecology majors. The body of research literature on natural selection was used to identify the content (five facts about natural selection and their inferences) to be tested. The test was piloted, and some of its content was revised, and the test was administered again. The test, which contained pictures of finches, guppies, and lizards, consisted of 20 questions with four possible choices. The final version of the concept inventory was tested with 206 non-major students enrolled in general biology. Authors reported item difficulty ranges within acceptable limits (14.5-80.6%, average 46.4%; three items were below suggested levels) and discrimination as point biserial values (.16–.52; six items were below the desired limit). Principal component analysis was

performed to measure internal validity of the final version of the instrument. The Kuder-Richardson (KR) formula 20 coefficient was found to be .58 for one course section and .64 for the other section.

Assessment interviews were conducted with seven students who took the test in order to establish face validity. Scores from interviews and tests were within 15% difference. Because the test was rather lengthy, the authors attempted to establish its readability with 23 community college students. About every seventh word was deleted from the test (the careful selection of the deleted words was done in order not to distort the content) and students were asked to fill in the gaps. Students scored about 50% on this task, which claimed to be appropriate for the target audience.

Digital Logic Concept Inventory

Herman, Zilles, and Loui (2014) reported the development of a Digital Logic Concept Inventory. The authors used digital logic misconceptions from two literature sources. A survey of instructors using Delphi consensus rating was employed to identify the most important and difficult topics. In the alpha testing of the instrument, the researchers gathered additional misconceptions by administering a version of the instrument with an option for an open response. Follow-up interviews with 11 students were conducted using think-aloud protocol. The final version of the instrument consisted of 19 items, covering topics from four main categories. The majority of items had four alternatives, some had five, and a few had six. Students were given 25 minutes to complete the test.

Authors reported the reliability that was estimated by KR-21 formula as .505 for students in computer science class ($N = 92$) and .639 ($N = 111$) for students in computer

engineering class; however, the mean differences were not significant (8.2 and 8.3, respectively). Minimum and maximum scores were reported to ensure full scale coverage. A chi-square test was performed to ensure the absence of bias on the individual item level.

An attempt to establish expert content validity was made. Experts were asked to answer each item, reflect on whether it reflects the concept students should know after completing the course, and rate the quality of the item. Experts were also asked to express their opinions on the Digital Logic Concept Inventory as a whole. Experts were asked about content relevancy and coverage. The student sample was divided into quintiles and response patterns were examined for individual items. The majority of experts agreed that the inventory reflects core conceptual knowledge and can be widely adopted as a formative assessment.

Enzyme-Substrate Interaction Concept Inventory

A brief article was published (Bretz & Linenberger, 2012) about the development of an Enzyme-Substrate Interaction Concept Inventory; however, the information about the preliminary stages of the process of development was limited**.** A doctorial dissertation of one of the authors described the process in detail (Linenberger, 2011).

Twenty-five undergraduate and graduate students from biochemistry courses were interviewed to reveal their misconceptions about enzyme-substrate interactions. Two models of representation were used: “induced-fit” and “lock-and-key.” The results of the interviews (concept maps and word frequency analysis) were used to create a pilot version of the instrument that was given to 108 students enrolled in a biochemistry course. A subsample of 10 students was interviewed using a think-aloud protocol. The

test was revised based on interview outcomes and item analysis. According to Linenberger (2011), an expert panel was involved in the development process but had not reached an agreement on correct answers to some of the questions. Some concerns regarding content, nature of the concepts, and their generalizability were raised as well.

The final version of the instrument consisted of 15 multiple–choice items with four options for each item; eight questions referred to a pictorial representation of enzyme-substrate interactions. The final version of the instrument was administered to 788 biochemistry students from 16 institutions. A sample of 707 students who answered all questions was analyzed. Demographics and the academic majors of the participants were reported. The mean of 8.32 and Ferguson's δ (a measure of the breadth which is the distribution of the test scores compared to the possible range of the scores) of .949 were reported. The distribution of the scores produced by this inventory was not normal, which forced the authors to use non-parametric statistics. A reliability coefficient of .53 measured by Cronbach's α indicated that most likely items were not measuring the same construct. The authors claim that students often had disconnected ideas and incorrectly related concepts. Their claim was supported by the interview data. Misconceptions and corresponding percentages of students possessing them as measured by distractor analysis from the inventory were reported.

The authors made an attempt to establish concurrent validity. A Kruskal-Wallis test showed a significant difference between performance of the non-major students and students majoring in chemistry, biochemistry, or molecular biology.

Mole Concept Test

Krishnan and Howe (1994) used a four-stage process to develop the Mole Concept Test. In the first stage, they constructed a mole concept map using current textbooks and previously published research. Science educators and chemistry instructors validated the definitions of the topics that composed the concept map. In the second stage, the concepts were redefined as learning objectives. Five learning objectives were formulated. In the third stage, a list of student misconceptions was constructed based on five previously published research studies. Experienced chemistry instructors were asked to confirm whether the misconceptions that were obtained from the literature are true misconceptions. The last stage involved development of the test items based on students' misconceptions. A final version of the instrument, along with a table of distribution of the test items according to the learning objectives, was presented to an expert committee for clarification purposes. The final version of the instrument consisted of 20 questions of various types. Four types of test items were developed: three simple multiple-choice questions, two two-tier true-false questions with reasons, five two-tier multiple-choice items, and 10 open-ended questions.

The Mole Concept Test was given to a sample of 20 sophomore chemistry majors. Authors reported a quite high Kuder-Richardson formula 20 reliability coefficient of .81, acceptable ranges of difficulty indices (.4–.8), discrimination indices (.46–.86), and biserial coefficients (.22–.68). The authors mentioned the potency of using this test for assessment in different courses.

Test to Identify Student Conceptualization

Voska and Heikkinen (2000) developed the Test to Identify Student Conceptualization (TISC). The test covered applications of Le Châtelier's principle. The test consisted of 10 items each with four possible choices and was two-tiered. Every question was followed by the second tier open-response question in which students were asked to provide a reason for their response.

Common student misconceptions about chemical equilibria were taken from five literature sources. Two chemistry educators identified concepts and terms that characterize chemical equilibrium and validated that propositional knowledge is congruent with selected chemistry concepts.

The test was administered to 95 second-semester general chemistry students. The Kuder-Richardson formula 20 reliability coefficient was found to be .79. Factor analysis yielded a two-factor structure. Probabilities that the test correctly identifies answers and reasons expressed in the students' interviews were calculated. The prevalence of 11 misconceptions was established with the TISC.

Three chemistry professors examined the content of the TISC. They agreed that each item was matched to one of the three chemical equilibrium topics. Construct validity was established by comparison of students' responses from the interviews and scores from the test.

The Relativity Concept Inventory

Aslanides and Savage (2013) created a list of concepts that are taught in an introductory relativity course. Expert feedback was obtained from 30 international participants by the means of an online survey. Only concepts that received more than

75% agreement were used to develop the 24 multiple-choice questions. Expert feedback, instructors' interviews, and think-aloud protocols with students solving the questions were used to continuously improve the draft Relativity Concept Inventory (RCI). The RCI consisted of 24 one-tier items each with 2-4 possible responses (including true/false). The RCI was given as a pre- and post-test. The RCI was given along with a 5-point confidence scale measuring how certain the student was when answering the specific question.

Item difficulty indices (mean of .71) and discrimination indices (mean of .24) were reported and laid within a desired range. The Kuder-Richardson formula 20 coefficient was found to be .74. Correlations between students' responses to various questions from the inventory were also reported. Due to the small sample size ($N = 53$), the authors used a Monte-Carlo simulation method. Item response theory and factor analysis were also used to examine functioning of the items and the relationship between them. Gender was found to be a significant predictor for four individual questions from the RCI.

CHAPTER III

METHODOLOGY

Research Design

A mixed method approach was used for the development of the Stereochemistry Concept Inventory. An exploratory sequential design, which is a type of design in which qualitative and quantitative data are collected and analyzed separately but used to inform the individual steps, was used. The first phase of the study was a qualitative exploration of the students' incorrect ideas, for which interview data were collected from students enrolled in an organic chemistry course. The qualitative phase was necessary because very few published studies exist on incorrect ideas in stereochemistry. Also in the first phase, data about which topics organic chemistry instructors cover in their instruction were collected by means of a national survey. The data in the first phase were used to construct the pilot version of the Stereochemistry Concept Inventory. The second, quantitative phase followed the qualitative phase for the purpose of refining of the Stereochemistry Concept Inventory. In the quantitative phase, data were collected from students enrolled in an undergraduate organic chemistry course. Psychometric analysis was used to assess the quality of items, and response process validity study was used for item refinement. Data from the response process validity study and psychometric analysis

were used help to refine the instrument in an iterative manner. A flowchart of the study is presented in Figure 3.1.

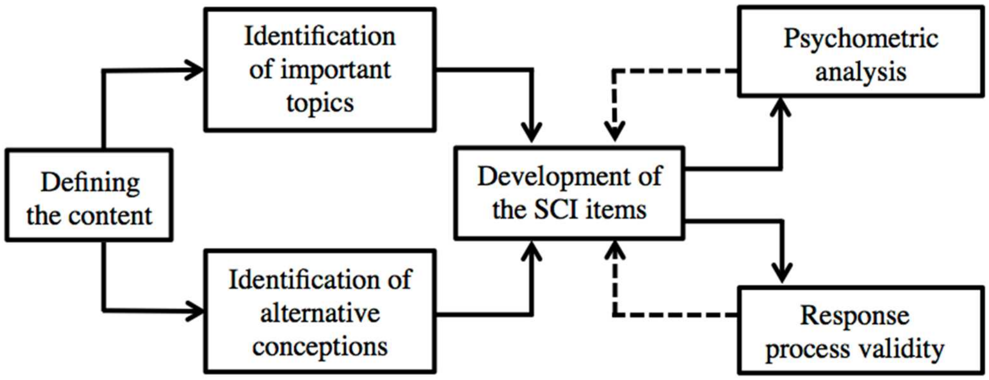


Figure 3.1. Flowchart of the study.

Theoretical Framework

This study was placed in the framework of constructivism. A famous quote from David Ausubel states: “The most important single factor influencing learning is what the learner already knows” (Ausubel, Novak, & Hanesian, 1978). Students in chemistry classes build new concepts based on their existing knowledge, beliefs, and expectations (Bodner, 1986).

Much of the research involving incorrect ideas has a theoretical basis in the work of David Ausubel, who developed a theory of verbal learning and “advanced organizers” that ultimately led to his assimilative theory of cognitive learning (Ausubel, 1960). Ausubel developed a theory of reception learning. Reception learning occurs when the content of the learning task is presented to the learner rather than independently discovered by the learner. Ausubel claimed that concepts are continuously brought into the mind’s conceptual framework and integrated with older conceptual structures. He made a distinction between rote learning and meaningful reception learning. In rote reception learning, knowledge is not associated with previously learned concepts. Rotely

learned knowledge is isolated and thus is more likely to be lost because it is not integrated into existing cognitive structures. Meaningful reception learning occurs when new knowledge and new concepts are associated with ideas or concepts already in the learner's cognitive structure. Rote-reception learning and meaningful learning are not dichotomously different, but they are rather on a continuum. Ausubel studied brief instructional episodes that were intended to organize relevant concepts together so that they could be easily assimilated into the student's existing cognitive structure. New knowledge is accepted when it is associated with a general subsuming concept that is already present.

Cognitive constructivism (Bodner, 1986) derived from the work of Jean Piaget, a Swiss developmental psychologist. His theory of cognitive development proposed that students could not be "given" information, which they immediately understood and used. In other words, direct transfer of information is not a plausible root that leads to learning. Instead, students have to "construct" their own knowledge. They build their knowledge through experience and using existing mental structures. These experiences enable them to create mental schemes. These schemes are modified, enlarged, and become more interconnected as instruction progresses.

One of the most fruitful outcomes of Ausubel's theory is concept mapping, invented by Joseph Novak in the 1970s during his tenure at Cornell University. Concept mapping is a technique for representing knowledge in a pictorial format (Novak, 1990). These knowledge graphs consist of networks of concepts. The networks consist of nodes (points, vertices) and links (edges, arcs). Each node represents a concept, and each link represents a relationship between concepts. In a meta analysis of 55 studies, Nesbit and

Adesope (2006) concluded that compared with activities such as attending lectures or participating in class discussions, concept mapping activities are more effective for knowledge retention and transfer. A concept map of stereochemistry concepts was used in the early phases of this project to ensure that concepts that were included both in the survey and the Stereochemistry Concept Inventory were interconnected.

Ausubel's and Novak's theoretical contributions stimulated the development of classification schemes that bring structure to chemistry curriculum. Alex H. Johnstone of the University of Glasgow proposed the most famous approach called Johnstone's triangle. Johnstone's triangle includes three domains that are representative of chemistry knowledge: symbolic, particulate, and macroscopic (Johnstone, 1993) as illustrated in Figure 3.2. Johnstone's theory provides the means to interconnect different pieces of chemistry knowledge into a coherent framework. The current study proposed a development of a diagnostic tool which was meant to assess student knowledge in all three domains of chemistry knowledge. As a part of an organic chemistry course, stereochemistry involves concepts across all three domains.

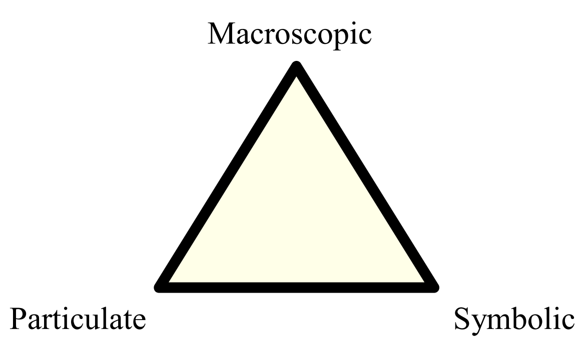


Figure 3.2. The three conceptual levels of chemistry.

Personal Stance

Qualitative research methodologists recommend discussing the study rigor as well as personal stance of the researcher before the study is conducted to address researcher's personal beliefs and biases. The expertise of a primary researcher is also important to ensure that the appropriate level of content knowledge is possessed.

The primary researcher holds a master's degree in Organic Chemistry and received extensive training in the area of stereochemistry as well as in educational research methods during his doctoral studies in the Chemistry Education program at the University of Northern Colorado. During his undergraduate and graduate studies, he took courses in stereochemistry and completed additional coursework in a variety of possible applications of stereochemistry. Courses in Stereochemistry, Organic Chemistry I (aliphatic compounds), II (cyclic compounds), and III (bioorganic chemistry), Advanced Organic Chemistry (both Synthesis and Mechanisms), Spectroscopy, Group Theory and Advanced Inorganic Chemistry were completed with A-grades. His first research paper (Leontjev, Vasiljeva, & Pivnitsky, 2004) was devoted to a stereochemical research topic and described the stereochemistry of reduction of various steroidal ketones with amine-borane complexes. Given all the above, the primary researcher considers stereochemistry a cornerstone of chemistry. This claim is partially supported by the abundance of stereochemistry concepts in the upper-division courses.

In all instances where the primary researcher's content knowledge was limited or not comprehensive enough, a "stereochemistry bible" by Eliel et al. (1994) was consulted. To date, this is the most comprehensive, fully referenced book on the subject of stereochemistry written by world-recognized experts in the field.

The primary researcher taught general and organic chemistry laboratories for eight years in the United States and for two years in Russia. During his laboratory instruction, he always tried to facilitate (occasionally successfully) students in seeking answers to the question “why does this happen?” Years of instructional experience helped him to understand learning progressions of students. These skills are important in the interview process where the greatest emphasis is placed not only on what students know but also on the students’ reasoning behind their answers.

Setting

The Stereochemistry Concept Inventory is primarily intended for students enrolled in organic chemistry classes. At the University of Northern Colorado, stereochemistry is covered in the middle of the first semester of organic chemistry (usually the 7^{th} or 8^{th} week). This topic is usually taught during three lecture periods. During the semester, students are given weekly quizzes. Students are encouraged to engage in solving end-of-chapter problems, but the problems are not graded. Students are graded based on weekly quizzes, midterm and final exams, and web-based homework. The laboratory course is separate from the lecture course. Laboratory activities in the first semester organic chemistry course do not involve concepts of stereochemistry. In the second semester, a lab devoted to a reduction of a carbonyl compound addresses some of the stereochemical topics taught in the first semester.

The participants for the survey were instructors who were teaching organic chemistry at various types of institutions. Due to the different settings of institutions and their missions, instructional processes can vary significantly. Instructors from community

colleges, liberal arts colleges, and research universities were asked to complete a survey. Their demographic information was also collected.

Step 1. Defining the Content

The methodology for the development of concept inventories was outlined in the paper by Treagust (1988). This project followed the methodology proposed by Treagust with some content-specific details. Treagust proposed three broad areas required for creating these tests: defining the content, obtaining information about students misconceptions, and developing a diagnostic test.

Content

To identify relevant content as well as potential sources of incorrect ideas, an analysis of the nine most widely used organic chemistry textbooks (Houseknecht, 2010) was conducted. The analysis focused on two major themes and content. To make the Stereochemistry Concept Inventory appropriate for use in any organic chemistry course, it should cover the content shared among the most used textbooks. Textbooks were assessed for content overlap. Only the content that was presented in all organic chemistry textbooks was used.

Concept Map

A concept map was taken from Solomons' Organic Chemistry textbook (Solomons, Fryhle, & Snyder, 2014) with some modifications. For example, since organic reactions are covered after the stereochemistry topics in the majority of textbooks, all concepts related to reactions were eliminated from the concept map. Some concepts were reformulated to ensure clarity. Three organic chemistry instructors were asked to validate the scientific correctness of the content, as well as whether they covered

this content in their organic chemistry instruction. The goal of having the concept map early in the developmental process was to define the topics that were being tested and to show that they were interconnected.

Propositional Knowledge Statements

A set of propositional knowledge statements (also called learning objectives, learning goals, or instructors' expectations) were developed based on syllabi and textbooks for organic chemistry classes. Textbooks and several syllabi were analyzed for learning goals regarding stereochemistry. Learning goals involving stereochemistry topics from the concept map were considered for future steps. The information about the propositional knowledge statements helped the researcher to choose an appropriate format and content for the questions.

Step 2. Stereochemistry Instruction Survey

A survey (the Stereochemistry Instructional Survey, Appendix B) was developed with the intent of covering the entire spectrum of stereochemistry topics that may be taught in undergraduate organic chemistry as well as learning goals pursued by instructors. Two organic chemistry instructors were asked to validate the survey content. The survey was piloted with 12 organic chemistry instructors from various institutions to ensure clarity and content relevancy.

The survey was then distributed to organic chemistry instructors by email using Qualtrics to obtain information about the content covered in stereochemistry instruction, the complexity of problems used by instructors, instructors' expectations of student

knowledge and skills, and instructors' perceptions of difficulties that students experience regarding stereochemistry concepts.

The participants for the study were selected from three major institution types determined by the highest chemistry degree offered at their institution: *two-year* (Associate's degree), *four-year* (Bachelor's and/or Master's degree), and *doctoral* (Doctoral degree).

The email that instructors received contained the purpose and rationale for the study followed by the link to Qualtrics. If the instructor decided to participate, he/she clicked on the link. The first page contained the consent form. If the instructors were willing to participate, they pressed the "NEXT" button. They had an option to quit at any moment by just closing the window in the browser.

Data Analysis

The results of the survey were analyzed using basic descriptive statistics. For each category (either learning objective or topic of instruction) frequencies were calculated. The percentage of instructors who agreed that a particular topic was important informed the researchers whether they should include this topic in the Stereochemistry Concept Inventory. Topics and learning objectives that received the highest agreement measured as percentage of respondents served as a content basis for the Stereochemistry Concept Inventory.

The demographic data were used to describe the sample of participants. Participants were classified based on the type of institution to which they belonged, the highest degree earned, years of teaching experience, and area of expertise.

Step 3. Identification of Incorrect Ideas

This phase of the study was qualitative and pertained to the detection of incorrect ideas students possess about stereochemistry.

Student Interviews

Qualitative research usually utilizes interviews as a main source of data (Creswell, 1998). Several types of interviews are used in qualitative research. Among these are the informal conversational interview, the interview guide approach, the standardized open-ended interview, and the closed, fixed-response interview (Patton, 1990). These approaches differ in sequence, wording, and context of the questions asked during interviews. In this study, the interview guide approach was used. For this approach, topics to be covered are specified beforehand, but the interviewer selects the sequence and wording of the questions. Data collected in this approach is more systematic than in the informal conversational approach. This approach also allows for reformulating the question if a participant is unfamiliar with terminology. Advantages of this approach are that it allows for flexibility in tailoring questions to particular individuals.

The data collection method for this part of the project consisted of interviews with the students who were enrolled in the organic chemistry course and volunteered to participate. The interview guide approach was used which allowed for slight modification of the order of the questions and their wording. This allowed deeper investigation into some of the students' ideas. When several questions addressing the same idea were asked, the chance that they revealed the student's conceptions was greater.

The interviews consisted of two parts, and each student answered questions from both parts. In the first part, students were asked conceptual questions about

stereochemistry and to provide their reasoning for their answers. In the second part, students were given a set of several problems on stereochemistry. These problems required a short answer, and students were asked to explain why they chose a particular answer. A sample of students enrolled in an organic chemistry course was interviewed. The interview guide (Appendix C) was designed based on the concepts of stereochemistry that are covered in a traditional organic chemistry course.

Think-aloud techniques were used during the data collection. This technique is based on listening to learners, and it is a valuable method in science education because it provides insights into what occurs in the learner's mind. This information is useful to educational researchers because it provides them with better understanding of students' learning strategies and allows practitioners to develop new teaching methods as well as assessment tools to gauge students' understanding (Bowen, 1994).

Quizzes

Another method was used to investigate the abundance of incorrect ideas in larger samples of students. Conceptual open-response questions similar to some interview questions were included in lecture and laboratory quizzes. The responses were analyzed for frequencies after coding them across major themes. The data collected this way provided more support for generalizability of the findings.

Curricula for organic chemistry lecture and laboratory courses are quite different and cover different aspects of organic chemistry. A lecture course typically covers theoretical aspects of organic chemistry, while a laboratory course is devoted to applied aspects of organic chemistry. A laboratory course is also limited to a specific set of experiments that can be performed in the laboratory with equipment that is available in a

particular school setting. By administering quizzes in both settings, we attempted to reduce the effect of a certain setting (lecture or laboratory) on student responses. Two quizzes (one in lecture, one in the laboratory) were administered one in the first semester of organic chemistry, and the second one in the second.

Data Analysis

Interviews were analyzed for the presence of general themes that resemble a particular incorrect idea. Some misconceptions have been addressed in previous research (Krylova, 1997; Lyon, 1999). The interview transcripts and responses to the open-ended questions were analyzed for the presence of incorrect ideas that are also found in the literature. Our coding rubric included the incorrect ideas previously found by other researchers as well as those that emerged from our data.

A general approach called "constant comparative analysis" was used in the data analysis. Originally developed by Glaser and Strauss (1967), this strategy involves taking one piece of data (for example one theme from an interview) and comparing it with all others that have similar themes. This approach helps to understand possible relations between various pieces of data. For example, if the answers in different interviews involved the same stereochemistry topic, these answers were compared with each other for general themes. As a result of this data analysis method, a list of incorrect ideas was generated. Student responses to quizzes were analyzed using a similar approach.

Development of the coding schemes and analysis of the data were conducted by the primary researcher. Two graduate students were asked to comment on emerging themes. A list of student-generated statements representing a certain incorrect idea was given to the graduate students, and they were asked to comment on whether these

statements were representative of a certain incorrect idea. Appendix D contains examples of codes for students' ideas about different biological activity of enantiomers.

Based on the findings from quizzes and interviews, a list of incorrect ideas was produced. Incorrect ideas were checked to verify that they did not involve any material beyond the initial concept map. A limited set of incorrect ideas allowed for the inclusion of distractors that were based on these ideas in multiple test items. Also, selection of prevalent incorrect ideas produces an assessment tool useful for assessment in a variety of settings.

Step 4. Item Development

The Stereochemistry Concept Inventory was composed of items in a multiple-choice format. This format involved the examinees' selecting the best response from a set of options. Multiple-choice test items were developed to address important concepts of stereochemistry established in the previous steps. One response option was composed to address the correct conception. Other response options (also called foils or distractors) were written to address common incorrect ideas. An item-writing taxonomy developed by Haladyna, Downing, and Rodriguez (2002) was used to guide the item writing process. This taxonomy consists of 31 guidelines that are cited in textbooks on educational testing and supported by evidence. Similar guidelines from the same authors are used by the American Chemical Society Exam Institute to write their diagnostic exams. The guidelines are included in Appendix A. In this taxonomy, guidelines 1-8 address content concerns, guidelines 9 and 10 address formatting issues, guidelines 11-13 address style concerns, guideline 14-17 reflect the stem writing, and guidelines 18-31 reflect the construction of choices.

Test Blueprint

Thorndike and Thorndike-Christ (2011) recommended constructing a table of specification (also called test blueprint) prior to constructing a test. A table of specification serves as an explicit plan that should lead a test construction. The basic dimensions of a table of specification are cognitive processes and the description of the content that is covered by the test. These two dimensions are matched to show which process relates to which segment of the content. This approach helps to avoid subjectivity in test construction, for example, overemphasizing one particular topic or lacking coverage for other topics. Using various forms of the test blueprint is recommended to strengthen the evidence for the content-related validity. A blueprint for the pilot version Stereochemistry Concept Inventory is presented in Appendix E, while is the blueprint for the final version, is presented in Chapter V (Table 5.7). The final version of the instrument is given in Appendix F.

Content Validity

Organic chemistry instructors teaching at the college level were asked to participate in the study through solicitation by email. The email addresses were taken from the universities' web-sites. The participants were presented questions from the Stereochemistry Concept Inventory. They were asked to comment on clarity, appropriateness of wording, and scientific content. The participants were also asked to answer the Stereochemistry Concept Inventory items. The survey was administered by means of Qualtrics. The responses from instructors were put into a spreadsheet, and percent agreement was calculated for each of the questions from the Stereochemistry Concept Inventory. The data served as the evidence for content validity.

Step 5. Psychometric Analysis

The items of the Stereochemistry Concept Inventory addressed only concepts included in the concept map and considered important by organic chemistry instructors participating in the national survey. Distracters for the items were written to address incorrect ideas uncovered in Step 3. Multiple distracters from multiple items addressed each alternative conception.

The Stereochemistry Concept Inventory was piloted with students at the University of Northern Colorado. The inventory was piloted with students enrolled in organic chemistry courses and in the upper-level courses such as biochemistry, which use stereochemistry topics as foundational concepts. The students who participated in the pilot test and students who participated in the initial interviews were coming from different cohorts. Students were encouraged not only to mark the option they considered correct, but also to provide comments or to phrase their option in their own words. The main purpose of the pilot test was to assess plausibility of the distractors and difficulty of the questions.

Organic chemistry instructors from other schools were asked to participate in the pilot and final parts of this study. Also, in the electronic survey (Part 1), the participants were asked if they were willing to participate in testing of the instrument, whether by validating its content by experts' opinions or by administering it to students. If they agreed, they provided an email and/or other contact information. Depending on the situation (mainly whether the instructor agreed to sacrifice classroom time for this), the Stereochemistry Concept Inventory was administered with either electronic or paper-and-pencil versions. Stability of the scores was assessed using test-retest reliability

coefficients. If reliability coefficients for both tests were similar, a conclusion about stability of the scores could be made. A second administration of the test happened at the end of the semester.

Item Analysis

Item analysis of the dataset was performed. These included item discrimination, item difficulty, and response pattern which are described below.

Distracter Analysis

The number of options used for each item in the multiple-choice test may differ from item to item. Moreover, the number of answer choices can vary within the same test. For example, Rodriguez (2005) argues that three response options are enough to ensure psychometric quality of test scores. Thorndike and Thorndike-Christ (2011) suggested that distractors should be functioning and effective, meaning that they should be attractive to some students.

Items were analyzed for the frequency of responses selected by participants. The main reason for this analysis was to check for plausibility of distractors. Two types of actions were taken after this analysis, as suggested by Haladyna (2004). If students did not choose the distractor, then it was removed or reworded. If one of the distractors appeared to be more attractive than the correct answer to most students, then this choice may have been indicative of either a confusing item stem or a confusing distracter.

The Stereochemistry Concept Inventory aims to test conceptual understanding, thus distracters were formulated based on students' incorrect ideas. The percentages of students choosing specific distracters are indicative of the prevalence of certain incorrect ideas in the student population. The results of the distracter analysis were summarized in

a list of incorrect ideas with corresponding percentages. Some incorrect ideas were represented by multiple distractors.

Item Difficulty

Item difficulty is the percentage of the population that correctly answered the item. The item difficulty index informs whether an item is too difficult or too easy for the population being tested. Item difficulty ranges from 0 (none of the students answered the item correctly) to 1 (all of the students answered the item correctly). The optimal item difficulty depends on the number of possible distracters and of the question type; there is no consensus on an acceptable range. For example, Kline (2005) recommended a range of .25–.75 for concept assessment. A range of .30–.70 was recommended by Kaplan and Saccuzzo (2012) as best for providing information about differences between students.

To obtain maximum spread of student scores, it is best to use items with moderate difficulties. For the questions with a different number of response options, moderate difficulty can be defined as the point halfway between a perfect score and chance score. However, Lord (1952) recommended that ideal difficulties should be .1 units above the value calculated using the point halfway approach in order to maximize item discrimination ability. The calculations of ideal difficulties for questions with different numbers of responses are presented in Table 3.1.

Table 3.1.

Ideal Difficulty Estimates for Items with Different Numbers of Response Options.

Number of Choices	Chance score	Optimal Difficulty	Ideal Difficulty (Lord, 1952)
2	.50	.75	.85
3	.33	.67	.77
4	.25	.63	.73
5	.20	.60	.70

There are no absolute rules for item selection; however, items with a too low or too high difficulty index do not provide adequate information and should be replaced when possible. Items on the Stereochemistry Concept Inventory were analyzed for difficulty levels. If items fell outside of the desirable limits (.3-.7), they were carefully scrutinized and modified before subsequent testing.

Most of the published concept inventories that reported individual item analysis contained few items with difficulty outside the desired limits. Some studies did not report individual item difficulty indices, but instead reported the average difficulty or the mean of the total score. However, easy items may be beneficial for proper test construction. Commonly, instructors deliberately put a few easy items at the beginning of the test to increase students' confidence so that they continue taking the test. Putting difficult items first may result in lower motivation and elevated anxiety, which may lead to decreased achievement. Conversely, items with low difficulty levels may represent a deep-rooted incorrect idea that is not easily reversed by instruction.

Item Discrimination Index

The discrimination index (D) is computed by subtracting the number of students in the lower group (N_L) who answered the item correctly from the number of students

answered the item correctly in the upper group (N_U) and dividing the difference by the total number in the group (N). The discrimination index can be computed as follows:

$$D = \frac{N_U - N_L}{N}$$

When sample size is large enough (more than 200), analysis is carried out with 27% of the upper and the lower students because these groups are doubtfully different with respect to the trait in question (Kelley, 1939). If the sample size is smaller than 200, then an even split in two groups by 50% (Thorndike & Thorndike-Christ, 2011) is an appropriate procedure.

For the pilot study, we removed or modified items that had discrimination indices lower than .2 on the pilot test. For the final version of the Stereochemistry Concept Inventory, discrimination indices for the distracters were also computed. Negative values for the distracters and positive values for the correct option indicate an adequately functioning item.

Reliability

Reliability was estimated using the Kuder-Richardson formula 21. For paired cases (both pre and post scores are available), a test-retest reliability was estimated.

Response Process Validity

Response process validity evaluates how students understand the question and wording of responses as well as what content knowledge is used to arrive at the selected answer. Information gathered from a response process validity study provides meaningful information for the development and refinement of items.

Retrospective interviews were conducted with the students (*N* = 13) who already completed the Stereochemistry Concept Inventory. Participants were asked why they had selected a certain response option. In addition, participants were asked why certain response options were incorrect. Probing questions were used to identify what knowledge students were using to choose their answers or to eliminate responses.

Students' answers along with their reasoning were reviewed. If there was a mismatch (student possessed correct knowledge but selected an incorrect response or student possessed incorrect knowledge, but selected a correct response), the wording of an item stem or choices were changed.

Human Subjects Protocol

An IRB approval was obtained for each of the phases prior to collecting data. Students' participation in all phases was voluntary. Students had the option not to disclose their responses or interview data if they preferred. Participants who were interviewed for their participation were rewarded with a two-hour group review session by the researcher before their final or midterm exam in organic chemistry.

IRB approvals were obtained from the University of Northern Colorado and other participating schools (either as a result of IRB transfer or submission of a new application). All data collected in this study were through voluntary participation. Some participants received feedback on their performance on the Stereochemistry Concept Inventory (if they requested that) and/or a review session. IRB approvals from the University of Northern Colorado and other schools are given in Appendix G.

CHAPTER IV

ARTICLE 1: INCORRECT IDEAS IN STEREOCHEMISTRY

Abstract

The central goal of this study was to uncover and classify incorrect ideas that students have within the realm of stereochemistry. For that purpose, we conducted a qualitative study where interviews and open-ended questions were employed to elucidate incorrect ideas held by organic chemistry and biochemistry students. A wide range of incorrect ideas including conceptions related to chirality, optical activity, biological activity of enantiomers, various structural representations, and physical properties of stereoisomers were revealed as a result of analysis of the collected data. The incorrect ideas were found to be dependent on the prompt that was presented to the students. For example, students were more likely to explain a difference in biological activity of enantiomers by optical activity when they were presented the information about optical activity, while the students who were presented with the information about absolute configuration of enantiomers were more likely to employ the structural characteristics of molecules when explaining biological activity of enantiomers. Implications for practice and research are discussed.

Introduction

A large body of research has been performed to identify college students' conceptions in the area of general chemistry and introductory chemistry courses. Several detailed reviews are available (Barke et al., 2008; Taber, 2002) and cover more than two hundred misconceptions; however, studies of misconceptions in chemistry classes beyond general chemistry are scarce. A limited number of studies have reported misconceptions in organic chemistry, which is one of the chemistry classes with the largest student enrollment.

For this study, the term "incorrect ideas" is used when discussing students' conceptions that are different from scientifically accepted notions. The original terms (misconceptions or difficulties) are used in instances addressing the original literature.

Organic chemistry students often possess misconceptions of general chemistry topics such as structure-properties relationships (Cooper et al., 2013), Lewis structures (Cooper et al., 2010), and hydrogen bonding (Henderleiter et al., 2001; Taagepera & Noori, 2000). Several topics from organic chemistry have been reported as being the most challenging to students (Duis, 2011), which include reaction mechanisms, acid-base chemistry, synthesis, and stereochemistry. The most explored misconceptions of organic chemistry are organic acid strength (Bhattacharyya, 2006; Cartrette & Mayo, 2011; McClary & Talanquer, 2011) and organic reaction mechanisms (Bhattacharyya & Bodner, 2005; Grove, Cooper, & Cox, 2012; Grove, Cooper, & Rush, 2012). Students often rely on their eidetic memory when they try to reproduce organic mechanisms, which does not promote deep learning, especially for students with a limited working memory.

A meta analysis conducted by Bhattacharyya (2014) revealed two of the most prevalent students' misconceptions in a body of research literature conducted with organic chemistry students. According to this meta analysis, students across different studies considered that organic reactions always yield the lowest energy product. Another problem often encountered by students in the aggregated sample was a failure to recognize multiple reaction sites.

A study (Schmidt, 1992) of German high school students reported conceptual difficulties with isomerism. Students were inclined to restrict the concept of isomerism to compounds within the same class. Students also believed that isomeric molecules must have branched structures. A potential explanation for these beliefs lies in the discrepancy between the scientific definition of isomerism and instructional practices that overuse iso-nomenclature.

Chemistry misconceptions do not disappear after students complete an organic chemistry course. A study (Rushton et al., 2008) done with fourth-year chemistry students reported a wide range of alternative conceptions that students hold regarding reactivity and stability of organic compounds, as well as an understanding of "curved arrow" notation in reaction mechanisms. Even chemistry graduate students have misconceptions regarding foundational ideas in organic chemistry (Bhattacharyya, 2006). Turkish authors reported misconceptions held by prospective science teachers about alkenes (Şendur, 2012) and by university and high school students about aromaticity (Topal et al., 2007). A study by Taagepera and Noori (2000) revealed common misconceptions among organic chemistry students. Some misconceptions relate to

general chemistry topics, for example, beliefs that bond polarities depend solely on absolute electronegativity of atoms, regardless of whether they are connected or not.

Practitioners (Mdachi, 2012; Zoller, 1990) have reflected on their own classroom experience with teaching stereochemistry. According to their empirical observations, students often have difficulties identifying a plane of symmetry. Some stereochemistry misconceptions have been uncovered in two chemistry education research dissertations (Krylova, 1997; Lyon, 1999). While the primary purpose of these dissertations was not to uncover misconceptions, the dissertations contain the interview transcripts from which can be concluded that students in these studies have difficulties with a visualization of organic molecules and *R/S* nomenclature rules. A textbook analysis by Kumi et al., (2013) suggests that the origin of students' difficulties with visualization of organic molecules, especially with translation from 2D to 3D representations, are potentially reinforced by the diagrams presented in organic chemistry textbooks. Several studies reiterate students' difficulties with representations of organic molecules (Koutalas et al., 2014; Olimpo et al., 2015), emphasizing preference to static images and students' inflexibility to perform rotational tasks. However, a comprehensive review of research on students' learning (Graulich, 2015) emphasized the lack of generalizable findings and an insufficient amount of research on students' understanding of organic chemistry.

Theoretical Framework

Educational research of students' incorrect ideas and learning difficulties is rooted in the theory of constructivism. According to constructivism, students' do not simply receive new knowledge; rather, they construct it using pre-existing knowledge (Bodner, 1986). If constructivism can be factorized into cognitive constructivism and social constructivism, this study is closer to Piaget's cognitive constructivism as opposed to Vygotsky's social constructivism. Chemical knowledge cannot simply be transferred from the minds of instructors to students. On the contrary, students are actively engaged in construction of new knowledge. Instruction and prior knowledge contribute to the process of forming new knowledge. When content of instruction contradicts the existing knowledge base, concepts (either new or previous) must be changed. Bodner (1986) emphasized that knowledge must function accordingly within the context when this knowledge is introduced. Research on students' understanding of science often reveals that students possess many incorrect ideas about scientific phenomena. The theoretical contribution of the Ausubel's meaningful learning theory (Ausubel et al., 1978) stimulated the development of classification schemes that bring structure to chemistry curricula.

Construction of knowledge may involve formation of both ideas that are correct and incorrect. Not all incorrect ideas are the same. One of the classification systems of students' knowledge in chemistry was proposed by Zoller (1990, 1996) and consists of three distinct categories: "misconceptions", "misunderstandings", and "no conceptions". Zoller (1996) provides several examples of misconceptions. An idea that phosphoric acid H_3PO_4 is a base because it contains hydroxyl groups is a misconception. Switching the

direction in which a reaction proceeds is an example of misunderstanding. An example of a "no conceptions" case is when a student does not apply concepts of aromaticity or stabilization in an attempt to explain why mononitration of 1-naphthol results in the formation of 2-nitro-1-naphthol or 4-nitro-1-naphtol which involves drawing resonance structures. To a certain extent, a classification scheme proposed and implemented by Zoller (1990, 1996) resembles the SOLO (Structure of Observed Learning Outcomes) taxonomy (Biggs & Collis, 1982); however, Zoller's scheme is more specific to chemistry and more applicable for the ideas that students develop in chemistry classes. The SOLO taxonomy describes the levels of increasing complexity in a student's understanding of a topic and includes five stages: pre-structural, unistructural, multistructural, relational, and extended abstract. Stages are characterized by the number and meaningfulness of the connections that a student makes.

An alternative approach was proposed by Talanquer (2006). This approach is based on the hypothesis that the conceptual difficulties of most science learners result from reasoning based on "common sense." Learners who follow "common sense" tend to generate explanations of natural phenomena based on intuition and broad generalizations, unconsciously and often erroneously applying reasoning patterns. The commonsense chemistry explanatory framework relies on a set of five assumptions about the characteristics of phenomena in the natural world: *continuity*, *substantialism*, *essentialism*, *mechanical causality*, and *teleology*. *Continuity* refers to beliefs that matter can be continuously divided into smaller pieces which have the same qualitative properties as the macroscopic object. An example of a misconception based on this belief can be found in the statement that copper atoms have a red color and expand when

heated. *Substantialism* is an attribution of properties of material substances to abstract concepts. For example, students often consider that heat behaves like a liquid. Here, they substantiate an abstract idea of heat as a thermal energy with properties of a liquid which can be seen in everyday life. *Essentialism* refers to a belief that objects and materials have an essential set of properties that remain unchanged. The idea that rust is a specific type of iron is an example of essential thinking. *Mechanical causality* is an attribution of any changes that happen to a system with external factors. A belief that any chemical reaction happens because of an added reagent illustrates mechanical causality. *Teleology* refers to the belief that subjects and processes behave in order to satisfy a certain need. For example, students often believe that a reaction proceeds to the minimum energy level.

Understanding chemistry, which involves the formation of ideas, is not uniform, but rather can be subdivided to three domains (Johnstone, 1982, 1991, 1993): symbolic, particulate or submicroscopic, and macroscopic. When students construct chemistry knowledge, they need to interconnect all three domains. These domains and their interconnections are often presented as the Johnstone's triangle in chemistry education research literature. Johnstone's theory provides the means to interconnect different pieces of chemistry knowledge into a coherent framework. In this study, we have attempted to uncover students' incorrect ideas across these three domains.

Stereochemistry as a part of an organic chemistry course involves concepts across particulate, symbolic, and macroscopic domains. Designations of absolute configurations of stereogenic centers as *R* and *S* represent symbolic domain, as well as designation of a dextrorotatory compound as (+) and levorotatory compound as (–). The particulate

domain in stereochemistry is represented by various structural representations: wedge-dash projections, Newman and Fischer notations, ball-and-stick models, etc. Wedge-dash projections and projections that are showing only wedge or dash are the most commonly used in modern scientific literature. Newman and Fischer projections are mostly used in instructional materials for organic chemistry classes. The macroscopic domain is also represented in stereochemistry, however to a lesser extent. Most organic chemistry textbooks present Pasteur's resolution of enantiomers with tweezers. Salts of tartaric acid that he studied form crystals unique to each enantiomer. Crystals are different in their appearance and thus can be visually distinguished. However, it represents an exception rather than a common rule. Most racemic mixtures form only one phase. Another example of the macroscopic representation would be the plane of polarized light being rotated by solutions of chiral compounds. This can be easily observed with a polarimeter. To a certain extent, different biological properties constitute the macroscopic domain. Two compounds, (+)- and (–)-carvon, while being mirror images of each other, have different smells. (+)-Carvon smells like peppermint and (–)-carvon smells like caraway seeds. A summary of various levels of representations as applied to stereochemistry concepts is presented in Table 4.1.

Table 4.1.

The Three Conceptual Levels of the Johnstone's Triangle and Corresponding Concepts from Stereochemistry.

Macroscopic	**Particulate**	**Symbolic**
- Chirality of crystals - Rotation of plane-polarized light - Biological activity (e.g. different organoleptic properties for carvon enantiomers)	- Wedge-dash, sawhorse, Newman, and Fischer projections - Ball-and stick-models	- *R/S* configuration descriptors for stereogenic centers - (+)/(–) notation for dextrorotatory and levorotatory compounds

Purpose of the Study

There is a significant body of research available that uncovers students' incorrect ideas and difficulties that relate to organic chemistry mechanisms and visualization of organic molecules. However, to our knowledge, there are no attempts to systematize and classify incorrect ideas that relate to stereochemistry, which was identified (Duis, 2011) as one of the most challenging and foundational concepts of organic chemistry. Failure to master stereochemistry concepts can be a serious barrier to mastering organic chemistry, because most organic chemistry utilizes stereochemistry knowledge as foundational. However, research on students' incorrect ideas in stereochemistry is limited. Both practitioners and researchers must be familiar with students' incorrect ideas to avoid potential pitfalls in teaching. The following research question, which will be addressed, guided this study:

Q2 What incorrect ideas do organic chemistry students hold regarding stereochemistry?

We deliberately use the term "incorrect ideas" when we describe the findings of this study, because it encompasses a broader range of beliefs. The terms "misconceptions" (Cheung et al., 2009), "alternative conceptions" (McClary & Bretz, 2012), "alternate ideas" (Mulford & Robinson, 2002), "naïve ideas" (Stavy, 1990), and "incorrect ideas" (Villafañe et al., 2011) are often used interchangeably when describing the phenomena of students possessing knowledge that is different from scientifically accepted notions. The debates concerning the use of certain terms began more than 25 years ago (Abimbola, 1988) and are still in progress.

Our major intent in this study was to uncover problematic stereochemistry concepts of students, rather than estimate percentages of students who possess them. We

attempted to “saturate” themes with representative quotes from the participants representing multiple dimensions of the phenomenon.

Participants and Setting

This study was conducted at a coeducational public institution in the Rocky Mountain region of the United States. The university has an enrollment of approximately 10,000 undergraduates. The population of the institution consisted of 62% female, 38% male, and 15% minority students at the time of the study.

Two different professors (neither of the authors) taught two separate sections of the two-semester organic chemistry course. The instructors had 17 and 22 years of teaching experience. The textbook used in the course was *Organic Chemistry* by Carey and Giuliano (2010). The class met four times per week for 50 min lecture periods. The students were given quizzes on a weekly basis and two exams in one section and four exams in the another section. The final exam used in both classes was the American Chemical Society Organic Chemistry Exam (OR08 version). Several teaching assistants who were at the time pursuing either a master’s degree in Chemistry or a doctoral degree in Chemical Education taught the laboratory sections. At the time of the study, the first author (AL) taught one of the seven sections of laboratory. Typically, each laboratory section enrolls 12-18 students.

This study utilized a purposeful sampling technique (Creswell, 1998) in which participants were purposefully selected to inform the phenomena being studied. A variety of students representing different majors, academic background, and gender participated in the study. A total of 172 students participated in this study.

Data Collection Methods

This study utilizes a qualitative approach (Creswell, 1998; Merriam, 2009). Primary data were collected by interviews and by administering open-ended questions on three quizzes in both lecture and lab settings. All interview sessions typically lasted for approximately one hour. They were recorded using an iPhone. All drawings were collected as a part of analysis. A sample of questions asked in the interview is presented in Figures 4.1 and 4.2. We used interviews of two types: an interview with open-ended questions and an interview with closed-ended, multiple-choice questions, in which students were asked to explain their choice. A total of 24 interviews were used in this study both for generation of initial themes and as a source of quotes to illustrate incorrect ideas. A total of 11 students were interviewed with open-ended questions, while 13 students were interviewed with multiple-choice questions. Interviews with open-ended questions occurred in the middle of the second semester of organic chemistry. Interviews with multiple-choice questions occurred at the end of the first semester of organic chemistry. Participants were coming from different cohorts. None of the students who participated in the interviews with open-ended questions were interviewed with multiple-choice questions. Those students who participated in the interviews were offered a review session to prepare them for the final exam as a compensation for their time and effort. In addition, all of the students who participated in the interviews were debriefed after the interview. All incorrect ideas that students expressed pertaining to stereochemistry were addressed, and the students were given correct answers to the questions that were not answered correctly. Students generally highly valued both feedback and the review session.

- What are enantiomers? Could you provide an example of an enantiomer?
- How can you separate two enantiomers?
- Do enantiomers have different melting points? Boiling points? Any other physical properties?
- Do enantiomers have different biological properties? Why?
- How can you separate two diastereomers?
- What is a *meso*-compound? Could you provide an example?

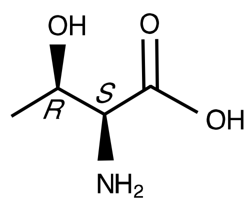
Figure 4.1. Examples of questions that were used in the interviews with open-ended questions.

1. Which of the following substituents has the highest priority according to the Cahn-Ingold-Prelog rules?

–OCH_3, –F, or –Cl

A. –Cl, because it has the highest atomic number
B. –F because it is the most electronegative element
C. –OCH_3 because it has more atoms

2. In which direction does the following compound rotate plane polarized light?



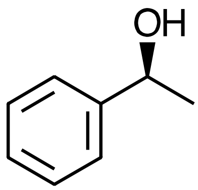
(2*S*, 3*R*)-(–)-2-amino-3-hydroxybutanoic acid

A. Counterclockwise because the α-carbon has an *S* configuration
B. Counterclockwise because it is levorotary
C. It is not optically active because *R* and *S* cancel the rotation of light

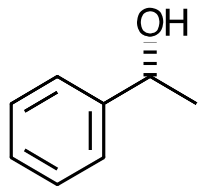
Figure 4.2. Examples of questions that were used in the interviews with multiple-choice questions.

To collect the information from multiple sources and in different settings, several quizzes were administered to students both laboratory and lecture part of instruction to collect additional evidence of their thinking. A sample of the questions asked on the quizzes is presented in Figure 4.3.

Quiz 1 given in the lab setting



I



II

1. Assign the configuration (*R* or *S*) for the chiral center in the compounds **I** and **II**. Label the chiral canters with an arrow and a letter *R* or *S*.

2. According to the Cahn-Ingold-Prelog rules, what is the order of priority of the substituents C_6H_5, CH_3, OH, and H. Arrange them in order of increasing priority.

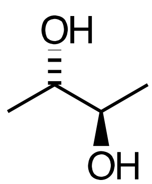
Lowest priority _____ < _____ < _____ < _____ *Highest priority*

3. Do you expect **I** and **II** to have the same boiling point? Why or why not?

4. Can you tell which compound (**I** or **II**) rotates plane of polarized light to the left? Explain your answer.

Quiz 2 given in the lecture setting

Which of the following structure(s) indicate a meso compounds?



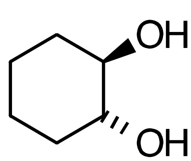


Figure 4.3. Examples of questions that were used in Quiz 1 and Quiz 2.

Quiz 1 was administered in a lab setting as a prelab assignment for the organic chemistry laboratory in the second semester. A total of 15 questions were included on Quiz 1. Quiz questions covered aspects of stereochemistry that were relevant to the

laboratory activity the students were performing. The laboratory exercise that students conducted after the quiz was the reduction of acetophenone with sodium borohydride with subsequent analysis of the mixture of enantiomers using gas chromatography with a chiral column. This was the second laboratory exercise where students analyzed reaction mixtures using a gas chromatograph. A total of 87 students took Quiz 1. Consent forms were received from 86 students, and their responses were used in our data analysis.

Quiz 2 was administered in a lecture setting at the beginning of the second semester of organic chemistry. Two questions on stereochemistry were included in the lecture quiz along with three questions on the reactions of esters provided by the instructors. Instructor-written questions were part of the routine assessment and are not included in this analysis. An example of one of the stereochemistry questions is included in Figure 4.3 (Quiz 2). Consent forms were received from all 29 students who took Quiz 2 in one lecture setting.

Quiz 3 was administered to students enrolled in a biochemistry course during the first class period of the semester. All of the students who took this quiz completed the organic chemistry course with a grade of C or higher, which is a prerequisite for the biochemistry course. Students majoring in chemistry or biochemistry usually take this course, as well as those who are enrolled in a pre-health program. The quiz contained 16 multiple-choice questions; the last question was open-ended and probed students' understanding of different biological activities of enantiomers. A total of 32 students completed Quiz 3, all of which gave their consent. The current study reports an analysis of only the open-ended question. We plan to report an analysis of the multiple-choice

questions which constituted a pilot version of a Stereochemistry Concept Inventory in a separate manuscript.

Creswell (1998) explained the idea of data validation by saying that triangulation is the process of corroborating evidence from different individuals and various types of data such as artifacts collected from students and interviews. A synthesis of evidence culminates in proposing themes as a result of qualitative research. In this study, we used triangulation when collecting information from different sources (interviews with open-ended questions and closed-ended questions, and open-ended questions on the quizzes, as well as field notes). We purposefully administered multiple quizzes in both lecture and lab settings to account for specific aspects of chemical knowledge that may be highly relevant to a lab setting (for example, separation of enantiomers on a chiral column by GC) but less relevant to a lecture setting.

The first author took field notes while observing both chemistry instructors' class periods that were devoted to stereochemistry. A total of nine class periods were observed. This was done to ensure that the content of the questions did not exceed the material covered by the instructors. Field notes that were taken during the lectures served as additional evidence for the proposed themes. These field notes were used to formulate questions and to ensure that the terminology used was in accordance with the material presented in lectures.

Data Analysis

Several themes emerged from the analysis of interview transcripts, field notes, and quiz responses. The themes are centered around certain incorrect ideas that some students possess. Two graduate students (one in chemical education and one in a traditional chemistry area) and three organic chemistry instructors were asked to comment on themes. After iterative refinement using constant comparison of the themes, we examined the collected data to select the most representative examples of quotes for the manuscript and calculated how many times a certain incorrect idea occurred in the interviews or quiz responses. The themes that emerged are presented in the following sections and, when appropriate, the relative occurrence of certain incorrect ideas is presented. During the interview process, students inevitably change their responses as a result of prompt or clarification questions; thus, initial incorrect ideas were counted toward the cumulative count.

We have also used information emerging from the literature to support some of the incorrect ideas found in this study. While the primary purpose of the literature was not to uncover incorrect ideas about stereochemistry, some publications (Krylova, 1997; Lyon, 1999) provided transcripts of interviews that contain some of the students' conceptions about stereochemistry. A general approach called "constant comparative analysis" was used in the data analysis. Originally developed by Glaser and Strauss (1967), this strategy involves taking one piece of data (for example one theme from an interview) and comparing it with all others that have similar themes. This approach helps to understand possible relations between various pieces of data. For example, if the answers in different interviews involved the same stereochemistry topic, these answers

were compared with each other for general themes. As a result of this data analysis method, a list of incorrect ideas was generated. Student responses to quizzes were analyzed using a similar approach.

Results and Discussion

The interview transcripts and responses to the open-ended questions were analyzed for the presence of incorrect ideas that are available in the literature. Seven themes are presented that illustrate student difficulties with stereochemistry topics. The themes are listed below and dimensions of these themes are presented to show continuum of ideas in the categories. Some themes or their dimensions were found in previous research studies, and some are novel.

Ranking Substituents for *R/S* assignment

To assign an absolute configuration, substituents must be ranked first. The conventionally used Cahn-Ingold-Prelog system utilizes a straightforward approach where substituents are ranked based on atomic numbers of the atoms directly attached to the chiral center. However, several students in our study assigned priority of the substituents based on their polarity, bulkiness, "molecular weight" of the *entire* substituent, number of hydrogens present, or any other method that is inconsistent with the Cahn-Ingold-Prelog priority system.

In the initial interviews, many of the students (9 of 24) assigned priority correctly; however, some of them relied on the memorized patterns from their prior experience:

> I remember oxygen has larger priority but I am not exactly sure why. Just whenever we do this oxygen is always number one.

Students also assigned priority of the substituents based on their polarity or electronegativity (7 of 24), bulkiness (2 of 24), “molecular weight” (2 of 24), number of hydrogen atoms present (1 of 24), or number of other atoms in the chain (3 of 24):

> F is small and the most electronegative that is why it has higher priority.
>
> Carboxylic group has a priority [over OCH_3] because it has two oxygens, so it has heavier mass.

It is worth noting that elicited patterns may depend on the examples that are used in the interviews. Often students exhibited multiple patterns of reasoning, blending several incorrect ideas together:

> Oxygen is larger by size than carbon, that is why it takes priority. To the right of carbon is larger, to the left of carbon is smaller. Ethyl is larger than methyl, so it takes priority. Ethyl has more hydrogens than vinyl, so it takes priority. More atoms, larger priority.

or blending correct and incorrect reasoning together:

> O is one and N is two because O is more electronegative and it has a large atomic number.

On Quiz 1, students were asked to rank the order of priority of the substituents **C_6H_5**, **CH_3**, **OH**, and **H**, and in the subsequent question to assign the configuration for the chiral center in the enantiomers of 1-phenyethanol. Ninety one percent (78 of 86 students) of our participants performed the priority assignment correctly, but only 78% (61 of 78) participants who ranked the substituents correctly also assigned the configuration correctly. Eight students did not assign the priority correctly with the most common (6 of 8) mistake being assigning the highest priority to the phenyl group. However, four out of the eight students who assigned the priority incorrectly still assigned the correct configuration to the chiral center in stereoisomers of 1-phenyethanol.

One of the possible explanations for these incorrect ideas might be that students are often asked to do ranking tasks that involve either a concept of polarity or bulkiness, because these are two primary factors that drive organic chemistry reactions. The concepts of polarity and electronegativity are introduced earlier in general chemistry. The concept of size of the substituent or any chemical moiety is less common in general chemistry but quite abundant in organic chemistry. Students may be so accustomed to the task of ranking entities based on polarity or size that they automatically transfer it to the ranking of the substituents for the Cahn-Ingold-Prelog system.

Our findings are consistent with previous investigations in organic chemistry knowledge acquisitions. Krylova (1997) reported that students rank substituents based on their bulkiness, polarity, and number of hydrogen atoms present. Wathen (2008) noted that students sometimes assigned priority by molecular weight of the substituents. However, in some rare cases ranking is done based on mass to distinguish between isotopes of the same element such as hydrogen and deuterium.

With few exceptions, most organic chemistry texts place the Cahn-Ingold-Prelog system of *R/S* nomenclature in the beginning of the stereochemistry chapter. Chamberlain (2012) reflected on a positive experience of postponing teaching *R/S* nomenclature until after the foundations of stereochemistry were covered. Students were taught various methods of how to determine if the molecule is chiral and how to classify pairs of molecules according to their stereochemical relationship in the absence of *R* and *S* designations. According to his observations, students performed better on the problems that ask if molecules are identical, enantiomers, or diastereomers.

R/S Assignment

While students may have no problem with ranking of the substituents according to the Cahn-Ingold-Prelog rules, some have difficulty assigning the configurational descriptors to the molecule. Most of the students expressed the idea that no matter how the molecule is drawn, the hydrogen is always facing behind the plane of paper away from the viewer:

> Hydrogen is always on the back.
>
> You need to keep hydrogen at the back when you are doing CIP rules.
>
> Number four is always pointing backwards.

While most of the students tried to rearrange the molecule moving other substituents, three students from our sample of 24 people who were interviewed assigned the incorrect configuration assuming that the lowest priority substituent is always pointing backwards; however, on the examples they were presented it was either located in the plane of the paper or pointing towards them. Most of these mistakes were made by students working with Fischer projections. Sometimes students admitted their confusion when the substituent with the lowest priority is not facing toward the viewer:

> I know that you put H on a dashed line. H are the lowest priority and they stick at me; that's what throws me off.

Several students reversed the order of the substituents such that the one with the lowest priority became the one with the highest priority. However, this mistake was not consistent within specific interviews (students switched to the correct ranking in other examples), and this was classified either as misunderstanding (Zoller, 1990) or lack of attention.

When students were given a molecule to assign a configuration, one of the students made a statement that hydrogen is always facing toward the back or to the side. Kuo, Jones, Pulos, and Hyslop (2004) found that students generally perform better at assigning correct configuration when presented with a structure in which the substituent with the lowest priority was behind the plane of the paper. We noticed the tendency of our participants to use this strategy as well. Most of the mistakes that were made in *R/S* assignments were associated with incorrect rearrangement of the molecule. Twenty of 24 participants tried to rearrange molecules (often unsuccessfully, eight students tried to switch only two substituents) to place the substituent with the lowest priority on the back, while four students exhibited some level of higher order thinking and used other strategies such as looking at the molecule from different perspectives. It is worth noting that none of these four students equated *R/S* configuration to the specific rotation of the compound, which was one of the most predominant incorrect ideas (described below).

Specific Rotation is Often Equated to *R/S* configuration

Students are taught that chiral compounds can be either dextro- or levorotatory about the same time they are introduced to the idea that chiral centers can have two possible configurations. These two concepts can become interconnected, and some students believed that the sign (+ or –) of the optical rotation of a compound is equivalent to the absolute configuration (*R* or *S*). Approximately one-half of our interview participants (13 of 24) stated that arrangement of substituents is linked to the sign of rotation of plane-polarized light:

> R will rotate clockwise and S will rotate counterclockwise. That is how you determine which one is R and which one is S.

> Chiral center reflects light. R to the right, S to the left.
>
> I vaguely remember something about dextrorotary being R.
>
> I assume that the R-isomer rotates to the right or clockwise.
>
> Plus is R, minus is S.
>
> Whenever we are ranking the things, if it goes counterclockwise, it's S. My common sense tells me that if its S, it goes counterclockwise, if its R, [it goes] clockwise.

The other participants (11 of 24) clearly stated that an absolute configuration is not linked to the sign of optical rotation.

> We can look at the rotation of light, but it is not the way to tell which one is R and S.
>
> You cannot tell which one rotates light in which direction. You have either that direction or other direction.

The question addressing the same idea was included on Quiz 1 (Figure 4.3, question 4). We used more leading wording and phrased it "Can you tell which compound (*R* or *S*) rotates plane-polarized light to the left?" Approximately 60% (N = 52) of students stated that the *S*-isomer rotates plane-polarized light to the left. Several examples of students' reasoning for this relationship are provided below:

> S – the 1, 2, 3 circle to determine R or S goes to the left.
>
> Compound I [R-isomer] because it's substituents move to the left. Each isomer can bend light in a certain direction because of the way the $^{-}$OH group is facing.
>
> Yes, using the Cahn-Ingold-Prelog rules you can determine that compound II [S-isomer] rotates light to the left.
>
> Compound II rotates to the left because it has the S (counter-clockwise) configuration.
>
> S isomer would rotate in the plane of polarized light because when it rotates, you will be able to see the OH group.

As follows from these reasoning statements, students referred to the arrangement of the substituents in the chiral molecule to determine whether the compound is dextrorotatory or levorotatory. Some students probably do not understand that these properties are quite separate from each other. However, drawing a circular arrow when determining an absolute configuration is symbolically similar to the idea of rotating light clockwise or counterclockwise. This incorrect idea can be an example of *substantialism* (Talanquer, 2006) where students substantiate the abstract idea of rotation of light with visible properties of the molecular structure. Under Zoller's classification scheme, this would be a clear example of a misconception.

From our classroom observations, we noticed that both instructors repeatedly said that optical rotation and absolute configuration are two different concepts that are not to be confused with each other. Many textbooks explicitly state that there is no obvious correlation that exists between the configurations of the enantiomers and the direction in which they rotate plane-polarized light (e.g., Solomons, Fryhle, & Snyder, 2014, p. 210; Wade, 2005, p. 183). However, the textbook used by students in the current study did not state explicitly that *R/S* descriptors are not related to the sign of optical rotation.

Biological Activity

Eliel, Wilen, and Mander (1994) factorized stereochemistry into static and dynamic domains. *Static stereochemistry* deals with the stereochemistry of molecules as well as with their energy, spatial arrangement, physical properties, and most of their spectral properties. *Dynamic stereochemistry* deals with stereochemistry of reactions as well as with their stereochemistry requirements and outcomes. We have specifically limited the scope of our research to *static stereochemistry* to avoid the increasing

complexity of *dynamic stereochemistry*. However, there is a concept that is covered quite extensively in the introduction to stereochemistry chapter in most of organic chemistry textbooks. This concept involves the difference in biological activity of enantiomers. This relates to *dynamic* stereochemistry, which is explained by different interactions between chiral molecules of the substance and chiral molecules of the organism.

In the interviews, students were asked about biological activity of different stereoisomers and about why they think stereoisomers may have different activities. In our interviews, all of the students remembered that enantiomers had different biological properties; however, none of them could provide an explanation why they have different biological properties. When probed for that, students produced statements about the role of the structure of the molecule or its optical rotation:

> It maybe has something to do how it reflects the light. One is R and second is S, but that's not reflecting of the light. Reflecting of the light is what you can test by shooting light on it. It does not correspond to R or S. They have the same pH, acidity… So the only thing that is different is how they rotate light… Maybe one reflects the light and second turns it into a body and that cause damage?

In their organic chemistry lectures, students were presented with an example of thalidomide, which may have led to the development of the idea that one enantiomer causes harm, while the other is safe:

> I expect it to be different but I don't know why. One enantiomer could be toxic – if you breathe or something it is bad for you. The other one is fine. This is something different.

Surprisingly, the quote from the student above include some correct ideas. For example, the student knew that configuration and optical rotation are two distinct features, and that enantiomers have the same physical properties. However, the student still could not provide an explanation that was based upon interactions of chiral molecules with a chiral environment. Students' conceptions of different biological activity of stereoisomers fall

under the “no conception” classification. In the study by Linenberger and Bretz (2014), stereochemistry was mentioned only by one of 24 students who were interviewed about various types of complementarity in enzyme-substrate interactions. This can suggest lack of comprehension of stereochemistry and subsequent failure to make connections to other topics such as enzyme-substrate interactions.

On Quiz 1, students were asked if they would expect the same biological activity for the enantiomers of a particular compound. They were also asked to elaborate with a potential explanation for the same or different biological activities for the enantiomers. Twelve students (16%) stated that the biological activity is different without providing sufficient explanation. Four students (5%) stated that they would expect the same type of activity but in the opposite direction. This is probably confusion between biological activity with optical activity. Many students (24 responses, 28%) provided an explanation that included structural characteristics of the molecule. Six students (7%) referred to the differences in the optical activity as a major reason for the discrepancy in biological activity. Only eight students (9%) referred to the differences that are caused by interactions of the molecules of the enantiomers with other compounds such as enzymes or receptors. Only one of the responses explicitly stated that biological environments are chiral. Twelve students (14%) stated that there is no difference between enantiomers in their biological activity because they are the same compound:

> Yes, because they are the same compound only arranged differently in space.
>
> Yes, because they are the same compound just different ways of which OH group is put on the molecule.

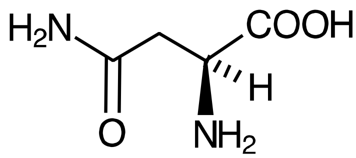
Some students referred to the position of the OH group also when trying to explain why two enantiomers have different biological activity:

> No, because compound I is more favored because the OH group has [the] highest priority and wants to face outward.

> No because of the way OH is facing.

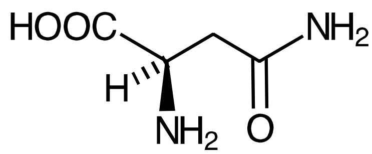
To investigate this incorrect idea more in depth, we administered Quiz 3 to students who had completed the two-semester organic chemistry sequence. On the quiz, students were presented with two pairs of enantiomers (see Figure 4.4) and asked why the biological activity as presented is different. One pair of enantiomers was presented with their trivial names and an *R* or *S* designation. Another pair of enantiomers was presented with a sign of optical rotation. This was done to elicit how the way information is presented affects students' responses.

Version A. Why is *S*-asparagine bitter and *R*-asparagine sweet?



S-asparagine

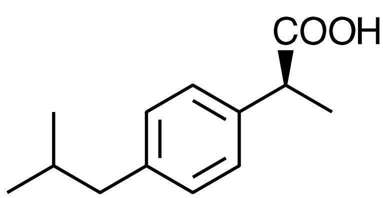
bitter taste



R-asparagine

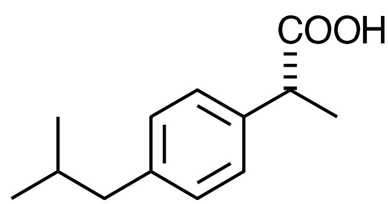
sweet taste

Version B. Why does (+)-ibuprofen relieve pain, but (–)-ibuprofen does not?



(+)-ibuprofen

pain reliever



(–)-ibuprofen

inactive

Figure 4.4. Two versions of a question on Quiz 3 that elicit students' ideas about reason for different biological activity of enantiomers.

All students' written responses that were given for both versions of the question on Quiz 3 were analyzed by the ideas mentioned in the explanations. Of the 16 students who attempted version A of this question, 12 (75%) mentioned the idea that the configuration is responsible for the different taste of these compounds. Two students (13%) explained this difference by different interactions of these molecules to receptors. Two (13%) students did not provide any explanation beyond "I do not know the answer." Of the 16 students who attempted version B of this question, only five (32%) referred to the configuration of the molecules in their answers, while another five (32%) explained the difference in biological activity by the optical activity. Six (38%) student explanations involved an idea of different interactions of enantiomers with enzymes, "pain signals", or "body parts". It is worth noting that this particular example includes a dextrorotatory compound that functions as a pain reliever. Two (13%) student responses included statements that the (+)-stereoisomer is active while the other is not. In these cases, possibly the (+) is viewed as a designation of a certain quality, either optical or biological activity:

> Because (+)-ibuprofen can fit into enzyme when (-)-ibuprofen cannot not.
>
> Because (+)-ibuprofen reacts with polarized light and (-)-ibuprofen cannot not.

Definition of Chirality

Chirality refers to the property of a rigid object that is non-superimposable on its mirror image. In terms of symmetry elements, chiral objects have no symmetry elements of a second kind, such as a mirror plane σ, a center of inversion i, or a rotation-reflection axis (Moss, 1996). This definition of chiral objects is given in most organic chemistry textbooks. However, students often take a shortcut approach and determine whether a

molecule is chiral or not based on the presence of chiral centers. In our interviews with open-ended questions, most of the participants (8 of 11) correctly defined chiral compounds as non-superimposable mirror images; however, later in these interviews the students used the presence of a carbon atom with four different substituents as a criterion for chirality of the molecules. Meso-compounds were assigned by nine of 11 participants as chiral compounds for the reason that they have a tetrasubstituted carbon, while chiral substituted allenes and biphenyls were assigned as achiral by all 11 participants due to the absence of a carbon atom "*with four different things.*" In the interviews with multiple-choice questions, nine of 13 students indicated that meso compounds are chiral and supported their reasoning by the presence of a chiral center.

One of the students described their approach as follows, which probably indicates the reasoning why most of the students consider a molecule chiral if they found chiral centers:

I am only focusing on the carbon, not the molecule entirely.

Five participants (two of 11 in the interviews with open-ended questions, and three of 13 in the interviews with multiple-choice questions) considered only adjacent groups when determining if the molecule is chiral. With this reasoning, 3-bromohexane would be considered an achiral molecule because two substituents (ethyl and propyl) do not differ if only groups that are adjacent to the center are considered. Both ethyl and propyl included CH_2 fragments that are directly connected to the chiral center. Interestingly enough, this reasoning can lead to the correct answer for incorrect reasons. For example, one of our interviewees claimed that 1*R*,3*S*-dibromocyclohexane is achiral because there are not four different groups located in close proximity of the central atom:

CH_2's are the same, so it's not chiral

The compound is indeed achiral due to being a meso form and containing an internal plane of symmetry. We prefer to classify this idea as incomplete rather than incorrect. This particular idea may be formed when students are not exposed to more complicated examples where differences on the substituents occur beyond the first connection point. For example, an instructional overuse of examples such as bromochlorofluoromethane CHBrClF may lead to formation of incorrect ideas such as this one. During a lecture observation, one of the students in the audience asked the lecturing professor if they need to look beyond the first atom when determining if the center is chiral, which represents supporting evidence for this incomplete idea.

A nitrogen-containing chiral compound was assigned as achiral by three of the 11 students who were interviewed with open-ended questions:

Only a carbon can be chiral center, nitrogen cannot.

Only sp^3 hybridized carbons can be chiral centers.

The origin for students reasoning can be purely semantically rooted. The phrases "chiral center" and "chiral molecule" share the adjective "chiral" which may be used as an indication of inclusivity and causality relationships between these two concepts. The following statement provided by one of our participants can support this claim:

Multiple chiral centers will make the whole thing chiral.

Failure to recognize a meso compound as achiral was also observed with examples provided in Quiz 2 (Figure 4.3). Only four (15%) students correctly identified both of the structures presented correctly. The majority of the students (65%) identified only one structure correctly, while 20% did not identify either of the structures correctly.

Krylova (1997) observed various ideas that students express about chiral centers and chirality. For example, students considered substituents as different only if they involve different chemical symbols. Causal statements that a chiral molecule is chiral because it has a chiral center, and an achiral molecule is achiral because it does not have chiral centers were also observed. Lyon (1999), who used concept maps to elicit students' ideas on stereochemistry, found that students think that only sp^3 hybridized moieties can be chiral atoms. Also, students in the Lyon's study were characterizing the presence of chiral atoms as a chemical property.

Physical Properties of Stereoisomers

During interviews, participants mentioned that enantiomers have different physical properties (4 of 11), similar physical properties (2 of 11), or the same physical properties (5 of 11). As for diastereomers, most of the participants (8 of 11) stated that they have the same physical properties, while a few (3 of 11) stated that they have different physical properties.

On Quiz 1, we presented structures of both enantiomers of 1-phenylethanol and asked the questions: "How can these two compounds be separated in the laboratory?" and "Do you expect that **I** and **II** have the same boiling point?" Seventy-two percent of the participants answered that they have different boiling points, while 28% claimed that they have the same boiling point:

> No, because the way the OH is bonded to the carbon affects the boiling point.
>
> No, because one is cis and one is trans, and typically cis and trans have different boiling points.
>
> No, because even though they have similar molecular weight and bond angles, they have different chemical properties.

As one can see, students support their claims by a wide range of beliefs about structure-property relationships. Not all of them are incorrect – for example it is true that *cis* and *trans* isomers typically have different boiling points – but these principles are rather misapplied to this situation. As for the question about separation of the enantiomers, 43% of the students suggested the idea of separating the two by boiling point or distillation, 12% suggested adding some reagent, 8% believed that they can be separated by means of extraction, 9% suggested use of gas chromatography, and 12% claimed that enantiomers cannot be separated. Since Quiz 1 was administered before the laboratory experiment in which students actually separated enantiomers on a chiral GC column, an idea of separation of enantiomers would be feasible and shows that students actually read a laboratory procedure before coming to class. The suggestion of adding a "special" reagent would also be correct, because enantiomers can be separated by conversion to diastereomers, but none of the students who mentioned this method suggested adding a *chiral* reagent:

> They can be seperated [original spelling] by adding ethanol.
>
> With an ether wash.
>
> The two products can be seperated [original spelling] by HCl.

Some students interpreted the verb "separated" as "distinguish" and brought up a wide range of ideas, both correct and incorrect:

> These two products can be separated based on the way they bend polarized light.
>
> Painstaking labor of using electron microscope to individually inspect each molecule and separate them accordingly.

It is worth noting that students often consider enantiomers to be the same compound and automatically attribute equal properties to them:

> They are the same compound, the OH is just placed onto the C at a different angle.

We posed questions about physical properties of diastereomers in our interviews and on Quiz 2, but the analysis revealed that students have a difficult time remembering what diastereomers are. Most of the students could be classified as having no conceptions (Zoller, 1990) about diastereomers:

> Diastereomers would act in somewhat similar manner as enantiomers.

> Diastereomer is the same compound just drawn differently.

Those students who had the correct understanding of diastereomers only applied this concept to chiral molecules:

> Diastereomers are molecules that have multiple chiral centers, and are not superimposable.

> Diastereomers are not mirror images, but they have the same things connected to the same atoms.

Only one student recalled cis-trans isomers as examples of diastereomers. Both instructors teaching the organic chemistry course spent only 5-10 min of their lecture time talking about diastereomers when introducing stereoisomers.

Total Number of Stereoisomers

In our study, we found that students incorrectly applied the formula for calculating the total number of stereoisomers or considered that every chiral atom implies one pair of enantiomers. On Quiz 1, we asked the question "For a compound with X chiral centers what is the total maximum number of stereoisomers?", where X was randomly assigned to be three, four, five, six, or seven. For each of the aforementioned versions of this question, we received 15, 14, 20, 17, and 18 responses, respectively.

We used the wording "chiral centers" (as opposed to "stereogenic centers") because this terminology was used by instructors in both sections. Responses were analyzed across several patterns. The results are presented in Table 4.2. We deliberately chose not to include two chiral centers in the questions because this example produces the same result (four stereoisomers) irrespective of what formula is used. Additionally, the example with two stereocenters was repeatedly used thoroughly during the lectures, which may stimulate simple recall.

Table 4.2.

Numbers of Students that are Using Various Formulas for Determining Total Number of Stereoisomers in Quiz 1.

Number of chiral centers	Total number of students*	Patterns		
		$2n$	n^2	2^n
3	15	2	4	9
4	14	6	8**	8**
5	20	10	6	4
6	17	3	5	9
7	18	10	6	2

* – The number of students who received a version of the question with corresponding number of chiral centers from the first category.
** – Number represents the same eight students.

Students who tried to list all possible combinations generally succeeded at this task and were able to provide a correct answer even if they did not remember the formula. We observed several students (two of 11 participants in the interviews and three among those 15 who received the version of Quiz 1 with three chiral centers) who worked out all of the possible combinations for three chiral centers. Also, it is the formula that provides a maximum number of enantiomers. Quite often, especially in an exam setting, students deal with a meso form. In this case, the formula *number of stereoisomers* = 2^n, where n is the number of chiral centers, predicts too many stereoisomers. Also, for four chiral centers (that give a maximum of 16 stereoisomers), it is not possible to diagnose which formula (n^2 or 2^n) students were using, as both formulas produce the same result.

In our interviews with open-ended questions, we observed a similar pattern of responses. Each student was asked about the total number of enantiomers for compounds with three, five, and seven chiral centers. This was done to ensure observation of a consistent pattern. The "$2n$" pattern (4 of 11), "n^2" pattern (3 of 11), and "2^n" pattern (4 of 11) were observed. Lyon (1999) observed the reasoning that one chiral atom implies one pair of enantiomers which is consistent with our "$2n$" pattern observation.

Summary

Incorrect ideas exist independently of this research study, but the ideas that were detected may depend on how the phenomenon of stereochemistry understanding is investigated. Students' incorrect ideas were influenced by the way the researchers decided to elicit them. For example, in our example with different biological activities of stereoisomers of asparagine and ibuprofen, students provided explanations that involved

some ideas about structure of stereoisomers when they were presented information about their configuration. Conversely, the ideas regarding differences in the rotation of light are more likely to be elicited when students were presented with the information about dextrorotation or levorotation of the stereoisomers.

Triangulation (Creswell, 1998; Merriam, 2009) was used at all three levels of this study. We used multiple sources of information about incorrect ideas, such as interviews and open-ended questions. We interviewed students who were enrolled either in the first or the second semester of organic chemistry for greater generalizability of findings. Two classification schemes, Zoller's classification scheme and Talanquer's commonsense chemistry were used to explain some of the incorrect ideas revealed in this study.

Table 4.3.

List of Incorrect Ideas Pertaining to Stereochemistry.

Incorrect ideas
1. Ranking of substituents for *R/S* configuration is based on electronegativity.
2. Ranking of substituents for *R/S* configuration is based on size.
3. Hydrogen is always pointing backwards in Fischer projections.
4. Specific rotation is equated to *R/S* configuration.
5. Different biological activity of enantiomers is explained by opposite rotation of plane-polarized light.
6. Enantiomers are the same compound, just drawn from different perspectives.
7. Diastereomers are different representation of the same compound.
8. A carbon atom with four different substituents makes the whole molecule chiral.
9. Enantiomers have different physical properties.
10. Enantiomers can be separated by physical methods such as distillation.
11. A total number of stereoisomers is calculated using n^2 or $2n$ formula where n is a number of stereogenic centers.

As can be seen from Table 4.3, students' incorrect ideas apply to all three domains of Johnstone's triangle. For example, incorrect ranking of groups for *R/S* assignment

represents misunderstandings in the symbolic domain, as well as unfamiliarity with the (+)/(–) notation. The conception that a hydrogen atom is always pointing away from the viewer applies both to the particulate domain, as Fischer projections are a form of representation of organic molecules, and to the symbolic domain because Fischer notation involves assumptions that substituents on the horizontal line are projecting towards and substituents on the vertical line are projecting away from the viewer. The conception about enantiomers and diastereomers being different representations of the same molecule is an example of an incorrect idea that occurs in the particulate domain. A cross-domain incorrect idea involves students making a direct connection of optical rotation (macroscopic domain) to *R/S* configuration (particulate domain). An assumption that physical properties are different for enantiomers is illustrative of an incorrect idea that occurs in the particulate domain.

Students possess a variety of incorrect ideas that are related to all area of stereochemistry. Some incorrect ideas may be caused by limited learning experiences of students, while other incorrect ideas may represent deeply rooted heuristics that are resistant to change. Other sources could lead to a formation of incorrect ideas is lack of instructional time and ambiguity of textbook explanation.

Teaching Implications

Practitioners may find the results of this study useful to transform and enhance their teaching approach. Numerous teaching strategies – both at the laboratory and lecture level – can be used to revert these incorrect ideas or suppress their development. For example, the fact that some students assign priority of the substituents based on their electronegativity or polarity may be explained by the abundance of tasks where students

rate entities based on their electronegativity. Another example that is very commonly used in organic chemistry textbooks is halogen substituted methane molecules. Even though this example is very obvious and simple, it has very little practical value. In fact, optical rotations for the enantiomers of bromochloroflorometHane, CHBrClF, were not measured until recently due to difficulties with their separation (Polavarapu, 2002). Although often mentioned as a prototype for chiral molecules, fully substituted bromochlorofluoroiodomethane, CBrClFI, has not been synthesized (Gilchrist, 1995, p. 228). These molecules can serve as examples at an earlier stage of stereochemistry instruction, but instructors should provide more examples with aliphatic substituents.

The confusion of optical rotation with absolute configuration can be alleviated by a laboratory experiment. For instance, students have an opportunity to measure optical rotation of compounds with known absolute configuration and directly observe that compounds with a certain configuration of a stereogenic center can be either levo- or dextrorotary. This type of activity would also provide students with experience operating a polarimeter, which was reported to be one of the skills desirable in chemistry industry (Fair, Kleist, & Stoy, 2014).

Instructors may benefit from teachings examples for organic chemistry that may be found in Massive Online Open Courses (Leontyev & Baranov, 2013). The variety of instructional materials from Massive Online Open Courses including videos, tutorials, problem sets can be used to enhance organic chemistry instruction and increase students' exposure to content. For example, in his teaching of stereochemistry, Michael McBride (2011), illustrates the difference in biological activity accompanied by shaking his right

hand with the right hand of the student and his left hand with the right hand of the student, This example is not only exceptionally clear, but involves kinesthetic learning.

The use of physical models has been emphasized many times in the literature. In addition, models provide a unique experience to expand the psychomotor domain of learning, which is a rare commodity in chemistry education. Models are especially effective for students that have problems visualizing molecules in three dimensions (Al-Balushi & Al-Hajri, 2014).

The current study would lead to a development of a Stereochemistry Concept Inventory. Incorrect ideas identified in this study will be used to write distracters for multiple-choice items that constitute the inventory.

CHAPTER V

ARTICLE 2: DEVELOPMENT OF A STEREOCHEMISTY CONCEPT INVENTORY

Abstract

Stereochemistry is one of the most difficult and fundamental concepts of organic chemistry. Failure to master stereochemical concepts can be a barrier for students' success in organic chemistry courses. Organic chemistry instructors may find themselves in need of an assessment tool that quickly gives information about their students' ideas (both correct and incorrect) related to stereochemistry.

We report a development process of a Stereochemistry Concept Inventory (SCI) that assesses organic chemistry students' knowledge and skills within the realm of stereochemistry. The test items were based on important topics identified by the Stereochemistry Instruction Survey, and distracters were based on students' difficulties and incorrect ideas identified in a qualitative study. An iterative process involving multiple test administrations and continuing item refinement was used to obtain psychometric qualities of the scores produced by the SCI. This paper outlines the analysis of data obtained from 439 students from 17 different institutions across the U.S. Multiple measures of detecting reliability of the scores produced by the SCI are discussed.

Introduction

As chemists always need high-quality measurement tools, chemistry education researchers also find themselves in the situation where they need instruments that produce reliable data and valid inferences. However, while the measurements in chemistry are objective and can be observed directly (e.g., melting point), in educational research, variables are latent and cannot be measured as a result of direct observation. Quite often a need for a measurement tool is absolutely crucial, for example, in quasi-experimental and experimental studies in which performances of two or more groups are compared to determine the effect of a certain pedagogical intervention.

Concept inventories are standardized diagnostic instruments that assess how well students' conceptual knowledge fits the commonly accepted knowledge in the discipline. From a methodological perspective, a concept inventory is a multiple-choice assessment test that probes the understanding of a single topic (Bailey, 2009), also called a construct (Wilson, 2005). Distracters for questions are composed typically from students' incorrect ideas that are also called in the science education research literature misconceptions, student difficulties, or alternative conceptions.

Concept inventories can be used for the assessment of large groups of students, both for educational and research purposes. Other methods to test understanding are also available, such as in-depth clinical interviews, knowledge trees, or concept maps. However, data collection by means of these other methods requires much time and effort. Interviews yield massive qualitative data, while analysis of concept maps provides both qualitative and quantitative data. Direct interpretation and analysis of such pieces of data are not always possible, and perhaps conducting interviews is not a reasonable approach

for the classroom instructor who teaches a large class. Thus, practitioners cannot always benefit from using interviews or concept map analysis in their classrooms, especially when dealing with large student populations. Concept inventories are efficient assessment methods for classroom use, and they provide interpretable data. High quality assessment can empower chemistry researchers with tools to reveal cognitive structures (Taagepera & Noori, 2000).

Existing Assessment Tools

A variety of concept inventories have been reported in different STEM fields: biology (Anderson et al., 2002), physics (Steif & Dantzler, 2005), astronomy (Bailey et al., 2012), computer sciences (Herman, Loui, & Zilles, 2010), and others. We have summarized the basic information about existing chemistry concept inventories in Table 5.1. The list represents measures of cognitive domain, although several instruments (Brandriet & Bretz, 2014a; McClary & Bretz, 2012) presented are intended to measure both content knowledge and confidence in that content knowledge. These instruments can be considered cross-domain measures because they measure both cognitive structures and affective components that are associated with cognitive structures.

Table 5.1.

List of Assessment Tools to Measure Chemistry-Related Concepts.

Brief description of the instrument(s)	References
Diagnostic instrument to evaluate concepts of covalent bonding	(Peterson, Treagust, & Garnett, 1989)
Instrument to asses understanding of mole concept	(Krishnan & Howe, 1994)
Two-tier instrument to assess alternative conceptions about chemical bonding	(Tan & Treagust, 1999)
Test to identify student conceptualization about chemistry equilibrium	(Voska & Heikkinen, 2000)
Concept inventory for assessment alternate conceptions among first-semester general chemistry students	(Mulford & Robinson, 2002)
Diagnostic instrument to assess understanding of qualitative analysis	(Tan, Goh, Chia, & Treagust, 2002)
Chemistry concept inventory for use in chemistry, material, and engineering courses	(Pavelich et al., 2013)
Diagnostic instrument to determine understanding of ionization energy	(Tan, Taber, Goh, & Chia, 2005)
Diagnostic instrument for evaluating students ability to describe and explain chemical reactions using multiple levels of representations	(Chandrasegaran, Treagust, & Mocerino, 2007)
Diagnostic instrument to assess students' understanding of separation of matter	(Tüysüz, 2009)
Concept inventories to diagnose understanding in a year-long organic chemistry course	(Cartrette & Dobberpuhl, 2009)
Diagnostic tool for assessing understanding of the particulate nature of matter	(Nyachwaya et al., 2011)
Structure and motion of matter survey to assess implicit assumptions about particulate nature of matter	(Stains, Escriu-Sune, Molina Alvarez de Santizo, & Sevian, 2011)
Chemistry concept reasoning test for measuring conceptual understanding and critical scientific thinking of general chemistry models and theories	(Cloonan & Hutchinson, 2011)
Systemic assessment question for meaningful understanding of organic chemistry	(Vachliotis, Salta, & Tzougraki, 2013; Vachliotis, Salta, Vasiliou, & Tzougraki, 2011)
Instrument to assess students' understanding of foundational concepts from general chemistry and biology before biochemistry network	(Villafañe, Bailey, Loertscher, Minderhout, & Lewis, 2011)
Three-tier test to assess high school students' understanding of acids and bases	(Cetin-Dindar & Geban, 2011)
Instrument for testing the ability to use implicit information from Lewis structures for various purposes	(Cooper et al., 2012)
Diagnostic tool to identify alternative conceptions related to acid strength held by organic chemistry students	(McClary & Bretz, 2012)

Table 5.1, continued.

Brief description of the instrument(s)	References
Concept inventory to assess understanding enzyme-substrate interactions	(Bretz & Linenberger, 2012)
Test to assess pre-service teachers' misconceptions about global warming, greenhouse effect, ozone layer depletion, and acid rain	(Arslan, Cigdemoglu, & Moseley, 2012)
Two-tier diagnostic instrument to assess solution chemistry concepts	(Adadan & Savasci, 2012)
The Thermochemistry Concept Inventory to test conceptual understanding of thermochemistry concepts by first-semester general chemistry students	(Wren & Barbera, 2013, 2014)
The Bonding Representation Inventory to identify students misconceptions related to covalent and ionic bonding representations	(Luxford & Bretz, 2014)
The Redox Concept Inventory to students' understanding of oxidation-reduction reactions in symbolic and particulate domain	(Brandriet & Bretz, 2014a, 2014b)
Assessment instruments measuring understanding of specific components of scale	(Gerlach, Trate, Blecking, Geissinger, & Murphy, 2014)
Two-tier diagnostic instrument of understanding of electrochemical cells	(Loh, Subramaniam, & Tan, 2014)
Visual-Perceptial Chemistry Specific assessment tool	(Oliver-Hoyo & Sloan, 2014)
Metabolic Pathways Visualization Skill Test	(dos Santos & Galembeck, 2015)
The General, Organic, and Biological Chemistry Knowledge Assessment to assess understanding of chemistry concepts relevant to nursing practice	(Brown, Hyslop, & Barbera, 2015)
Instrument to measure understanding of nanoscience and nanotechnology	(Schönborn, Höst, & Lundin Palmerius, 2015)
Chemical Representation Inventory for measuring students' knowledge of nomenclature, chemical equations, skeleton formulae, ball-and-stick models, and translations between various representations	(Taskin, Bernholt, & Parchmann, 2015)

The formats of concept inventories and diagnostic tools vary. Several published assessment tools have an open-ended format (Nyachwaya et al., 2011; Stains et al., 2011), and several are two-tiered tools; for example, the test published by Tan and Treagust (1999) includes both question and reason tiers. However, the majority of assessment tools is a collection of multiple-choice items. Some concept inventories (Luxford & Bretz, 2014) are composed of a mixture of two-tiered questions and single-tier questions. An example of the systemic assessment published by Vachliotis et al. (Vachliotis et al., 2013) includes concept maps with missing pieces that are supposed to be filled in by students.

Different strategies have been used in the development and validation of concept inventories. For example, Gerlach et al. (2014) used a previously proposed Scale Concept Trajectory and wrote items that address each stage and component of the trajectory. Brown et al. (2015) used results from a survey of nursing educators and chemistry instructors to determine the content basis for the General, Organic, and Biological Chemistry Knowledge Assessment tool. Adadan and Savasci (2012) used both Turkish national curriculum standards and commonly used textbooks in Turkey to establish the domain of students' understanding of solution chemistry. However, sometimes getting agreement between end-users can be difficult, and the outcomes of these processes depend on what method is used. The more rigorous the methodology used, the more trust can be put into results of the study, and greater is the likelihood of the assessment being recognized by the community. In an exemplary study by Streveler et al. (2011), the Delphi consensus method was used to uncover relative importance of thermal and

transport concepts during the developmental process of the Thermal and Transport Science Concept Inventory.

Inventories are also targeting different populations. Villafañe et. al., (2011) developed an assessment tool for upper division biochemistry students. High school students' conceptions of bonding were investigated by Tan and Treagust (1999) using a two-tier instrument. Another concept inventory developed by Bretz and Luxford (Luxford & Bretz, 2014) was intended for both college students and high school students, in order to investigate their ideas about different representations of covalent and ionic bonding. While the detailed comparison of assessment tools is not the purpose of this paper, we would like to reiterate that there is a great diversity in the development process and subsequent use of the assessment tools. Since the quality of the written research report may be discrepant with the quality of the research study, the individual differences between assessment tools is more likely to be even larger than it appears from the manuscripts that reported their development.

Although there is a relative abundance of assessment tools for general chemistry concepts and high-school chemistry, the assessment tools for organic chemistry are still scarce, and a tool for reliable diagnosis of specific organic chemistry concepts is a rare commodity. Despite the relative absence of validated tools, one can find a plethora of articles in the *Journal of Chemistry Education* that report various instruction practices that meant to enhance student understanding of organic chemistry concepts. However, few of these methods are supported by actual measures of cognitive outcomes. To our knowledge, only four assessment tools were reported that are suitable for conceptual assessment in the organic chemistry class. McClary and Bretz (2012) describe in detail

the development and implementation of a concept inventory that assesses students' knowledge of organic acids strength. Their instrument is cross-domained and assesses both knowledge (cognitive domain) and confidence (affective domain); the knowledge part is two-tiered, containing question-tier items and reason-tier items. Cartrette and Dobberpuhl (2009) reported the development of an assessment battery for a two-semester organic chemistry course. The items on their concept inventories are meant to assess the core concepts from both semesters of the organic chemistry course and are suitable for a self-assessment. Vachliotis et al. (2013) developed and validated an assessment instrument based on concept maps. The instrument was designed to capture 11th grade students' meaningful understanding of organic chemistry concepts. The authors used a systemic approach to teaching and learning as a guiding model in the development of their assessment schema. A recently published Chemical Representation Inventory (Taskin et al., 2015) is suitable for measuring students' knowledge of various representations that are used in organic chemistry and ability to interconvert between various representation.

Purpose and Benefits of the Study

From a survey of organic chemistry instructors, Duis (2011) reported that instructors identified stereochemistry as one of the most challenging and foundational areas of chemistry. For research and instructional purposes, it is beneficial to have an assessment tool that measures the level of understanding of the concepts that relate to stereochemistry. To our knowledge, there is no assessment tool reported in the chemistry education research literature that measures the understanding of stereochemistry. The purpose of this study was to develop an instrument to assess students' understanding of

the stereochemistry concepts considered important by experts. The instrument, measuring the presence and abundance of incorrect ideas, can be useful not only for organic chemistry instructors, but also for instructors whose courses require stereochemistry concepts as a prerequisite.

Methodology and Guiding Literature

Treagust (1988) in his seminal paper described a developmental process for diagnostic tests, which has been implemented to create a number of assessment instruments. He presented three broad areas required for creating these tests: "defining the content," "obtaining information about students' misconceptions," and "developing a diagnostic test." These areas can be subdivided into ten individual steps:

1. Identifying propositional knowledge statements.
2. Developing a concept map.
3. Relating propositional knowledge to a concept map.
4. Validating the content.
5. Examining related literature.
6. Conducting unstructured student interviews.
7. Developing multiple-choice content items with free response.
8. Developing the two-tier diagnostic tests.
9. Designing a specification grid.
10. Continuing refinements.

Many researchers who have developed concept inventories in different fields modified or even eliminated steps proposed by Treagust. A review by Lindell et al. (2007) summarized the design and validation methodologies for 12 concept inventories in the field of physics and astronomy. The authors report a lack of consistent methodology, differences in detecting the concept domain of interest, and discrepancies in detecting reliability and validity. Published reviews of chemistry assessment tools (Barbera & VandenPlas, 2011; Bretz, 2014) highlight and elaborate on discrepancies that occur in the

design process of assessment tools. To summarize these reviews, it can be said that no two instruments are alike, neither in the development process nor in the implication for practice.

In general, the development process follows an iterative design in which a version of the instrument is created and administered to a sample of participants. The obtained data are analyzed, and revisions are made based upon the outcome of the analysis. The subsequent version is then administered to another sample of participants, and the process repeated. The choice of analysis method can vary from basic percentages of correct answers (Mulford & Robinson, 2002) to advanced statistical techniques, such as, 2-parameter Item Response Theory (Oliver-Hoyo & Sloan, 2014), cluster analysis (Brandriet & Bretz, 2014a), Rasch modeling (dos Santos & Galembeck, 2015; Taskin et al., 2015; Wren & Barbera, 2014), or a confirmatory factor analysis (Villafañe et al., 2011). The choice of statistical technique is often based on a theoretical basis that underlines a construct of interest. For example, Villafañe et al. (2011) developed an instrument that tested eight concepts with three items per each concept. Thus, a confirmatory factor analysis was used to assess the fit of experimental data with the initially proposed structure. Brandriet and Bretz (2014a) used a cluster analysis to identify groups of students with similar response patterns in cognitive and affective measures.

The two most crucial aspects when an instrument is developed are validity and reliability. Arjoon, Xu, and Lewis (2013) evaluated the development process of 20 instruments that were used in chemistry education research since 2002. These instruments were evaluated against the Standards for Educational and Psychological Testing, jointly

developed by the American Educational Research Association (AERA), the American Psychological Association (APA), and the National Council on Measurement in Education (NCME). According to the evaluation, only 11 of these 20 instruments examined the evidence based on content, only six instruments included discussion of the response process, and only nine instruments reported evidence based on the internal structures of the instruments. However, 19 instruments included information about relationships to other variables. Information about internal consistency was reported for 15 instruments, while results of replicate administrations (temporal stability) were reported for only six instruments. Authors proposed that a chemistry education community should strive to collect various psychometric evidence for the instruments used in research studies to make chemistry education a theory-driven and data-driven field of science. Their findings are in accordance with an observation of Brown and Wilson (2011), which showed that most of the measures lacked a clearly stated model of cognition.

The development of the Stereochemistry Concept Inventory, described herein, consisted of several phases. Figure 5.1 illustrates individual phases of the instrument development. Data were collected for each stage by means of national surveys, interviews, and administrations of several versions of the Stereochemistry Concept Inventory to students of the targeted population. The study started with a qualitative exploration of the students' incorrect ideas and identification of topics considered important by organic chemistry instructors. The pilot version of the Stereochemistry Concept Inventory consisting of 30 questions, denoted as SCI-30Q in this manuscript, was developed and administered to students. Psychometric analysis, information from

students' interviews, and feedback from the practitioners were used to refine the items. Both modification and elimination of items were used as a result of refinement. The final version of the SCI consisting of 20 questions, denoted as SCI-20Q in this manuscript, was subjected to expert content review in addition to psychometric analysis.

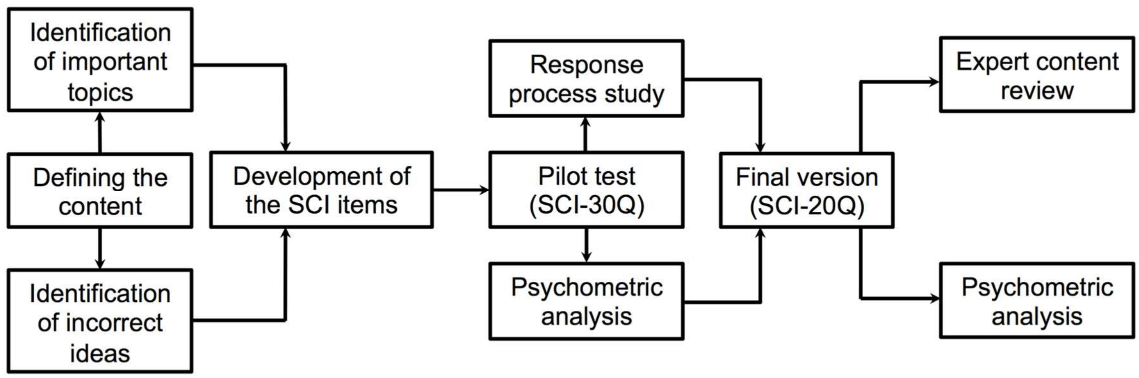


Figure 5.1. The stepwise model for development and evaluation of the Stereochemistry Concept Inventory (SCI).

From a methodological perspective, this study utilizes mixed-method sequential exploratory strategy (Creswell & Clark, 2007), starting with qualitative data collection. Subsequent steps are built on the results obtained from the qualitative phase. We initially started with the incorrect ideas revealed in the qualitative study and developed a test assessing these incorrect ideas. We chose a mixed-method methodology to achieve greater interpretability of our results, which provided a novel tool that can be used for the assessment of students' stereochemistry understanding. As Towns (2008) commented, mixed methodology allows chemistry education researchers to explore a phenomenon with significantly larger depth and breadth than a unimodal design.

We considered data-driven strategies for constructing multiple-choice questions for the SCI. Haladyna, Downing, and Rodriguez (2002) developed and validated

taxonomy of 31 multiple-choice item writing guidelines. Two sources of evidence were employed: the consensus achieved from reviewing recommendations from 27 textbooks on educational testing and the results of 27 research studies and reviews published since 1990. The taxonomy is intended for writing multiple-choice questions for wide assessment purposes. The taxonomy is also recommended for the development of test items for large-scale assessment tools. The taxonomy addressed content, formatting, and style questions, as well as writing the stem and the alternatives. Towns (2014) proposed a set of suggestions for writing multiple-choice questions in chemistry. Relevant examples and detailed guidelines of this review cover content, stem, and response construction of chemistry multiple-choice tests. To the extent possible, recommendations provided in both reviews (Haladyna et al., 2002; Towns, 2014) were followed when constructing the questions of the SCI.

Protection of Human Subjects

The University of Northern Colorado requires review and approval of all research projects involving human subjects. Institutional Review Board (IRB) approvals were obtained from the University of Northern Colorado and other participating schools (either as a result of IRB transferred or submission of a new application). All data collected in this study were through voluntary participation. Some participants received feedback on their performance and/or a review session.

Data Analysis

Data sets from the SCI administrations were analyzed using jMetrik (Meyer, 2013). Statistical tests were performed using SPSS (version 18.0.0.). Effect sizes were calculated using David B. Wilson's online effect size calculators from the Campbell Collaboration (Wilson, n.d.). Incomplete cases were removed from the final datasets (listwise deletion) which were analyzed for psychometrics, reliability coefficients, and group differences. Contingency tables (crosstabulations for two questions) were produced based on both complete and incomplete cases; incomplete cases were included because they contain responses for the two selected questions included in the contingency table.

Pilot Version of the SCI (SCI-30Q)

Stereochemistry Instruction Survey

A Stereochemistry Instruction Survey based on the content analysis of the organic chemistry textbooks was developed and administered to establish a content basis for the Stereochemistry Concept Inventory. To make the Stereochemistry Concept Inventory appropriate for use in any organic chemistry course, it should cover the content shared among the most used textbooks. To identify relevant content, we have conducted the content analysis of the nine most widely used organic chemistry textbooks (Houseknecht, 2010). A list of 34 topics was compiled as a result of the content analysis of the textbooks.

The Stereochemistry Instruction Survey attempted to cover the entire spectrum of stereochemistry topics taught in undergraduate organic chemistry as well as the learning goals pursued by instructors. The survey was sent to 1,028 organic chemistry instructors whose emails were retrieved from the institutions' websites. The survey started with a

consent form followed by an inclusion question (Do you currently teach or have taught within five years a two-semester organic chemistry course?) to assess eligibility of the participants. The participants were then presented with a list of 34 topics identified from the content analysis. The participants were asked to rate the presented topics as *important*, *optional*, and *not important*. The definitions of these categories were provided to the participants for clarity and unambiguous interpretations: *important* refers to topics that are considered relevant, always taught, and included as a part of assessment; *optional* refers to topics that may be taught if time allows or students are assigned to read about them in the course textbook; *not important* refers to topics that are never or almost never taught and are not considered relevant to stereochemistry instruction. The second part of the survey consisted of a list of 28 learning objectives identified from the same review of organic textbooks. The learning objectives were used for constructing the table of specification ("test blueprint") described later in this manuscript. At the end of the survey, participants were asked several demographic questions.

After two reminders (sent within two-week intervals), a total of 219 participants (a response rate of 21%) completed the Stereochemistry Instruction Survey. The participants taught at institutions offering doctoral (53%), Master's (10%), Bachelor's (28%), and Associate's (10%) as the highest degree in chemistry. Detailed demographic characteristics of the participants are presented in the supplemental material (Table 5.10, survey 1). Percentages of responses that indicate topics as *important*, *optional*, and *not important* are presented in Table 5.2.

Table 5.2.

Instructors' Rankings of Stereochemistry Topics.

Topics	Important, %	Optional, %	Not important, %
Constitutional isomers	98	2	0
Cis/trans isomers for compounds with a double bond	100	0	0
Cis/trans isomers for alicyclic compounds	91	8	1
Conformers	95	5	0
Enantiomers	100	0	0
Diastereomers	99	1	0
Meso compounds	93	7	0
Cahn-Ingold-Prelog priority rules	98	1	1
R/S nomenclature	99	0	0
+/- nomenclature	53	38	9
d/l nomenclature	26	55	20
D/L nomenclature	36	48	16
Erythro and threo nomenclature	13	45	42
Block diagram of polarimeter	50	39	11
Plane-polarized light	80	18	2
Optical rotation	89	10	1
Levorotatory and dextrorotatory compounds	69	25	6
Specific rotation	65	28	7
Chirality at atoms other than carbon (e.g., nitrogen)	39	50	10
Chirality of disubstituted cyclohexanes	82	14	4
Chirality of allenes	26	55	19
Chirality of substituted biphenyls	18	51	31
Chirality of helicenes	8	48	44
Resoluton of enantiomers through formation of diastereomers	67	29	4
Resoluton of enantiomers through enzymatic binding	30	50	20
Resoluton of enantiomers through chiral chromatography	28	51	21
Racemic mixtures	99	1	0
Enantiomeric excess	65	30	5
Equivalence of physical properties of enantiomers	92	6	1
Non-equivalence of physical properties of diastereomers	92	6	1
Different biological activities of stereoisomers	81	18	1
Total number of stereoisomers for a compound with *m* chiral centers	82	15	3
Enantiotopic and diastereotopic atoms	50	42	9
Prochirality	28	55	17

Note: Numbers in each column represent percentages of respondents who selected the corresponding category. Percentages may not add up to 100% due to rounding.

Anchoring Concept Map

Using a set of topics that received high rankings from the survey, a concept map (Figure 5.2) was constructed to show the interrelationship of these topics. A prototype for this concept map was taken from the Organic Chemistry textbook by Solomons, Fryhle, and Snyder (2014) with some modifications. For example, since organic reactions are covered after the stereochemistry topics in the majority of textbooks, all concepts related to reactions were eliminated from the original textbook concept map. Some concepts were reformulated to ensure clarity. Three organic chemistry instructors were asked to validate the scientific correctness of the content and whether they cover this content in their organic chemistry instruction. The goal of having the concept map early in the development process was to define the topics being tested. The major purpose of the concept map is not to include all topics that represent stereochemistry knowledge, but rather to show interconnection of concepts included in the pilot version of the SCI (SCI-30Q).

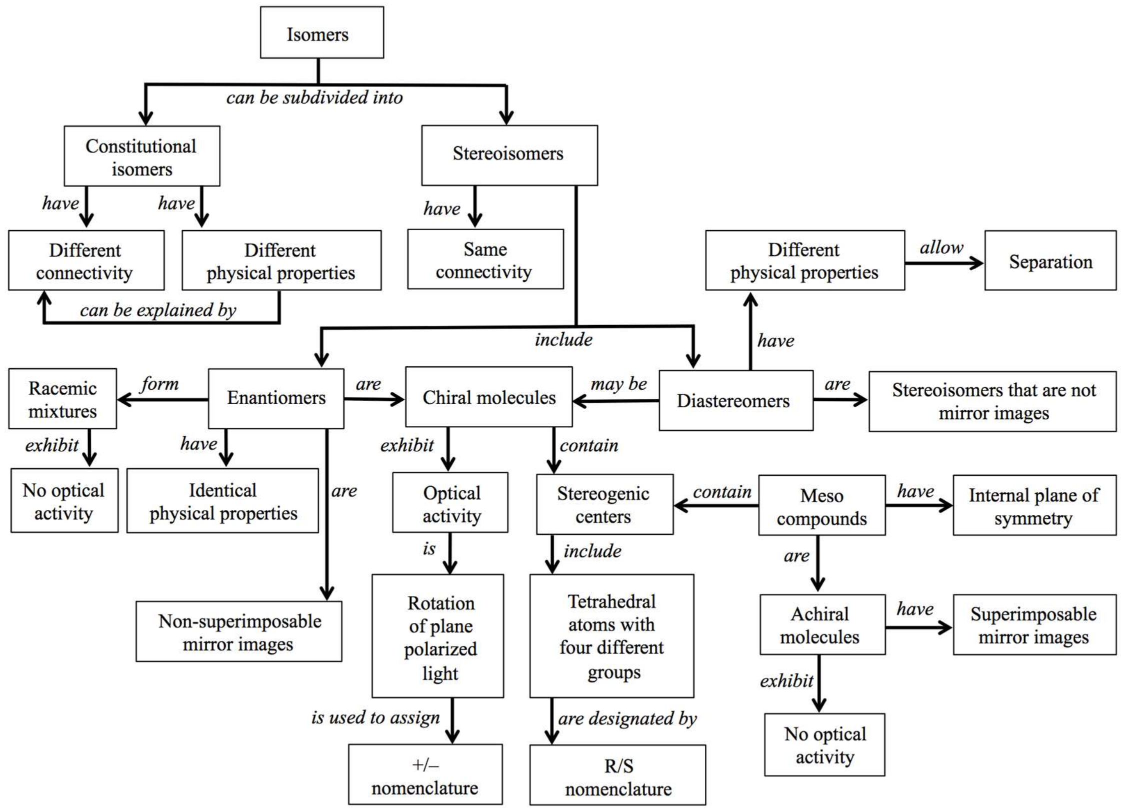


Figure 5.2. A stereochemistry concept map used for anchoring concepts included in the SCI-30Q. Adopted with modifications from Solomons et al. (2014).

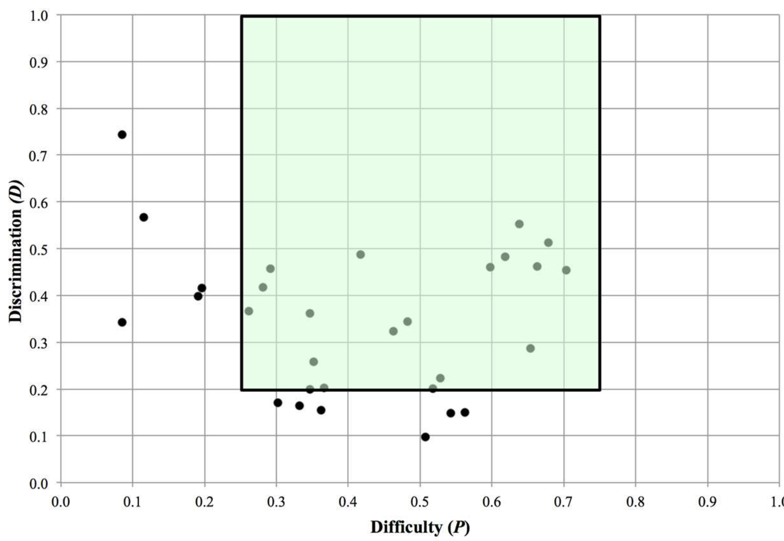
Pilot Testing of SCI-30Q

Thirty-two questions were written that address important concepts of stereochemistry as identified in the Stereochemistry Instruction Survey. These questions were administered to students ($N = 32$) enrolled in a biochemistry course during their first class period. Based on the collected response patterns, we eliminated the most difficult question ($P = 0.0$) and the easiest question ($P = 1.0$). The remaining set of 30 items constituted a pilot version (SCI-30Q) of the Stereochemistry Concept Inventory. This version was administered in a paper-and-pencil form at three different institutions: one doctoral/research university in the Mountain West region ($N = 114$) and two liberal arts colleges in South West ($N = 55$) and Mid-Atlantic Northeast regions ($N = 30$) of the U.S.

The SCI-30Q was administered within two weeks after stereochemistry was covered, and participants were given 25-30 min to complete the test in organic chemistry classes. For the psychometric analysis, all three datasets were combined. A total of 199 students answered all questions from the SCI-30Q. Three students submitted tests with missing responses; consequently, these were eliminated from the dataset (listwise deletion). The overall mean performance was 12.49 (SD = 3.45) out of 30, which constituted 42%. The scores ranged from 3 (10% correct) to 24 (80% correct). The overall reliability was 0.40 as estimated by the KR-21 coefficient. Individual items were also analyzed for their difficulty (percent of correct responses) and discrimination (ratio of participants who answered the item correctly in high-achieving and low-achieving groups).

The purpose of the pilot testing was to collect evidence about item functioning and use this evidence to modify or eliminate non-functioning items. We conducted item analysis to determine the items that functioned well. For educational assessment purposes, the recommended range of difficulty is .25−.75 (Kline, 2005) and discrimination above .20 (Ebel, 1972). Based on the psychometric analysis, five items were found to be too difficult, and six items did not have acceptable discrimination power. Figure 5.3A summarizes the information about item difficulty and discrimination indices for the pilot version of the instrument. We used this information supported by evidence from the response process study to make modifications leading to the final version of the SCI (SCI-20Q). The results of psychometric analysis of SCI-20Q are presented in Figure 5.3B for comparative purposes, while a detailed explanation is provided in the "Item analysis of SCI-20Q" section.

A.



B.

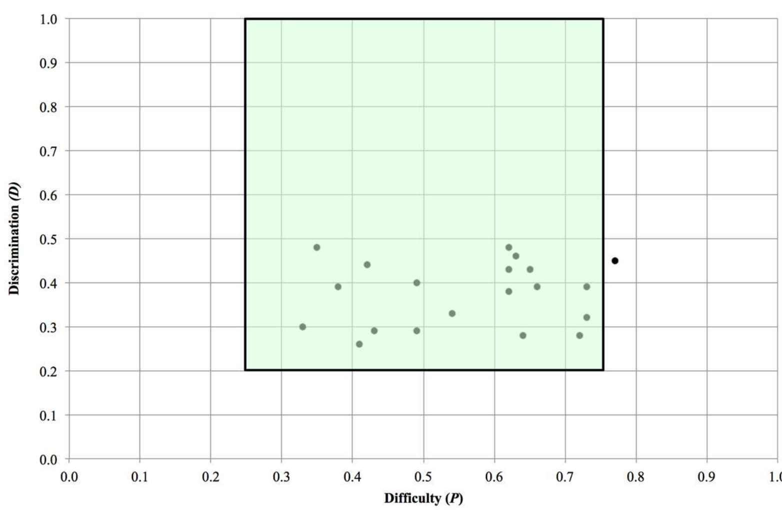


Figure 5.3. Difficulty (P) and discrimination (D) indices for both versions of the SCI. A: pilot version, SCI-30Q; B: final version, SCI-20Q. The highlighted area represents items with acceptable psychometric characteristics.

Validation Interviews for the SCI-30Q

Validation interviews are often used to ensure response process validity (Brandriet & Bretz, 2014b; Wren & Barbera, 2013). The goal of these interviews is to ensure that students interpret the question correctly and confirm that their responses align with their conceptual understanding. Specifically, students should choose the correct option if they have the correct understanding of a phenomena, and students who have a certain incorrect idea should be inclined to choose the distractor that represents that incorrect idea. Cases when students who have the correct idea select the distractor that represents an incorrect idea and cases when students who have incorrect knowledge select the correct option are indicative of problematic questions and are more likely to introduce construct-irrelevant variance to the scores that are produced by an assessment tool.

We interviewed 13 students who had taken the pilot, a 30-question version of the SCI. All students were enrolled in a doctoral/research university located in the West Mountain region of the U.S. All students were enrolled in one of two sections of the first semester of organic chemistry course; different instructors taught the two sections. As an incentive for the participation in the interviews, students were offered a review session before the final exam. All interviews took place during the 13^{th} week of the course. The interviews lasted for 45-60 min. During the interviews that occurred after students took SCI-30Q, students were asked to justify their answers and probed into their understanding.

Based on the analysis of students' reasoning, eight questions did not show strong evidence of the response process. It is worth noting that seven out of eight items that

were problematic from the response validity point of view also did not exhibit suitable psychometric characteristics (*D* = .25−.75; *P* > .2). Figure 5.4 shows two examples of the questions that were eliminated. The basis for the elimination of item #6 was that five of the participants did not know the meaning of the (+) notation used in the question. Students are more inclined to select option C because of their unfamiliarity with the (+/–) notation:

> I am not sure what this means (points at +) and that's why I chose C.

The reason for eliminating item #8 was that interviews showed evidence for construct irrelevant variance, such as test-wise strategy, which two of the participants used:

> Whenever you have definite like "always" I always count it out. So I chose B.
>
> I chose D because that is seeming wrong because of "always" [points at statement I]. I have eliminated this one [points at II] because I did not apply chiral or achiral to meso."

Item #6 (SCI-30Q)

What is the absolute configuration of a chiral atom in (+)-naproxen?

A. *R* because it rotates plane polarized light clockwise
B. *S* because it rotates plane polarized light counterclockwise
C. It can be either *R* or *S*

Item #8 (SCI-30Q)

Which of the following statements is (are) true:

I. Compounds with multiple chiral centers are always chiral
II. Meso compounds are achiral

A. **I** only
B. **II** only
C. **I** and **II**
D. Neither **I** nor **II**

Figure 5.4. Examples of two items that were eliminated from the SCI-30Q based on a lack of evidence from the response process validity study.

Final Version of the SCI (SCI-20Q)

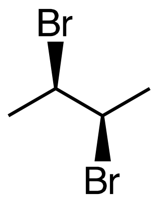
Development of the Final Version

SCI-20Q

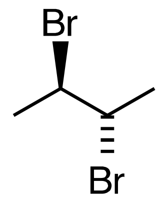
We have applied several modifications to the SCI-30Q version that yielded the final version of the instrument (SCI-20Q). The modifications included rewording or removal of some items. When deciding which items to remove, we considered multiple factors: psychometric characteristics, response process, and comprehensive coverage of important stereochemistry topics. We eliminated ten items from the SCI-30Q. Among these ten items, nine did not exhibit acceptable psychometric characteristics and significant evidence of the response process validity, and one item exhibited acceptable psychometrics but lacked evidence of response process. Modifications of the remaining items based on psychometric analysis included the removal of one of the non-functioning distracters (three items) and simplification of cognitive tasks needed to solve a question (two items). The final version contains 10 items with three response options (one correct and two distracters) and 10 items with four response options (one correct and three distracters). Our decision to include items with only three alternatives was based on a meta-analysis by Rodriguez (2005) that revealed test items with three-response options function similar to those with four-response options. One of the questions (see Figure 5.5) was modified to decrease the number of cognitive steps involved in the solution process. The first step was to determine the relationship between two given structures, and the second step was to decide the relative physical properties of these compounds. In our interviews, only two students were able to correctly determine the relationship between the two structures, so we decided to specify the relationship in the item stem.

Pilot (SCI-30Q) version

Which statements are true about the physical properties of the following compounds?



(2*R*,3*R*)-2,3-dibromobutane



(2*R*,3*S*)-2,3-dibromobutane

A. Boiling points are the same. Optical rotations are equal, but in opposite directions.
B. Boiling points are the same. Optical rotations are numerically different.
C. Boiling points are different. Optical rotations are numerically different.
D. Boiling points are different. Optical rotations are equal, but in opposite directions.

Final (SCI-20Q) version

Which statements are true about the physical properties of diastereomers that contain two stereocenters?

A. Boiling points are the same. Optical rotations are equal, but in opposite directions.
B. Boiling points are the same. Optical rotations are numerically different.
C. Boiling points are different. Optical rotations are numerically different.
D. Boiling points are different. Optical rotations are equal, but in opposite directions.

Figure 5.5. The pilot and final versions of one of the questions used in the Stereochemistry Concept Inventory. Revisions were done based on interviews and instructors' feedback.

We made a decision to keep or modify some of the items that did not show appropriate psychometric characteristics, because eliminating all of these items could lead to construct underrepresentation. While a test might exhibit excellent psychometric characteristics, it may not test all aspects of the domain of interest. Another reason to keep 20 items is to make the SCI convenient to administer within 20 min, which is less than half of a regular class period. During our recruitment of participants, we noted that most of the instructors were not willing to sacrifice an entire class period for data

collection but were more agreeable to administer the test if it takes only half of a class period or less. The shorter version also should produce less assessment fatigue among students.

Instructors' Feedback on SCI-20Q

We adopted a widely used method described in nursing education assessment (Polit & Beck, 2006) for the content validity study. A national survey of organic chemistry instructors was used to collect feedback on the 20-item instrument (SCI-20Q). The survey also contained an invitation to participate in the data collection by administering the final version of SCI at their schools. The survey was sent to 2,756 instructors. Their emails were retrieved from the institutions' websites. Each instructor received an email that described the purpose of the study and the nature of data collected. The electronic link in the survey led to the Qualtrics website. An informed non-signature consent form preceded the survey. It was followed by a screening question included to ensure that the participant was teaching or had been teaching a two-semester (or equivalent) organic chemistry course within the previous five years. After the screening question, the participants were presented with each of the 20 SCI items, accompanied by a question regarding the relevance of the item, and an open-ended feedback question (see Figure 5.6).

Relevancy Question

How relevant is this question with regard to your organic chemistry instruction?
- Not relevant
- Somewhat relevant
- Quite relevant
- Highly relevant

Feedback Question

If you have any feedback about this question, provide it in the box below. This may include comments on accuracy of scientific content, clarity of wording or drawings, and importance of the content.

Figure 5.6. Questions that accompanied each item of the SCI-20Q item on the feedback survey sent to organic chemistry instructors.

Each of the 20 items was presented in a random order to ensure equal coverage of the questions in case participants were to terminate the survey before completing it. Participants were also asked to select correct responses for all of the SCI-20Q questions. Following the completion of all of the SCI questions, participants were asked to provide general feedback on the instrument and if they are willing to participate in the field-testing. A demographic section was placed at the end of the survey and included questions about the highest degree in chemistry offered at their institution, their years of teaching experience, the highest degree possessed by the participant, and their primary area of expertise.

We received responses from 251 (response rate 9%) participants. The demographic characteristics of the sample are given in supplemental material (Table 5.10, survey 2). The demographic characteristics of participants for the content validity

study are fairly similar to those instructors who had completed the Stereochemistry Instruction Survey.

The results of the content validity study are presented in Table 5.3, along with the percentages of correct answers given by participants. The content validity indices were calculated as a proportion of participants who select "Quite relevant" and "Highly relevant" on the relevancy question.

Table 5.3.

Content Validity Indices and Percentages of Correct Answers Given by Instructors for the SCI-20Q Questions.

Q #	Content validity index, %	Correct answers, %
Q1	92	98
Q2	93	97
Q3	89	87
Q4	71	95*
Q5	72	97
Q6	89	100
Q7	89	99
Q8	94	97
Q9	95	92
Q10	87	100
Q11	87	92
Q12	86	97
Q13	93	85
Q14	77	92
Q15	79	85
Q16	54	98
Q17	74	98
Q18	93	95
Q19	65	96
Q20	63	93

* The item contained two correct options. The given percentage is a sum of percentages of instructors who selected both alternatives. The question was revised in the SCI-20Q version to ensure that only one answer is correct.

Overall, we received 674 text comments addressing items individually and as a set. The comments varied in the length and the amount of information and level of reflection. Of these comments, 119 (18%) were used in the revision process. The feedback was used to modify the wording and pictorial representation of three items to ensure clarity and unambiguity of the correct option. Minor grammatical changes were applied to five other items. We also analyzed responses to the SCI questions from the instructors. Most of the items were answered correctly by 95-100% of the participants. One item was ambiguous and contained two options that could be considered correct. The feedback given by instructors was also used to revise the graphical information in some of the items' stems to ensure scientific correctness of the question.

We received responses from 63 participants indicating their willingness to administer the SCI to their classes. Contact was made with these participants, and an IRB approval was sought from the instructors' institution. Data, collected from 17 diverse institutions, are described in the following section.

Multi-institute Data Collection with SCI-20Q

The final version of the Stereochemistry Concept Inventory was administered at 17 institutions. The participating institutions differed in type, size, setting, and location (see Table 5.10 of supplemental material). At three institutions, the SCI was administered as the paper-and-pencil version, while at 13 institutions it was administered as the online form. At one site, both electronic and paper-and-pencil versions were administered to different sections of an organic chemistry class that were taught by different instructors. All participating instructors were asked to review the set of questions of the Stereochemistry Concept Inventory prior to giving it to students to ensure the instructor

had covered the content of all items. The student participants from all sites, with the exception of two, were enrolled in the first semester of the two-semester organic chemistry course sequence. Participants at sites #6 and #9 were enrolled in a one-semester organic chemistry and a graduate organic chemistry course, respectively.

It was emphasized both during the recruitment phase and on the informed consent form that the purpose of the SCI is not to give students' a grade but rather to collect information about their performance. Instructors at three sites offered their students a small point incentive based on completion of the SCI (as determined by a print-out of the last page of the electronic version of the SCI). A paper-and-pencil version was administered in class, while the link to an online version was forwarded to students by their instructors with the information about the purpose and nature of the data collected.

The paper-and-pencil version (4 pages, 5 questions per page) was administered, within three weeks after the stereochemistry chapter was covered. Students who took the online version (one question per page) also completed the inventory within three weeks of receiving classroom instruction on stereochemistry. Demographic questions were placed at the end of both forms of the SCI. Information about major and gender was collected for both paper-and-pencil and online versions. The online version also asked about the school where participants were enrolled.

A total of 558 student responses were collected. After removing 102 incomplete tests and 17 tests without a signed consent form, a total of 439 student responses were analyzed. Most of the incomplete tests were from participants who did not finish the online version of the SCI. Ideally, percentages of missing responses should be compared across items to identify items that appear confusing for students. However, percentages of

missing responses were higher for the items appearing later on the test. For the item analysis, reliability, and correlation to other variables, only complete tests were included.

Item Analysis of SCI-20Q

For the purpose of data analysis, all complete responses from different sites were combined. There may be differences between participants from different institutions; however, the purpose of this study is not to identify these differences, but rather to create an assessment tool that can be used by researchers and practitioners. In addition, sample sizes from some institutions were small and represented a small fraction of a class, which does not allow legitimate comparison, and statistical tests that allow this comparison are underpowered. Also, the mean scores for the sites #11, #14, and #17, which had low response rates, may include students that are different from the overall population. We did not find significant differences ($p = .843$) between paper-and-pencil ($M = 11.27$, $SD = 3.52$) and online versions of the test ($M = 11.20$, $SD = 3.59$), which suggests equivalency of the forms. An effect size of 0.02 is indicative of an undetectable difference between students' performance on two modes of test delivery.

Difficulty and discrimination for the combined sample are presented in Table 5.4. For comparison with the pilot version (SCI-30Q), we presented a scatterplot of difficulty and discrimination values in Figure 5.3. As one can see, there was improvement from the pilot version in terms of difficulty and discrimination values fitting within the recommended ranges. Table 5.4 also presents an item analysis at the distractor level. For a proper functioning item, discrimination indices for distracters should be negative, indicating that students with lower ability are more likely to chose a distracter than students with higher ability. With the exception of one item (Q10), all distracters have

negative discrimination indices as expected. The distracter C for item 10 has a discrimination value of 0.01, which is indicative that there is almost no difference in the proportions of higher and lower achieving students who select this option.

Table 5.4.

Difficulty (P) and Discrimination (D) Indices for Each Response Option for the SCI-20Q items.

Item	Option	*P*	*D*
Q1	A	.05	–.11
	B	.07	–.20
	C	**.64**	**.28**
	D	.24	–.14
Q2	A	.11	–.17
	B	.24	–.22
	C	.03	–.23
	D	**.62**	**.38**
Q3	A	.04	–.24
	B	**.43**	**.29**
	C	.52	–.19
Q4	A	.10	–.17
	B	.07	–.17
	C	**.41**	**.26**
	D	.43	–.08
Q5	**A**	**.66**	**.39**
	B	.06	–.10
	C	.17	–.21
	D	.11	–.26
Q6	**A**	**.65**	**.43**
	B	.24	–.33
	C	.11	–.19
Q7	A	.13	–.29
	B	.04	–.10
	C	**.73**	**.39**
	D	.10	–.18
Q8	A	.07	–.22
	B	.11	–.27
	C	**.77**	**.45**
	D	.05	–.24
Q9	A	.09	–.18
	B	.13	–.20
	C	**.73**	**.32**
	D	.05	–.11
Q10	A	.26	–.37
	B	**.54**	**.33**
	C	.19	.01

Item	Option	*P*	*D*
Q11	**A**	**.42**	**.44**
	B	.11	–.26
	C	.47	–.27
Q12	A	.06	–.16
	B	**.72**	**.28**
	C	.21	–.21
Q13	A	.16	–.25
	B	**.62**	**.43**
	C	.22	–.28
Q14	A	.29	–.14
	B	.19	–.13
	C	**.33**	**.30**
	D	.20	–.06
Q15	**A**	**.49**	**.29**
	B	.45	–.25
	C	.07	–.08
Q16	A	.38	–.27
	B	**.38**	**.39**
	C	.24	–.14
Q17	A	.09	–.16
	B	.56	–.37
	C	**.35**	**.48**
Q18	A	.11	–.12
	B	.38	–.26
	C	.03	–.20
	D	**.49**	**.40**
Q19	**A**	**.63**	**.46**
	B	.19	–.34
	C	.18	–.23
Q20	A	.18	–.34
	B	**.62**	**.48**
	C	.06	–.16
	D	.14	–.19

The mean of the SCI-20Q scores was 11.22 (*SD* = 3.56), constituting 56.1%. The scores range from 2 (10% correct, *N* = 1) to 20 (100% correct, *N* = 3). The values of skewness (0.08) and kurtosis (–0.48) fall within the ±1 range indicating normality of the distribution of total scores. We have summarized distribution of the SCI scores obtained as a result of the multi-site administration in Table 5.5. Table 5.5 can be used by end-users of the SCI to compare results from their classes to a sample used in this study. We define a percentile rank here as the percentage of scores that fall both *at* and *below* a given score.

Table 5.5.

SCI Scores with Corresponding Frequencies and Percentile Ranks.

SCI score	Frequency	Cumulative frequency	Percentile rank
1	0	0	0
2	1	1	0
3	2	3	1
4	7	10	2
5	13	23	5
6	18	41	9
7	24	65	15
8	39	104	24
9	42	146	33
10	40	186	42
11	56	242	55
12	41	283	65
13	34	317	72
14	37	354	81
15	29	383	87
16	23	406	93
17	13	419	95
18	13	432	98
19	4	436	99
20	3	439	100

Discussion of Reliability and Validity

Reliability of SCI-30Q and SCI-20Q

Traditional measures of reliability such as Cronbach's alpha may not be the best measure of internal consistency. Several review papers address the limitation of the Cronbach's alpha coefficient. High values may be indicative of redundant items (Adams & Wieman, 2011), while lower values may present fragmented knowledge that students possess (Bretz & McClary, 2015). Since concept inventories often assess the ideas that are incomplete and fragmented, a Cronbach's coefficient is usually lower due to the unlikelihood of participants being the same or similar. Indeed, recently published concept inventories show quite low reliability coefficients: .28 in the diagnostic instrument of understanding of electrochemical cells (Loh et al., 2014); .26–.46 in the Thermochemistry Concept Inventory (Wren & Barbera, 2014); .41 in the ACID I (McClary & Bretz, 2012). However, there are concept inventories with appropriate reliability, so probably the argument about fragmented ideas selectively applies to some of the assessment instruments, but not the others. Several alternative ways to assess the reliability have been suggested, for example, use of the test-retest reliability. In order to address all of the concerns that have been raised about reliability, in addition to the internal consistency coefficient, we have estimated reliability based on temporal stability and reliability of alternate forms (Thorndike & Thorndike-Christ, 2011).

For the SCI-30Q pilot, we observed a low (KR-21 of .40) reliability coefficients. Rerunning the analysis without 12 problematic (both from psychometrically and from response process study) items yielded a higher reliability coefficient of 0.56 for the remaining set of 18 items from the SCI-30Q. It is worth noting here, that these reliability

coefficients (.40 and .56) are obtained from the same sample, thus the difference between them can be due to a sampling error.

In the development phase, the SCI-30Q was administered twice to the same student participants within a 17-day interval. From the two datasets, we obtained matched scores for 47 participants. The means of the scores obtained on the first administration ($M = 11.96$, $SD = 3.14$) did not differ significantly ($p = .185$) from the scores on the second administration ($M = 12.49$, $SD = 4.04$), although an effect size of $d = 0.15$ indicates that there might be a small gain in the scores that was not detected due to an underpowered sample. Gains are usually observed in the test-retest conditions (Barbera, 2013; Mulford & Robinson, 2002). A correlation between scores obtained on two administrations was found to be 0.74 ($p = .000$) which indicates similar performance of the participants on both administrations. Students did not receive any stereochemistry instruction nor the correct answers for the SCI questions in the time period elapsed between the two administrations.

For the final, SCI-20Q version, an overall reliability as measured by the KR-21 coefficient is 0.64. An overall reliability as measured by Cronbach's alpha coefficient is 0.67. We have also analyzed reliability for the gender and major subgroups, as well as for modes of administration (online and paper-and-pencil). The reliability coefficients for the individual subgroups are given in Table 5.6. As can be noted from Table 5.6, the higher values of reliability generally are associated with higher scores. The exception is the "Others" group that is composed of majors that were not listed in the demographic form and graduate students. Connection of reliability to overall test performance support the claim proposed by Bretz and McClary (2015) that lower reliability coefficients represent

fragmented inconsistent knowledge. Possibly, higher levels of ability of chemistry majors are indicative of more coherent knowledge structures that result in higher reliability coefficients.

Table 5.6.

Sample Sizes, Reliability Coefficients, Means and Standard Deviations of the SCI-20Q Scores for Individual Groups and Modes of Administration.

	N	Cronbach's alpha	Mean	*SD*
Modality				
Online	273	.69	11.20	3.59
Paper	166	.66	11.27	3.52
Gender				
Male	187	.70	11.84	3.65
Female	252	.64	10.77	3.43
Major				
Chemistry	95	.73	12.01	3.78
Physical/natural sciences	280	.65	10.83	3.47
Social sciences/humanities	22	.71	11.27	3.84
Others	42	.56	12.10	3.15
All complete cases combined	439	.67	11.22	3.56

To obtain additional measures of reliability, we developed an alternative form (form B) for the SCI-20Q. Each question was replicated with a slight change of the structures and compounds (such as changing CH_3 to CH_3CH_2 or Br to OH) that are used in the stems and response options. While it is not always possible to create an alternative question with the same difficulty, we would like to note that two questions on the alternative version were more difficult since they included an additional step. Both versions of the SCI were administered electronically to a sample of 110 students from

seven different institutions. The questions were presented to the participants in a random order. A correlation between scores obtained on two versions was found to be .785 (p = .000). It is worth noting that the scores (SCI-20Q: M = 10.99, SD = 3.62; SCI-20Q-B: M = 9.02, SD = 3.42) obtained by participants on two versions of the SCI are significantly different (p = .000; d = 0.56). The reliability coefficients (KR-21) for SCI-20Q and SCI2-20Q-B are .65 and .60, respectively, if computed separately, and the overall coefficient of the combined version is .79. To our knowledge, the SCI concept inventory was the first concept inventory in chemistry for which reliability evidence of alternate forms was collected.

Evidence Supporting Content

A table of specifications (also known as a test blueprint) is recommended to be constructed before creating an assessment instrument (Nitko & Brookhart, 2011). This approach ensures adequate coverage of different content areas and cognitive tasks that are expected from the examinees. The test blueprint is usually presented as a matrix of questions that are aligned towards topics and level of objective. The Bloom's taxonomy or revised Bloom's taxonomy is often used as a basis for the levels of educational objectives associated with certain topics. However, in chemistry as well as in other natural sciences, there is very little evidence that educational objectives align with Bloom's levels as noted by Cooper and Klymkowsky (2013).

When constructing a table of specifications, we decided to use concrete educational objectives that were endorsed by the majority of the organic chemistry instructors that participated in the national survey of stereochemistry instruction. Only the objectives that were endorsed by more than 75% percent of practitioners were included in

the blueprint. The detailed blueprint is presented in Table 5.7. We used several questions per topic to ensure that we addressed the content on different levels of representations. Two organic chemistry instructors were asked to comment on the questions and their alignment on the blueprint. According to their feedback, the set of the questions aligns with the blueprint and tests the concepts appropriately.

Table 5.7.

The Test Blueprint for the SCI-20Q Version.

Topic	Learning Objective	Question #s
Chirality	Recognize chiral molecules	1, 2, 3
	Recognize stereogenic centers	4
Stereoisomers	Determine the relationship of two given structures	10, 11, 13, 15
	Calculate total number of stereoisomers	5
Enantiomers	Recognize enantiomers	10, 11
	Know that enantiomers have identical properties	12, 16
Diastereomers	Recognize diastereomers	13
	Know that diastereomers have different properties	14, 16
Meso compounds	Recognize meso	15
	Know properties of meso	16
Optical activity	Know that optical activity is a property of chiral molecules	16, 17
	Identify relationship between structure and optical activity	16, 17
R, S nomenclature	Know the CIP rules	6, 7, 8
	Assign *R* and *S* descriptors to stereogenic centers	8, 9
Projections	Understand Fischer notations	19, 20
	Understand Newman notations	17

The distractors for multiple-choice questions come from the qualitative study of students' incorrect ideas in organic chemistry. Table 5.8 represents the major incorrect ideas that were found in the qualitative study aligned by SCI topics and corresponding percentages of students who possess these ideas. We plan to report incorrect ideas that

were revealed based on analysis of students' work, cognitive interviews, analysis of existing literature, and the sharing of experiences between instructors, as a separate manuscript. Table 5.8 includes several succinctly expressed predominant incorrect ideas that were used for the development of the SCI items. We present both percentages of students that expressed incorrect ideas in interviews and percentages of students who selected distracters on the SCI-20Q that represent these incorrect ideas.

Table 5.8.

Percentages of Students Who Possess a Certain Incorrect Idea as Detected by the SCI and in a Qualitative study.

Topic	Incorrect idea	Percentages of students selecting option(s) corresponding to the incorrect idea (*N* = 439), %	Percentage of students having this incorrect idea from a qualitative study (*N* = 24), %
Chirality	Defining chirality of a molecule by the presence of an atom with four different substituents	24	92
Optical activity	Equating specific rotation to *R/S* configuration	56	54
R/S nomenclature	Ranking substituents for *R/S* assignment based on irrelevant characteristics	31*	54
Projections	Considering that hydrogen is always facing backwards irrespective of spatial arrangement of a molecule	14	13

* Determined as the average of responses to two questions (#6 and #7).

As can be seen from Table 5.8, the percentages of students holding specific incorrect ideas about optical activity, *R/S* nomenclature and projections are fairly similar as detected by interviews and the SCI questions. Regarding the concept of chirality, one of the potential explanations for the discrepancy between the percent of students selecting

an option representing an incorrect idea on the SCI and those students from the qualitative study can be as follows. In interviews students were asked what makes a molecule chiral, and most of them (92%) made a statement that a molecule is chiral if it has an atom with four different substituents, rather than defining chirality as non-superimposable mirror images. This statement is not entirely incorrect, because for the molecules that contain only one atom with four different substituents, this is a correct statement. Most often students were providing examples with one tetra-substituted carbon as a chiral molecule. In the SCI we have used either a set of molecules from which they select chiral ones or gave a molecule and asked students to select a statement that describes its chirality.

The SCI-20Q version contained two questions that address students' familiarity with the Cahn-Ingold-Prelog rules. Figure 5.7 represents these items with distractors which address the same set of incorrect ideas, i.e., ranking of the substituents is based on electronegativity or their size.

Item #6 (SCI-20Q)

Which of the following substituents has the highest rank according to the Cahn-Ingold-Prelog rules?

$–OC(CH_3)_3$, –F, or –Cl

A. –Cl, because it has the highest atomic number.
B. –F, because it is the most electronegative element.
C. $–OC(CH_3)_3$, because it is the largest by size.

Item #7 (SCI-20Q)

Which of the following substituents has the higher priority according to the Cahn-Ingold-Prelog rules?

$–NH_2$ or –Br

A. $–NH_2$ because N is more electronegative than Br.
B. $–NH_2$ because it has more atoms than Br.
C. –Br because it has a higher atomic number than N.
D. –Br because it is larger than N.

Figure 5.7. Two questions from the SCI that address the same content area (ranking assignment of substituents) and the same incorrect ideas (ranking assignment based on electronegativity or size).

In order to assess consistency of ideas measured by these two questions, we composed a contingency table (Table 5.9) to examine students' responses for these two questions. A contingency table is a matrix format that displays the frequency distribution of the variables. The observed count represents the number of participants who had selected both options that correspond to the cell. The expected count represents what count will appear in the cell if the two variables are unrelated. As it can be seen from the contingency table, an observed count is always higher than expected in the cells that represent consistent answers that are correct (Item 6A and Item 7C) or answers that

represent incorrect ideas of ranking based on electronegativity (Item 6B and Item 7A) or ranking based on size (Item 6C and Item 7BD). This finding can serve as evidence of decision consistency among students, since 74% of respondents select alternatives on both options that share the same content.

Table 5.9.

A Contingency Table Representing Crosstabulation of Items 6 and 7.

Item 6		Item 7				Total
		A	B	C	D	
A	Count	8	7	**274**	13	302
	Expected Count	44.6	13.8	**211.6**	32.0	302
B	Count	**58**	4	38	23	123
	Expected Count	**18.2**	5.6	86.2	13.0	123
C	Count	5	**11**	25	**15**	56
	Expected Count	8.3	**2.6**	39.2	**5.9**	56
Total	Count	71	22	337	51	481
	Expected Count	71	22	337	51	481

Note: Cells containing bold numbers represent responses that are consistently chosen on both questions.

Evidence Based on the Relationship to Other Variables

As pointed by Arjoon et al. (2013), collective evidence on relationship of scores from a certain instrument can be instrumental in establishing a nomological network for a construct of interest. In this section, we report relationships of the SCI-20Q scores with demographic variables for all participants and with grades obtained in the organic chemistry course at site #1.

Along with responses to the SCI-20Q version, we collected basic demographic information. The overall scores on the SCI for the demographic groups are presented in Table 5.8. Means for males (M = 11.84, SD = 3.65) and females (M = 10.77, SD = 3.43) differed significantly (p = .002, d = 0.30). A comparison of chemistry majors (M = 12.01, SD = 3.78, N = 95) with the rest of the population (M = 11.01, SD = 3.46, N = 344) gives a marginally significant p-value of .02 and an effect size of 0.28. As Barbera and VandenPlas (2011) noted, an instrument meant to measure some chemistry knowledge should be able to discriminate between theoretically different groups of students. To a certain extent, higher performance of chemistry majors substantiates the claim that supports concurrent validity. A discrepant performance of males and females is in concordance with most research findings. However, in the absence of a meta-analysis on gender performances in chemistry we report this result without any substantial claims regarding effect of gender on scores. Meta analytic studies of gender performances are available, for example, for mathematical achievement (Hyde, Fennema, & Lamon, 1990), where a role of moderating variables, such as age, education level, year, demographic characteristics, complexity of tasks, was thoroughly investigated to explain observed discrepancies.

At one institution (#1 in Table 5.11 in the supplemental material), the SCI was administered with a midterm exam composed by the instructor. Approximately 20% of questions on this exam were devoted to stereochemistry. A matched dataset of 26 students was examined for students who successfully (grade C or higher) completed the course. The correlation of the scores obtained on the SCI with the scores received on the midterm exam was found to be .27 and not significant (p = .180). We also examined the correlation of the SCI scores with the final grade in the course. There was a relatively high correlation of .54 and it was significant (p = .005). To some extent, these correlations serve as supporting evidence for the convergent and predictive validity (Barbera & VandenPlas, 2011); however, a concept inventory that tests a narrow and specific area of knowledge is unlikely to be related to the overall grade in the organic chemistry course due to the fact that the scores on the SCI and overall grade represent related but different domains of knowledge, skills, and behavior. A correlation with the midterm exam grades is not expected to be high for the same reason. A higher correlation with the final exam may suggest that the SCI tests a representative set of skills that are essential to succeed in an organic chemistry course. Scores obtained on the SCI were not counted towards either midterm or final exams in this particular case.

Implications and Conclusions

Research Implications

The developed instrument can be used as a measure of cognitive outcomes for experimental and quasi-experimental designs that involve some type of instructional intervention aimed to enhance stereochemistry comprehension. The scores from the SCI can also be used as a measure of prior ability for more complex designs, for example, Solomons' four-group design or counterbalanced measures design (Gall et al., 2003). These designs include a pretest that is used as a covariate to account for inequivalence of the groups that are exposed to treatment or kept as controls. The scores obtained from the SCI given as a pretest may be used to determine equivalency or account for inequivalence in groups. The characteristics of the national sample presented in this study can serve as a comparison for the effectiveness of educational experiments that do not have a control group. However, this comparison should be used with caution due to limitations of the sample in this study. The scores from the SCI administered as pre- and post-tests can also be used for determining teaching effectiveness.

Opinions of stakeholders who are organic chemistry instructors were accounted for both in the development and validation phases. The instrument is more likely to be used by practitioners if they have a sense of ownership or contribution to the content. The feedback that was provided by practitioners made invaluable contributions to the development and revision of the content. Most of our respondents provided objective, unbiased suggestions that helped to develop better functioning items.

Teaching Implications

For diagnostic purposes, organic chemistry instructors who wish to quickly obtain data that can reveal students' understanding of stereochemistry can use the SCI as a formative assessment. Instructors who teach upper division courses such as biochemistry or advanced organic chemistry may use this instrument to determine if students need remedial instruction on stereochemistry.

Instructors can also identify areas that appear problematic for students and choose a teaching approach to help students attain a better understanding of stereochemistry concepts. A variety of teaching approaches have been reported, including card games (Costa, 2007), self-directed lessons (Cody et al., 2012), and demonstration experiments (Schwartz, Lepore, Morneau, & Barratt, 2011). Several teaching approaches combine modeling activities with laboratory experiments (Bandaranayake, 1980; Clausen, 2011). These activities can potentially help students to overcome confusion of absolute configuration with optical rotation, one of the most prevalent incorrect ideas that can be detected with the SCI. The efficiency of these activities can be tested by administering the SCI as a pre- and post-test and comparing not only the overall results but percentages of correct answers for the specific items that are testing concepts covered in activities.

Limitations of the Study

Being a multiple-choice test, the SCI has all of the limitations that these types of questions have, such as guessing and simple recall of relevant knowledge. Another limitation that the SCI has is limited coverage of stereochemistry concepts. The content that is covered in the SCI refers only to static stereochemistry (Eliel et al., 1994) which deals with the spatial arrangement of molecules and their properties. Dynamic

stereochemistry which deals with stereochemistry of reactions is not covered in the SCI. Development of another tool that assesses stereochemistry concepts within their applications to reactions can overcome this deficiency. The voluntary nature of participation for both versions of the SCI constitutes another limitation of the inferences that were drawn from the data. For sites with limited participation, it can be assumed that students with a higher desire to learn and explore attempted answering the SCI questions. For multiple sites where the SCI was administered, differences may come both from variations in curriculum and instruction, as well as individual differences of populations in reasoning ability, mental capacity, or spatial ability.

Evidence of response process validity was collected during the development stage to refine the quality of the questions, but not for the final SCI-20Q version after some questions were modified. As a lesser evidence of response process validity of the SCI-20Q instrument, we examined all students' copies of the paper-and-pencil version (students were allowed to write on the test) and did not find any drawings that represent misunderstanding of content intended to be covered by corresponding SCI questions. On the contrary, most of the drawings left on students' copies represent either incorrect ideas or proficiency of stereochemistry knowledge.

Summary

The number of concept inventories available in chemistry education have increased rapidly over the past decade. This paper presents a new instrument for assessing students' understanding of stereochemistry. Our goal was to provide and test a rigorous methodology that was used both in the development and content validation phase of the instrument development. The Stereochemistry Concept Inventory features

multiple-choice questions containing distracters that reflect common incorrect ideas students possess about stereochemistry. The instrument is composed of 20 questions that cover diverse aspects of stereochemistry such as chirality, stereoisomers, optical activity, nomenclature, and various types of projections. Psychometric characteristics of the SCI items are based on multi-site samples and show sufficient evidence for reliability and stability of the scores measured in several distinct ways. Instructors who are interested in administering the SCI to their students may contact the authors.

Supplemental Information for Article 2

Table 5.10.

Demographics of Instructors who Completed the Stereochemistry Instruction Survey (Survey 1) and Instructors Who Participated in the Content Validation Study (Survey 2).

	Stereochemistry Instruction Survey (Survey 1)	Content validity study (Survey 2)
Surveys sent	*N* = 1,028	*N* = 2,756
Surveys started	*N* = 249	*N* = 346
Surveys completed	*N* = 219	*N* = 251
Highest degree in chemistry offered at the institution		
Associate's Degree	10%	5%
Bachelor's Degree	28%	56%
Master’s Degree	10%	8%
Doctorate Degree	53%	31%
Teaching experience		
Mean, years	18.3	17.1
SD, years	13.0	12.4
Highest degree in chemistry received by the instructor		
Master's Degree	5%	4%
Doctorate Degree	95%	95%
Instructor’s primary area of expertise		
Analytical Chemistry	1%	1%
Biological Chemistry	4%	6%
Inorganic Chemistry	3%	4 %
Organic Chemistry	90%	86%
Physical Chemistry	2%	3%

Note: Percentages may not add to 100% due to rounding.

Table 5.11.

Description of the Institutions According to the Carnegie Classification of Institutions of Higher Education.

School code	Profile	Size	Setting	Basic classification	Region	Sample size	Mode
1	FT4–S–LTI	M4	HR	Master's L	South West	64	Both
2	FT4–MS–LTI	S4	HR	Bac/A&S	Mid-West	24	Paper
3	FT4–MS–LTI	S4	HR	Bac/A&S	Mid-West	70	Paper
4	PT2	VL2	–	Assoc/Pub-U-MC	West	16	Paper
5	FT4–S–LTI	L4	R	RU/H	Mid-West	9	Online
6	FT4–S–HTI	M4	R	Master's L	Mid-West	20	Online
7	FT4–S–LTI	L4	NR	Bac/Diverse	North West	35	Online
8	FT4–S–HTI	L4	R	RU/H	South East	10	Online
9	FT4–S–HTI	L4	R	RU/H	South West	13	Online
10	MFT4–S–HTI	M4	NR	Master's L	South East	30	Online
11	FT4–MS–LTI	M4	HR	Master's M	Mid-West	6	Online
12	FT4–S–HTI	L4	R	DRU	South West	53	Online
13	FT4–S–HTI	S4	R	Master's L	North East	21	Online
14	FT4–S–HTI	L4	R	Master's L	Mid-West	4	Online
15	FT4–S–HTI	S4	HR	Master's M	North West	26	Online
16	FT4–S–HTI	M4	R	Master's M	Mid-West	30	Online
17	FT4–S–HTI	M4	HR	Master's L	South West	8	Online

Note: Profile according to the Carnegie classification (the definitions of categories and the methodology of attribution can be found on http://carnegieclassifications.iu.edu/): FT4 – Full-time four-year; PT2 – Higher part-time two-year; MFT4 – Medium full-time four-year; S – selective; MS – more selective; LTI – lower transfer-in; HTI – lower transfer-in. Size: S4 – Small four-year; M4 – Medium four-year; L4 – Large four-year; VL2 – Very large two-year. Setting: R – primarily residential; HR - highly residential; NR - primarily nonresidential. Basic classification: Bac/A&S – Baccalaureate Colleges-Arts & Sciences; Assoc/Pub-U-MC – Associate's-Public Urban-serving Multicampus; Bac/Diverse – Baccalaureate Colleges-Diverse Fields; RU/H – Research Universities; Master's M: Master's Colleges and Universities (medium programs); Master's L: Master's Colleges and Universities (large programs); DRU – Doctoral/Research Universities.

CHAPTER VI

CONCLUSIONS, IMPLICATIONS, AND FUTURE RESEARCH

Conclusions

In this study, the following research questions were answered:

Q1 What stereochemistry topics do organic chemistry instructors consider important?

Q2 What incorrect ideas do organic chemistry students hold regarding stereochemistry?

Q3 Can the Stereochemistry Concept Inventory produce reliable data and valid inferences for the assessment of important concepts of stereochemistry?

Chapter IV includes a qualitative study that addressed Q2. Expressed succinctly, these incorrect ideas are:

1. Ranking of substituents for *R/S* configuration is based on electronegativity.
2. Ranking of substituents for *R/S* configuration is based on size.
3. Hydrogen is always pointing backwards in Fischer projections.
4. Specific rotation is equated to *R/S* configuration.
5. Different biological activity of enantiomers is explained by opposite rotation of plane-polarized light.
6. Enantiomers are the same compound, just drawn from different perspectives.
7. Diastereomers are different representation of the same compound.
8. A carbon atom with four different substituents makes the whole molecule chiral.
9. Enantiomers have different physical properties.

10. Enantiomers can be separated by physical methods such as distillation.
11. A total number of stereoisomers is calculated using n^2 or $2n$ formula where n is a number of stereogenic centers.

The uncovered incorrect ideas apply to all three domains of chemistry knowledge as represented by the particulate, macroscopic, and symbolic level of the Johnstone's triangle. Students' incorrect ideas can depend on the method by which they were detected.

Chapter V contains answers to research questions Q1 and Q3. Instructors ranked the topics that relate to chirality and stereoisomers as the most important (Table 5.2). These topics served as a content foundation for the Stereochemistry Concept Inventory, which was reviewed by a national sample of organic chemistry instructors and tested on a national sample of organic chemistry students. All of the questions in the final version of the Stereochemistry Concept Inventory showed acceptable psychometric characteristics. Multiple measures of reliability showed high stability of the SCI scores across time and different forms with acceptable levels of internal consistency. Sufficient evidence for the content validity was obtained from the national sample of organic chemistry instructors.

The uniqueness of the development process of the Stereochemistry Concept Inventory is in the accounting for opinions of stakeholders throughout the development, validation, and revision phases. Another unique feature of the Stereochemistry Concept Inventory is a warrant of psychometric quality of the scores that was obtained from multiple measures of reliability.

Implications for Research and Practice

The results of the national survey of organic chemistry instructors (Stereochemistry Instruction Survey) can be used by developers of assessment and instructional materials. Since the outcome of this survey is a list of topics that are considered important by the organic chemistry instructors, this can serve as a guide to develop more relevant assessment questions or activities that will be in demand by practitioners. A study of incorrect ideas can inform practitioners what difficulties organic students encounter when studying organic chemistry. Being familiar with these incorrect ideas can help instructors suppress their development by emphasizing the topics that are known to cause difficulties. Both results of the Stereochemistry Instruction Survey and findings from a qualitative study can be used to create alternative tools that assess a variety of leaning outcomes that are related to stereochemistry as a whole or to specific aspects of stereochemistry. For example, an alternative version of the Stereochemistry Concept Inventory can be created with pictorial images of ball-and-stick models instead of wedge-and-dash, Fischer, or Newman projections.

For practitioners, the Stereochemistry Concept Inventory may be used as a brief formative assessment tool at the end of the lecture module on stereochemistry. Instructors may use this inventory to test their students and quickly decide if any additional clarification on certain concepts is needed.

The Stereochemistry Concept Inventory can be used as a measure of cognitive outcomes for experimental and quasi-experimental designs that involve some type of instructional intervention aimed to enhance stereochemistry comprehension. A variety of instructional strategies have been reported to create a student-centered environment

(Weimer, 2013). These techniques include, but not limited to, Peer-led Team Learning, cooperative learning, Just-in-Time Teaching, inquiry activities, active learning, flipped learning, and various combinations of previously mentioned methods. An assessment of these techniques targeted for enhancing stereochemistry understanding can be performed using the Stereochemistry Concept Inventory. These research studies can contribute to the body of evidence-based research literature. One of the most influential chemistry education researchers, Melanie Cooper (2007), reported that faculty who develop educational strategies often try to promote them even if there is only anecdotal evidence of their effectiveness. Quite often they do no engage in collecting evidence to support the effectiveness of their method of instruction. Use of tests and surveys constructed by researchers themselves without evidence is quite common in educational research. This gap can be closed with high-quality assessment tools for a variety of learning outcomes and studies that utilize these tools in educational settings.

Future Research

Spatial Ability and Working Memory

The study described herein includes basic validation steps to ensure content validity, predictive validity, and concurrent validity. However, the relationships of stereochemistry competence with spatial ability and working memory capacity remain unexplored. Students can incorrectly answer stereochemistry questions involving visualization of organic molecules not only from their lack of spatial ability, but also from their lack of working memory. Pribyl and Bodner (1987) explored the relationship of spatial ability and organic chemistry competence. These relationships can be explored

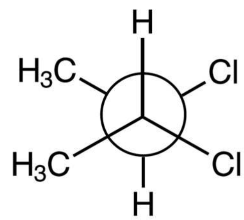
by means of structural equation modeling to estimate causal inferences. Spatial ability can be measured with the Paper Folding Test VZ2 or the Purdue Visualization of Rotations Test (Bodner & Guay, 1997). As the Stereochemistry Concept Inventory includes both questions that involve visual and textual questions, it adds another layer of complexity to this study. It can be hypothesized that questions containing graphical information may be more related to spatial ability measures, while questions that are textual are more related to working memory.

Chirality Concept Inventory

During this study, it became obvious that even narrowing the realm of stereochemistry to static concepts, the topic is too broad to be tested with a single measurement tool. However, one aspect of stereochemistry learning, understanding chirality, can be measured with a set of questions that can be aligned with the cognition model proposed by Brown and Wilson (2011). Being a fundamental concept of stereochemistry, chirality constitutes one of the major difficulties that students experience in organic chemistry. A diagnostic tool provides more interpretable results when it is highly focused on a specific topic. A set of questions asking if the presented molecule is chiral can constitute a Chirality Concept Inventory. These questions can be arranged in order of progressing difficulty as determined by the number of stereogenic centers which will allow for analyzing data using the Guttman scale (Bond & Fox, 2001). An alternative version of the Chirality Concept Inventory can be composed of questions that are two-tiered and include both question and reason tier. An example of a two-tiered question is shown in Figure 6.1.

Question tier

Is the following molecule chiral?



A. Yes B. No

Reason tier

The reason I chose my answer for the question above is

A. The molecule mirror image cannot be overlaid on itself.
B. The molecule contains a carbon atom with four different substituents.
C. The molecule is a racemic compound.
D. The molecule contains a plane of symmetry.

Figure 6.1. A two-tiered question for a proposed Chirality Concept Inventory.

Items for the Chirality Concept Inventory can also be based on specific theoretical visual-perceptual skills outlined by Oliver-Hoyo and Sloan (2014), with the items grouped by content and the representation mode (ball-and-stick models, wedge-dash projections, Fischer projections).

Item Response Theory

Analysis based on the Rasch or Item Response Theory (IRT) allows deeper insight into item functioning and how ability is linked to solving the item. A three-parameter IRT model allows differentiating between students who possess correct knowledge and those who have a correct response by guessing. As a preliminary step, the data from the SCI-20Q were examined by dichotomous Rasch analysis. The dataset was found to be satisfactory for both assumptions of the Rasch model as described by Bond and Fox (2001): unidimensionality (eigenvalue of less than 2 for the principal components analysis) and local independence (no residuals that correlate higher than 0.20) The results from Rasch analysis are summarized in Table 6.1.

Table 6.1.

Properties of the SCI-20Q Scale as Estimated by the Rasch Model.

Statistic	Items	Persons
Observed Variance	0.4509	0.8214
Observed *SD*	0.6715	0.9063
Mean Square Error	0.0117	0.2716
Root Mean Squared Error	0.1083	0.5212
Adjusted Variance	0.4392	0.5498
Adjusted *SD*	0.6627	0.7415
Separation Index	6.1179	1.4228
Number of Strata	8.4906	2.2304
Reliability	0.9740	0.6694

An assessment of fit (Bond & Fox, 2001) of students' responses to the Rasch model focuses on identifying observations that are outliers to the data set or unexpected response patterns in the dataset. The infit and outfit MNSQ statistics (see Table 6.2) are both well within the range recommended by Bond and Fox (2001) of 1.00 ± 0.50. The difficulty measures (*b*-parameter) generally range from –4.0 to 4.0, the more negative value indicating the easier the item. For the SCI-20Q, difficulty measures ranged from –1.10 to 1.17, which does not indicate the presence of any items that are extremely difficult or easy.

Table 6.2.

Item-level Psychometric Estimates for Rasch Model Analysis.

Question	Difficulty measure	Infit MNSQ	Outfit MNSQ
Q1	–0.38	1.08	1.07
Q2	–0.26	0.99	0.95
Q3	0.63	1.08	1.15
Q4	0.76	1.13	1.22
Q5	–0.46	0.97	0.94
Q6	–0.44	0.94	0.93
Q7	–0.83	0.96	0.90
Q8	–1.10	0.90	0.77
Q9	–0.83	1.01	1.07
Q10	0.11	1.05	1.03
Q11	0.67	0.95	0.94
Q12	–0.82	1.05	1.12
Q13	–0.24	0.95	0.95
Q14	1.17	1.07	1.24
Q15	0.36	1.09	1.10
Q16	0.87	1.01	0.98
Q17	1.03	0.91	0.87
Q18	0.38	0.99	1.02
Q19	–0.33	0.92	0.86
Q20	–0.28	0.90	0.87

Assessment of Faculty Workshops

Chemistry Collaborations, Workshops, and Community of Scholars (ccwcs.org) is an initiative sponsored by the National Science Foundation intended to facilitate faculty development workshops, including the Active Learning in Organic Chemistry (ALOC) workshop. The ALOC workshop is offered every year which provides access to large samples of students. The assessment of the effectiveness of the workshops follows an iterative model, presented in Figure 6.2.

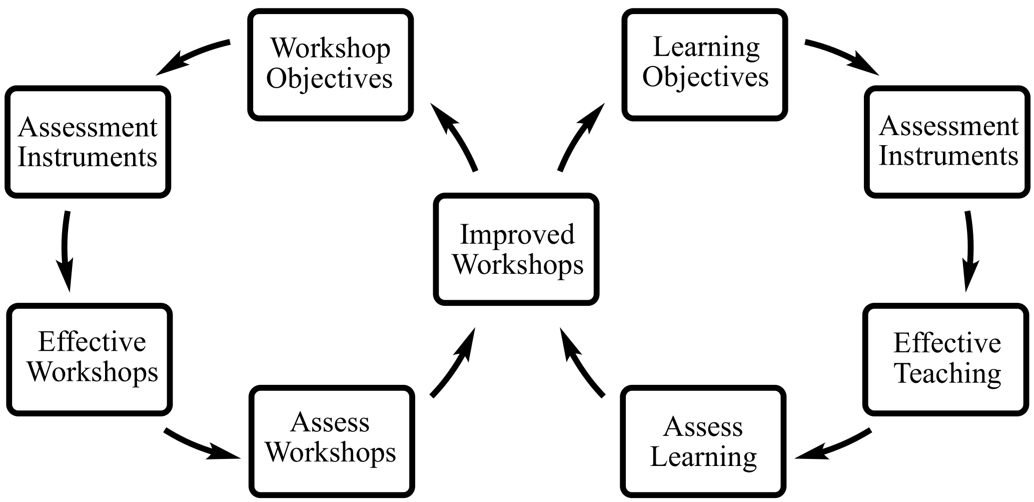


Figure 6.2. The chemistry CWCS assessment model (used with permission from Justin Houseknecht).

This assessment model includes both direct assessment of faculty participants and indirect assessment of their students. The Stereochemistry Concept Inventory, along with other assessment tools, will be used for the assessment of students enrolled in classes of the ALOC attendees. The data obtained from students enrolled in 25 institutions will be analyzed using Item Response Theory. As of June 2015, the data were collected from 135 students from seven institutions. This data will serve as a control comparison for future studies. These data were not included in this dissertation.

Affective Measures

Instructors may benefit from information about robustness of students incorrect ideas, whether students purposefully choose an answer representing a certain (correct or incorrect) idea or they simply guess. To measure students' confidence, a confidence tier can be added to each of the items in a concept inventory. Measurement of confidence can be performed in several ways. For example, in the development of a three-tier diagnostic test for assessment of misconceptions that relate to environmental science, a dichotomous confidence tier question was used (Arslan et al., 2012), while in the Redox Concept Inventory, Brandriet and Bretz (2014b) used a continuous scale to measure students' confidence.

Each SCI-20Q item can be given along with a confidence scale to assess students' affective construct. This procedure would provide additional insights not only into the prevalence of incorrect ideas, but also into their persistence. If used in an assessment battery, confidence measures would allow detecting a disequilibrium state of Piaget's cognitive development process.

An overall performance can be compared with overall interest in organic chemistry. For this purpose, an item measuring interest can be placed at the end of the SCI-20Q. A positive correlation of the SCI scores and interest scale strengthens validity evidence obtained for the SCI. Currently, these data are being collected for the students of the instructors who attend the ALOC workshop mentioned in the previous section. The version of the interest question is shown in Figure 6.3.

How interested are you in organic chemistry?

- o Very uninterested
- o Uninterested
- o Interested
- o Very interested

Figure 6.3. The interest question included in the electronic version of the SCI-20Q version.

As was shown in the study that reported the development of the instrument to measure understanding of nanoscience (Schönborn et al., 2015), the scores obtained from the interest question are valid predictors of an overall performance on the instrument.

References

Abimbola, I. O. (1988). The problem of terminology in the study of student conceptions in science. *Science Education*, *72*(2), 175–184.

Adadan, E., & Savasci, F. (2012). An analysis of 16–17-year-old students' understanding of solution chemistry concepts using a two-tier diagnostic instrument. *International Journal of Science Education*, *34*(4), 513–544.

Adams, W. K., & Wieman, C. E. (2011). Development and validation of instruments to measure learning of expert-like thinking. *International Journal of Science Education*, *33*(9), 1289–1312.

Al-Balushi, S. M., & Al-Hajri, S. H. (2014). Associating animations with concrete models to enhance students' comprehension of different visual representations in organic chemistry. *Chemistry Education Research and Practice*, *15*(1), 47–58.

Anderson, D. L., Fisher, K. M., & Norman, G. J. (2002). Development and evaluation of the conceptual inventory of natural selection. *Journal of Research in Science Teaching*, *39*(10), 952–978.

Arjoon, J. A., Xu, X., & Lewis, J. E. (2013). Understanding the state of the art for measurement in chemistry education research: Examining the psychometric evidence. *Journal of Chemical Education*, *90*(5), 536–545.

Arslan, H. O., Cigdemoglu, C., & Moseley, C. (2012). A three-tier diagnostic test to assess pre-service teachers' misconceptions about global warming, greenhouse effect, ozone layer depletion, and acid rain. *International Journal of Science Education*, *34*(11), 1667–1686.

Asirvatham, M. R. (2009). *Clickers in action: Increasing student participation in general chemistry*. New York, NY: W. W. Norton.

Aslanides, J. S., & Savage, C. M. (2013). The Relativity Concept Inventory: development, analysis and results. *Physical Review Special Topics - Physics Education Research*, *physics.ed*(9), 1–10.

Ausubel, D. P. (1960). The use of advance organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology*, *51*(5), 267–272.

Ausubel, D. P., Novak, J. D., & Hanesian, H. (1978). *Educational Psychology: A Cognitive View*. Holt McDougal.

Bailey, J. M. (2009). Concept inventories for ASTRO 101. *The Physics Teacher*, *47*(7), 439–441.

Bailey, J. M., Johnson, B., Prather, E. E., & Slater, T. F. (2012). Development and validation of the star properties concept inventory. *International Journal of Science Education*, *34*(14), 2257–2286.

Bandaranayake, W. M. (1980). Determination of the absolute stereochemistry of secondary alcohols: An advanced organic chemistry experiment for undergraduate students. *Journal of Chemical Education*, *57*(11), 828–829.

Barbera, J. (2013). A psychometric analysis of the Chemical Concepts Inventory. *Journal of Chemical Education*, *90*(5), 546–553.

Barbera, J., & VandenPlas, J. R. (2011). All assessment materials are not created equal: The myths about instrument development, validity, and reliability. In D. M. Bunce (Ed.), *Investigating Classroom Myths through Research on Teaching and Learning* (pp. 177–193).

Barke, H. D., Hazari, A., & Yitbarek, S. (2008). *Misconceptions in Chemistry*. Berlin: Springer.

Bauer, C. F. (2005). Beyond “student attitudes”: Chemistry self-concept inventory for assessment of the affective component of student learning. *Journal of Chemical Education*, *82*(12), 1864–1870.

Beauchamp, P. S. (1984). “Absolutely” simple stereochemistry. *Journal of Chemical Education*, *61*(8), 666–667.

Bhattacharyya, G. (2006). Practitioner development in organic chemistry: how graduate students conceptualize organic acids. *Chemistry Education Research and Practice*, *7*(4), 240–247.

Bhattacharyya, G. (2014). Trials and tribulations: student approaches and difficulties with proposing mechanisms using the electron-pushing formalism. *Chemistry Education Research and Practice*, *15*(4), 594–609.

Bhattacharyya, G., & Bodner, G. M. (2005). “It gets me to the product”: How students propose organic mechanisms. *Journal of Chemical Education*, *82*(9), 1402–1407.

Biggs, J., & Collis, K. (1982). *Evaluating the Quality of Learning: The SOLO Taxonomy*. New York: Academic Press.

Bodner, G. M. (1986). Constructivism: A theory of knowledge. *Journal of Chemical Education*, *63*(10), 873–878.

Bodner, G. M., & Guay, R. B. (1997). The Purdue Visualization of Rotations Test. *The Chemical Educator*, *2*(4), 1–17.

Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch Model*. Mahwah, New Jersey: Lawrence Erlabaum Associates.

Bowen, C. W. (1994). Think-aloud methods in chemistry education: Understanding student thinking. *Journal of Chemical Education*, *71*(3), 184–190.

Bowen, C. W., & Bodner, G. M. (1991). Problem-solving processes used by students in organic synthesis. *International Journal of Science Education*, *13*(2), 143–158.

Bowling, B. V., Acra, E. E., Wang, L., Myers, M. F., Dean, G. E., Markle, G. C., … Huether, C. A. (2008). Development and evaluation of a genetics literacy assessment instrument for undergraduates. *Genetics*, *178*(1), 15–22.

Brand, D. J., & Fisher, J. (1987). Molecular structure and chirality. *Journal of Chemical Education*, *64*(12), 1035–1038.

Brandriet, A. R., & Bretz, S. L. (2014a). Measuring meta-ignorance through the lens of confidence: Examining students' redox misconceptions about oxidation numbers, charge, and electron transfer. *Chemistry Education Research and Practice*, *15*(4), 729–746.

Brandriet, A. R., & Bretz, S. L. (2014b). The development of the Redox Concept Inventory as a measure of students' symbolic and particulate redox understandings and confidence. *Journal of Chemical Education*, *91*(8), 1132–1144.

Bretz, S. L. (2014). Designing assessment tools to measure students' conceptual knowledge of chemistry. In D. M. Bunce (Ed.), *Tools of Chemistry Education Research* (pp. 155–168). Washington, DC: American Chemical Society.

Bretz, S. L., & Linenberger, K. J. (2012). Development of the enzyme-substrate interactions concept inventory. *Biochemistry and Molecular Biology Education*, *40*(4), 229–233.

Bretz, S. L., & McClary, L. (2015). Students' understandings of acid strength: How meaningful is reliability when measuring alternative conceptions? *Journal of Chemical Education*, *92*(2), 212–219.

Brown, C. E., Hyslop, R. M., & Barbera, J. (2015). Development and analysis of an instrument to assess student understanding of GOB chemistry knowledge relevant to

clinical nursing practice. *Biochemistry and Molecular Biology Education*, *43*(1), 13–19.

Brown, N. J. S., & Wilson, M. (2011). A model of cognition: The missing cornerstone of assessment. *Educational Psychology Review*, *23*(2), 221–234.

Canpolat, N., Pinarbasi, T., & Sözbilir, M. (2006). Prospective teachers' misconceptions of vaporization and vapor pressure. *Journal of Chemical Education*, *83*(8), 1237–1242.

Carey, F. A., & Giuliano, R. M. (2010). *Organic Chemistry* (8th ed.). McGraw-Hill.

Cartrette, D. P., & Dobberpuhl, M. R. (2009). The concept inventory as a diagnostic of understanding and achievement in a year-long organic chemistry course. *The Chemical Educator*, *14*(4), 180–189.

Cartrette, D. P., & Mayo, P. M. (2011). Students' understanding of acids/bases in organic chemistry contexts. *Chemistry Education Research and Practice*, *12*(1), 29–39.

Cetin-Dindar, A., & Geban, O. (2011). Development of a three-tier test to assess high school students' understanding of acids and bases. *Procedia Social and Behavioral Sciences*, *15*, 600–604.

Chamberlain, B. (2012). R,S nomenclature: Help or hindrance in teaching configurational isomerism? Retrieved from http://teachthemechanism.com/2012/12/04/rs-nomenclature-help-or-hindrance-in-teaching-configurational-isomerism/

Chandrasegaran, A. L., Treagust, D. F., & Mocerino, M. (2007). The development of a two-tier multiple-choice diagnostic instrument for evaluating secondary school students' ability to describe and explain chemical reactions using multiple levels of representation. *Chemistry Education Research and Practice*, *8*(3), 293–307.

Cheung, D., Ma, H., & Yang, J. (2009). Teachers' misconceptions about the effects of addition of more reactants or products on chemical equilibrium. *International Journal of Science and Mathematics Education*, *7*(6), 1111–1133.

Clausen, T. P. (2011). Combining a standard Fischer esterification experiment with stereochemical and molecular-modeling concepts. *Journal of Chemical Education*, *88*(7), 1007–1009.

Cloonan, C. A., & Hutchinson, J. S. (2011). A Chemistry Concept Reasoning Test. *Chemistry Education Research and Practice*, *12*(2), 205–209.

Cody, J. A., Craig, P. A., Loudermilk, A. D., Yacci, P. M., Frisco, S. L., & Milillo, J. R. (2012). Design and implementation of a self-directed stereochemistry lesson document format. *Journal of Chemical Education*, *89*(1), 29–33.

Cook, D. L. (1962). The Hawthorne effect in educational research. *The Phi Delta Kappan*, *44*(3), 116–122.

Cooper, M., & Klymkowsky, M. (2013). Chemistry, life, the universe, and everything: A new approach to general chemistry, and a model for curriculum reform. *Journal of Chemical Education*, *90*(9), 1116–1122.

Cooper, M. M. (2007). Data-driven education research. *Science*, *317*, 1171.

Cooper, M. M., Corley, L. M., & Underwood, S. M. (2013). An investigation of college chemistry students' understanding of structure-property relationships. *Journal of Research in Science Teaching*, *50*(6), 699–721.

Cooper, M. M., Grove, N., Underwood, S. M., & Klymkowsky, M. W. (2010). Lost in Lewis structures: An investigation of student difficulties in developing representational competence. *Journal of Chemical Education*, *87*(8), 869–874.

Cooper, M. M., Underwood, S. M., & Hilley, C. Z. (2012). Development and validation of the implicit information from Lewis structures instrument (IILSI): Do students connect structures with properties? *Chemistry Education Research and Practice*, *13*(3), 195–200.

Costa, M. J. (2007). Carbohydeck: a card game to teach the stereochemistry of carbohydrates. *Journal of Chemical Education*, *84*(6), 977–978.

Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five designs*. Thousand Oaks, CA: Sage.

Creswell, J. W., & Plano Clark, V. L. (2007). *Designing and Conducting Mixed Methods Research*. Thousand Oaks: Sage Publications, Inc.

Dos Santos, V. J. S. V., & Galembeck, E. (2015). Metabolic pathways visualization skills development by undergraduate students. *Biochemistry and Molecular Biology Education*, n/a–n/a.

Duis, J. M. (2011). Organic chemistry educators' perspectives on fundamental concepts and misconceptions: An exploratory study. *Journal of Chemical Education*, *88*(3), 346–350.

Ebel, R. L. (1972). *Essentials of educational measurement*. Englewood Cliffs, NJ: Prentice Hall.

Eisen, Y., & Stavy, R. (1988). Students' understanding of photosynthesis. *The American Biology Teacher*, *50*(4), 208–212.

Eliel, E. L., Wilen, S. H., & Mander, L. N. (1994). *Stereochemistry of Organic Compounds*. New York: Wiley.

Emenike, M. E., & Holme, T. A. (2012). Classroom response systems have not "crossed the chasm": Estimating numbers of chemistry faculty who use clickers. *Journal of Chemical Education*, *89*(4), 465–469.

Fair, J. D., Kleist, E. M., & Stoy, D. M. (2014). A survey of industrial organic chemists: Understanding the chemical industry's needs of current bachelor-level graduates. *Journal of Chemical Education*, *91*(12), 2084–2092.

Fisher, K. M., Williams, K. S., & Lineback, J. E. (2011). Osmosis and diffusion conceptual assessment. *Cell Biology Education*, *10*(4), 418–429.

Gall, M. D., Gall, J. P., & Borg, W. R. (2003). *Educational Research: An Introduction*. Boston, MA: Allyn and Bacon.

Gay, L. R. (1992). *Educational Research: Competencies for Analysis and Application* (4th ed.). Englewood Cliffs, NJ: MacMillan.

Gerlach, K., Trate, J., Blecking, A., Geissinger, P., & Murphy, K. (2014). Investigation of absolute and relative scaling conceptions of students in introductory college chemistry courses. *Journal of Chemical Education*, *91*(10), 1526–1537.

Gilchrist, T. L. (Ed.). (1995). *Comprehensive Organic Functional Group Transformations. Synthesis: Carbon with Three or Four Attached Heteroatoms*. Elsevier.

Glaser, B. G., & Strauss, A. L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New York, NY: Aldine De Gruyter.

Granville, M. F. (1985). Student misconceptions in thermodynamics. *Journal of Chemical Education*, *62*(10), 847–848.

Graulich, N. (2015). The tip of the iceberg in organic chemistry classes: how do students deal with the invisible? *Chemistry Education Research and Practice*, *16*(1), 9–21.

Grove, N. P., Cooper, M. M., & Cox, E. L. (2012). Does mechanistic thinking improve student success in organic chemistry? *Journal of Chemical Education*, *89*(7), 850–853.

Grove, N. P., Cooper, M. M., & Rush, K. M. (2012). Decorating with arrows: Toward the development of representational competence in organic chemistry. *Journal of Chemical Education*, *89*(7), 844–849.

Hake, R. R. (1997). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, *66*(1), 64–74.

Haladyna, T. M. (2004). *Developing and Validating Multiple-Choice Test Items*. Mahwah, New Jersey: Lawrence Erlabaum Associates.

Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, *15*(3), 309–333.

Haslam, F., & Treagust, D. F. (1987). Diagnosing secondary students' misconceptions of photosynthesis and respiration in plants using a two-tier multiple choice instrument. *Journal of Biological Education*, *21*(3), 203–211.

Hazel, E., & Prosser, M. (1994). First-year university students' understanding of photosynthesis, their study strategies & learning context. *The American Biology Teacher*.

Henderleiter, J., Smart, R., Anderson, J., & Elian, O. (2001). How do organic chemistry students understand and apply hydrogen bonding? *Journal of Chemical Education*, *78*(8), 1–5.

Herman, G. L., Loui, M. C., & Zilles, C. (2010). Creating the digital logic concept inventory. In *SIGCSE '10: Proceedings of the 41st ACM technical symposium on Computer science education* (pp. 102–106). New York, New York, USA: ACM Request Permissions.

Herman, G. L., Zilles, C., & Loui, M. C. (2014). A psychometric evaluation of the digital logic concept inventory. *Computer Science Education*, *24*(4), 277–303.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, *30*(3), 141–158.

Holme, T., Bretz, S. L., Cooper, M., Lewis, J., Paek, P., Pienta, N., … Towns, M. (2010). Enhancing the role of assessment in curriculum reform in chemistry. *Chemistry Education Research and Practice*, *11*(2), 92–97.

Houseknecht, J. B. (2010). Topic sequence and emphasis variability of selected organic chemistry textbooks. *Journal of Chemical Education*, *87*(6), 592–597.

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, *107*(2), 139–155.

Johnstone, A. H. (1982). Macro and microchemistry. *School Science Review*, *64*, 377–379.

Johnstone, A. H. (1991). Why is science difficult to learn? Things are seldom what they seem. *Journal of Computer Assisted Learning*, *7*(2), 75–83.

Johnstone, A. H. (1993). The development of chemistry teaching: A changing response to changing demand. *Journal of Chemical Education*, *70*(9), 701–705.

Kaplan, R., & Saccuzzo, D. (2012). *Psychological Testing: Principles, Applications, and Issues*. Wadsworth Publishing.

Kelley, T. L. (1939). The selection of upper and lower groups for the validation of test items. *Journal of Educational Psychology*, *30*(1), 17–24.

Kline, T. J. B. (2005). *Psychological Testing: A practical approach to design and evaluation*. Thousand Oaks, CA: Sage.

Köse, S. (2008). Diagnosing student misconceptions: using drawings as a research method. *World Applied Sciences Journal*.

Koutalas, V. G., Antonoglou, L. D., Charistos, N. D., & Sigalas, M. P. (2014). Investigation of students' ability to transform and translate 2D molecular diagrammatic representations and its relationship to spatial ability and prior chemistry knowledge. *Procedia - Social and Behavioral Sciences*, *152*, 698–703.

Krishnan, S. R., & Howe, A. C. (1994). The mole concept: Developing an instrument to assess conceptual understanding. *Journal of Chemical Education*, *71*(8), 653–655.

Krylova, I. (1997). *Investigation of causes of differences in student performance on the topics of stereochemsitry and reaction mechanisms in an undergraduate organic chemistry course*. ProQuest Dissertations and Theses.

Kumi, B. C., Olimpo, J. T., Bartlett, F., & Dixon, B. L. (2013). Evaluating the effectiveness of organic chemistry textbooks in promoting representational fluency and understanding of 2D–3D diagrammatic relationships. *Chemistry Education Research and Practice*, *14*(2), 177–187.

Kuo, M.-T., Jones, L. L., Pulos, S. M., & Hyslop, R. M. (2004). The relationship of molecular representations, complexity, and orientation to the difficulty of stereochemistry problems. *The Chemical Educator*, *9*, 321–327.

Kurbanoglu, N. I., Taskesenligil, Y., & Sözbilir, M. (2006). Programmed instruction revisited: a study on teaching stereochemistry. *Chemistry Education Research and Practice*, *7*(1), 13–21.

Leontjev, A. E., Vasiljeva, L. L., & Pivnitsky, K. K. (2004). Reduction of steroidal ketones with amine—boranes. *Russian Chemical Bulletin*, *53*(3), 703–708.

Leontyev, A., & Baranov, D. (2013). Massive open online courses in chemistry: A comparative overview of platforms and features. *Journal of Chemical Education*, *90*(11), 1533–1539.

Liew, C. W., & Treagust, D. F. (1995). A Predict-Observe-Explain teaching sequence for learning about students' understanding of heat and expansion liquids. *Australian Science Teachers Journal*, *41*(1), 68–71.

Lindell, R. S., & Olsen, J. P. (2002). Developing the Lunar Phases Concept Inventory. In *Physics Education Research Conference* (pp. 1–4). Boise, ID.

Lindell, R. S., Peak, E., & Foster, T. M. (2007). Are they all created equal? A comparison of different concept inventory development methodologies. In *Physics Education Research Conference* (Vol. 883, pp. 14–17). AIP Publishing.

Linenberger, K. J. (2011). *Biochemistry Students' Understandings of Enzyme-Substrate Interactions as Investigated through Multiple Representations and the Enzyme-Substrate Interactions Concept Inventory*. Miami University.

Linenberger, K. J., & Bretz, S. L. (2014). Biochemistry students' ideas about shape and charge in enzyme-substrate interactions. *Biochemistry and Molecular Biology Education*, *42*(3), 203–212.

Loh, A. S. L., Subramaniam, R., & Tan, K.-C. D. (2014). Exploring students' understanding of electrochemical cells using an enhanced two-tier diagnostic instrument. *Research in Science & Technological Education*, *32*(3), 229–250.

Lord, F. M. (1952). The relation of the reliability of multiple-choice tests to the distribution of item difficulties. *Psychometrika*, *17*(2), 181–194.

Luxford, C. J., & Bretz, S. L. (2014). Development of the bonding representations inventory to identify student misconceptions about covalent and ionic bonding representations. *Journal of Chemical Education*, *91*(3), 312–320.

Lyon, G. L. (1999). *College students' understanding of stereochemistry: Difficulties in learning and critical junctures*. Louisiana State University.

MacArthur, J. R., & Jones, L. L. (2008). A review of literature reports of clickers applicable to college chemistry classrooms. *Chemistry Education Research and Practice*, *9*(3), 187–195.

Marbach-Ad, G., McAdams, K. C., Benson, S., Briken, V., Cathcart, L., Chase, M., … Smith, A. C. (2010). A model for using a concept Inventory as a tool for students' assessment and faculty professional development. *Cell Biology Education*, *9*(4), 408–416.

McBride, M. (2011). Freshman Organic Chemistry I. Retrieved from oyc.yale.edu/chemistry/chem-125a

McClary, L. M., & Bretz, S. L. (2012). Development and assessment of a diagnostic tool to identify organic chemistry students' alternative conceptions related to acid strength. *International Journal of Science Education*, *34*(15), 2317–2341.

McClary, L., & Talanquer, V. (2011). College chemistry students' mental models of acids and acid strength. *Journal of Research in Science Teaching*, *48*(4), 396–413.

Mdachi, S. J. M. (2012). Giving thought to students' alternative conceptions in stereochemistry: one teacher's basis for pedagogical content knowledge improvement. *Journal of Turkish Science Education*, *9*(4), 22–34.

Merriam, S. B. (2009). *Qualitative Research: A giude to Design and Implementations*. John Wiley & Sons.

Meyer, P. J. (2013). jMetrik.

Moss, G. P. (1996). Basic terminology of stereochemistry. *Pure and Applied Chemistry*, *68*(12), 2193–2222.

Mulford, D. R., & Robinson, W. R. (2002). An inventory for alternate conceptions among first-semester general chemistry students. *Journal of Chemical Education*, *79*(6), 739–744.

Nakhleh, M. B. (1992). Why some students don't learn chemistry: Chemical misconceptions. *Journal of Chemical Education*, *69*(3), 191–196.

Nakhleh, M. B. (1994). Student's models of matter in the context of acid-base chemistry. *Journal of Chemical Education*, *71*(6), 495.

Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research*, *76*(3), 413–448.

Nitko, A. J., & Brookhart, S. M. (2011). *Educational Assessment of Students*. Boston, MA: Pearson/Allyn & Bacon.

Novak, J. D. (1990). Concept mapping: a useful tool for science education. *Journal of Research in Science Teaching*, *27*(10), 937–949.

Nyachwaya, J. M., Mohamed, A.-R., Roehrig, G. H., Wood, N. B., Kern, A. L., & Schneider, J. L. (2011). The development of an open-ended drawing tool: an alternative diagnostic tool for assessing students' understanding of the particulate nature of matter. *Chemistry Education Research and Practice*, *12*(2), 121–132.

Olimpo, J. T., Kumi, B. C., Wroblewski, R., & Dixon, B. L. (2015). Examining the relationship between 2D diagrammatic conventions and students' success on representational translation tasks in organic chemistry. *Chemistry Education Research and Practice*, *16*(1), 143–153.

Oliver-Hoyo, M., & Sloan, C. (2014). The development of a Visual-Perceptual Chemistry Specific (VPCS) assessment tool. *Journal of Research in Science Teaching*, *51*(8), 963–981.

Osborne, R. J., & Cosgrove, M. M. (1983). Children's conceptions of the changes of state of water. *Journal of Research in Science Teaching*, *20*(9), 825–838.

Patton, M. Q. (1990). *Qualitative evaluation and research methods*. Thousand Oaks, CA: SAGE Publications.

Pavelich, M., Jenkins, B., Birk, J., Bauer, R., & Krause, S. (2013). Development of a Chemistry Concept Inventory for use in chemistry, materials and other engineering courses. In *Proceedings of the 2004 American Society for Engineering Annual Conference & Exposition* (pp. 1–8).

Peterson, R. F., Treagust, D. F., & Garnett, P. (1989). Development and application of a diagnostic instrument to evaluate grade-11 and -12 students' concepts of covalent bonding and structure following a course of instruction. *Journal of Research in Science Teaching*, *26*(4), 301–314.

Polavarapu, P. L. (2002). The absolute configuration of Bromochlorofluoromethane. *Angewandte Chemie International Edition*, *41*(23), 4544–4546.

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, *29*(5), 489–97.

Pribyl, J. R., & Bodner, G. M. (1987). Spatial ability and its role in organic chemistry: A study of four organic courses. *Journal of Research in Science Teaching*, *24*(3), 229–240.

Rodriguez, M. C. (2005). Three options are optimal for multiple□choice items: A meta□ analysis of 80 years of research. *Educational Measurement: Issues and Practice*, *24*(2), 3–13.

Rushton, G. T., Hardy, R. C., Gwaltney, K. P., & Lewis, S. E. (2008). Alternative conceptions of organic chemistry topics among fourth year chemistry students. *Chemistry Education Research and Practice*, *9*(2), 122–130.

Schmidt, H. J. (1992). Conceptual difficulties with isomerism. *Journal of Research in Science Teaching*, *29*(9), 995–1003.

Schönborn, K., Höst, G., & Lundin Palmerius, K. (2015). Measuring understanding of nanoscience and nanotechnology: Development and validation of the nano-knowledge instrument (NanoKI). *Chemistry Education Research and Practice*.

Schwartz, P. M., Lepore, D. M., Morneau, B. N., & Barratt, C. (2011). Demonstrating optical activity using an iPad. *Journal of Chemical Education*, *88*(12), 1692–1693.

Şendur, G. (2012). Prospective science teachers' misconceptions in organic chemistry: The case of alkenes. *Journal of Turkish Science Education*, *9*(3), 186–190.

Singer, S. R., Nielsen, N. R., & Schweingruber, H. A. (2012). *Discipline-Based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering*. Washington, DC: The National Academies Press.

Smith, M. K., Wood, W. B., & Knight, J. K. (2008). The Genetics Concept Assessment: a new concept inventory for gauging student understanding of genetics. *Cell Biology Education*, *7*(4), 422–430.

Solomons, G. T. W., Fryhle, C. B., & Snyder, S. A. (2014). *Organic Chemistry* (11th ed.). New York, NY: John Wiley & Sons.

Stains, M., Escriu-Sune, M., Molina Alvarez de Santizo, M. L., & Sevian, H. (2011). Assessing secondary and college students' implicit assumptions about the particulate nature of matter: Development and validation of the Structure and Motion of Matter Survey. *Journal of Chemical Education*, *88*(10), 1359–1365.

Stamm, K. R., Clark, F., & Eblacas, P. R. (2000). Mass communication and public understanding of environmental problems: the case of global warming. *Public Understanding of Science*, *9*(3), 219–237.

Stavy, R. (1990). Children's conception of changes in the state of matter: From liquid (or solid) to gas. *Journal of Research in Science Teaching*, *27*(3), 247–266.

Steif, P. S., & Dantzler, J. A. (2005). A Statics Concept Inventory: Development and psychometric analysis. *Journal of Engineering Education*, *94*(4), 363–371.

Stone, A., Allen, K., Rhoads, T. R., Murphy, T. J., Shehab, R. L., & Saha, C. (2003). The statistics concept inventory: a pilot study. In *Frontiers in Education Conference. Proceedings* (Vol. 1, p. T3D–T31). IEEE.

Streveler, R. A., Miller, R. L., Santiago-Roman, A. I., Nelson, M. A., Geist, M. R., & Olds, B. M. (2011). Rigorous methodology for concept inventory development: Using the "Assessment Triangle" to develop and test the Thermal and Transport Science Concept Inventory (TTCI). *International Journal of Engineering Education*, *27*(5), 968.

Taagepera, M., & Noori, S. (2000). Mapping students' thinking patterns in learning. *Journal of Chemical Education*, *77*(9), 1224–1229.

Taber, K. (2002). *Chemical misconceptions: Prevention, diagnosis and cure* (Vol. 1). Royal Society of Chemistry.

Taber, K. S. (1998). An alternative conceptual framework from chemistry education. *International Journal of Science Education*, *20*(5), 597–608.

Talanquer, V. (2006). Commonsense chemistry: A model for understanding students' alternative conceptions. *Journal of Chemical Education*, *83*(5), 811–816.

Tan, K.-C. D., Goh, N.-K., Chia, L.-S., & Treagust, D. F. (2002). Development and application of a two-tier multiple choice diagnostic instrument to assess high school students' understanding of inorganic chemistry qualitative analysis. *Journal of Research in Science Teaching*, *39*(4), 283–301.

Tan, K.-C. D., Taber, K. S., Goh, N.-K., & Chia, L.-S. (2005). The ionisation energy diagnostic instrument: a two-tier multiple-choice instrument to determine high school students' understanding of ionisation energy. *Chemistry Education Research and Practice*, *6*(4), 180–197.

Tan, K.-C. D., & Treagust, D. F. (1999). Evaluating students' understanding of chemical bonding. *School Science Review*, *81*(294), 75–84.

Taskin, V., Bernholt, S., & Parchmann, I. (2015). An inventory for measuring student teachers' knowledge of chemical representations: design, validation, and psychometric analysis. *Chemistry Education Research and Practice*.

Thorndike, R. M., & Thorndike-Christ, T. M. (2011). *Measurement and Evaluation in Psychology and Education* (8th ed.). Pearson Education Inc.

Topal, G., Oral, B., & Özden, M. (2007). University and secondary school students' misconceptions about the concept of "aromaticity" in organic chemistry. *International Journal of Environmental and Science Education*, *2*(4), 135–143.

Towns, M. H. (2008). Mixed methods designs in chemical education. In D. M. Bunce (Ed.), *Nuts and Bolts of Chemical Education Research* (pp. 135–148). Washington, DC: American Chemical Society.

Towns, M. H. (2014). Guide to developing high-quality, reliable, and valid multiple-choice assessments. *Journal of Chemical Education*, *91*(9), 1426–1431.

Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science. *International Journal of Science Education*, *10*(2), 159–169.

Tüysüz, C. (2009). Development of two-tier diagnostic instrument and assess students' understanding in chemistry. *Scientific Research and Essay*, *4*(6), 626–631.

Vachliotis, T., Salta, K., & Tzougraki, C. (2013). Meaningful understanding and systems thinking in organic chemistry: Validating measurement and exploring relationships. *Research in Science Education*, *44*(2), 239–266.

Vachliotis, T., Salta, K., Vasiliou, P., & Tzougraki, C. (2011). Exploring novel tools for assessing high school students' meaningful understanding of organic reactions. *Journal of Chemical Education*, *88*(3), 337–345.

Villafañe, S. M., Bailey, C. P., Loertscher, J., Minderhout, V., & Lewis, J. E. (2011). Development and analysis of an instrument to assess student understanding of foundational concepts before biochemistry coursework. *Biochemistry and Molecular Biology Education*, *39*(2), 102–109. doi:10.1002/bmb.20464

Villafañe, S. M., Loertscher, J., Minderhout, V., & Lewis, J. E. (2011). Uncovering students' incorrect ideas about foundational concepts for biochemistry. *Chemistry Education Research and Practice*, *12*(2), 210.

Voska, K. W., & Heikkinen, H. W. (2000). Identification and analysis of student conceptions used to solve chemical equilibrium problems. *Journal of Research in Science Teaching*, *37*(2), 160–176.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wade, L. G. (2005). *Organic Chemistry* (5th ed.). Prentice Hall.

Wandersee, J. H., Mintzes, J. J., & Novak, J. D. (1994). Research on Alternative Conceptions in Science. In D. L. Gabel (Ed.), *Handbook of Research on Science Teaching and Learning* (pp. 177–210). New York: MacMillan.

Wathen, M. D. (2008). *Effect of molecular representation on chemsitry students' understanding of R and S configuration*. University of Northern Colorado.

Weimer, M. (2013). *Learner-Centered Teaching* (2nd edition). John Wiley & Sons.

Wenning, C. J. (2008). Dealing more effectively with alternative conceptions in science. *Journal of Physics Teacher Education Online*, *5*(1), 11–19.

Wilson, D. B. (n.d.). The Campbell Collaboration: Practical Meta Analysis Effect Size Calculator. Retrieved June 20, 2015, from http://www.campbellcollaboration.org/resources/effect_size_input.php

Wilson, M. (2005). *Constructing Measures: an Item Response Modeling Approach*. London: Lawrence Erlabaum Associates.

Wren, D., & Barbera, J. (2013). Gathering evidence for validity during the design, development, and qualitative evaluation of Thermochemistry Concept Inventory items. *Journal of Chemical Education*, *90*(12), 1590–1601.

Wren, D., & Barbera, J. (2014). Psychometric analysis of the thermochemistry concept inventory. *Chemistry Education Research and Practice*, *15*(3), 380–390.

Wu, H.-K., & Shah, P. (2004). Exploring visuospatial thinking in chemistry learning. *Science Education*, *88*(3), 465–492.

Yezierski, E. J., & Birk, J. P. (2006). Misconceptions about the particulate nature of matter. *Journal of Chemical Education*, *83*(6), 954–960.

Zoller, U. (1990). Students' misunderstandings and misconceptions in college freshman chemistry (general and organic). *Journal of Research in Science Teaching*, *27*(10), 1053–1065.

Zoller, U. (1996). The use of examinations for revealing and distinguishing between students' misconceptions, misunderstandings and "no conceptions" in college chemistry. *Research in Science Education*, *26*(3), 317–326.

APPENDIX A

TAXONOMY OF WRITING

MULTIPLE-CHOICE

QUESTIONS

1. Every item should reflect specific content and a single specific mental behavior, as called for in test specifications (two-way grid, test blueprint).
2. Base each item on important content to learn; avoid trivial content.
3. Use novel material to test higher level learning. Paraphrase textbook language or language used during instruction when used in a test item to avoid testing for simple recall.
4. Keep the content of each item independent from content of other items on the test.
5. Avoid overly specific and overly general content when writing multiple-choice items.
6. Avoid opinion-based items.
7. Avoid trick items.
8. Keep vocabulary simple for the group of students being tested.
9. Use the question, completion, and best answer versions of the conventional multiple-choice, the alternate choice, true-false, multiple true-false, matching, and the context-dependent item and item set formats, but avoid the complex multiple-choice (Type K) format.
10. Format the item vertically instead of horizontally.
11. Edit and proof items.
12. Use correct grammar, punctuation, capitalization, and spelling.
13. Minimize the amount of reading in each item.
14. Ensure that the directions in the stem are very clear.
15. Include the central idea in the stem instead of the choices.
16. Avoid window dressing (excessive verbiage).
17. Word the stem positively, avoid negatives such as NOT or EXCEPT. If negative words are used, use the word cautiously and always ensure that the word appears capitalized and boldface.
18. Develop as many effective choices as you can, but research suggests three is adequate.
19. Make sure that only one of these choices is the right answer.
20. Vary the location of the right answer according to the number of choices.
21. Place choices in logical or numerical order.
22. Keep choices independent; choices should not be overlapping.
23. Keep choices homogeneous in content and grammatical structure.
24. Keep the length of choices about equal.
25. None-of-the-above should be used carefully.
26. Avoid All-of-the-above.
27. Phrase choices positively; avoid negatives such as NOT.
28. Avoid giving clues to the right answer
29. Make all distractors plausible.
30. Use typical errors of students to write distractors.
31. Use humor if it is compatible with the teacher and the learning environment.

Adapted from

Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied measurement in education*, *15*(3), 309-333.

APPENDIX B

STEREOCHEMISTRY INSTRUCTIONAL SURVEY

1. Inclusion Criteria

Do you currently teach or have taught within 5 years two-semester organic chemistry course?

- o Yes, I teach or have taught two-semester organic chemistry course
- o No, I do not teach two-semester organic chemistry course, but I still would like to provide my opinion on stereochemistry instruction

2. Content of Stereochemistry Instruction

Consider your typical organic chemistry course. For each of the topics listed below, select the appropriate response indicating if you consider the topic to be important, optional, or not important. For clarity, we define the terms as following: "Important" refers to topics you consider important, always teach, and include as a part of your assessment. "Optional" refers to topics that you may teach if time allows, or assign students to read in the course textbook. These topics are not a part of your regular assessment. "Not Important" refers to topics that you never or almost never teach and do not consider relevant to your stereochemistry instruction.

Different types of isomers and stereoisomers.

Constitutional isomers
Cis-trans-isomers for compounds with double bond
Cis-trans-isomers for alicyclic compounds
Conformers
Enantiomers
Diastereomers
Meso-forms

Nomenclature and rules

Cahn-Ingold-Prelog priority rules
R/S nomenclature
+/- nomenclature
d/l nomenclature
D/L nomenclature
M/P nomenclature
Erythro and threo nomenclature

Properties of light

Block diagram of a polarimeter
Circularly polarized light
Plane-polarized light
Optical rotation
Levorotatory and dextrorotatory compounds
Specific rotation

Resolution of enantiomers

Resoluton of enantiomers through formation of diastereomers
Resoluton of enantiomers through enzymatic binding
Resoluton of enantiomers through chiral chromatography

Representation types

Wedge-dash structures

Fischer projections
Newman projections
Sawhorse projections
Haworth projections

Symmetry elements

Plane of symmetry
Symmetry axis
Inversion center
Improper rotation axis

Specific examples of chirality

Stereoisomerism of disubstituted cyclohexanes
Chirality at atoms other than carbon (nitrogen, sulfur, or phosphorus)
Chirality of octahedral complexes of metals
Allene chirality
Chirality of substituted biphenyls
Chirality of helicenes
Chirality of natural compounds

Various

Racemic mixtures
Enantiomeric excess (ee)
Relationship between stereoisomers with multiple chiral centers
Total number of stereoisomers for a compound with m chiral centers (n = 2^m)
Physical properties (melting points, densities, and solubility) of stereoisomers
Equivalence of physical properties of enantiomers
Non-equivalence of physical properties of diastereomers
Non-equivalence of physical properties of meso-forms and racemic mixtures
Different biological activities of stereoisomers
Enantiotopic and diastereotopic atoms
Prochirality

3. Learning Objectives (selected)

Below is a list of possible learning objectives. Select which skills you expect your students to develop.

Assign R/S descriptors to stereocenters
Identify stereogenic atoms and chiral molecules
Identify molecules that are mirror images (enantiomers)
Describe the stereochemical relationships between molecules
Know and apply Cahn-Ingold-Prelog priority rules
Know the block diagram of a polarimeter
Recognize enantiotopic and diastereotopic atoms
Recognize all elements of symmetry in molecules
Recognize a plane of symmetry in the molecule
Compare and differentiate physical properties of enantiomers and diastereomers
Understand the different reactivity of diastereoisomers
Understand why enantiomers have different biological properties
Interconvert various types of representations
Recognize chirality at non-carbon stereogenic centers
Recognize chirality in allenes or substituted biphenyls
Know the relationship between enantiomers and their specific rotations
Calculate enantiomeric excess knowing % of enantiomers in a mixture
Interconvert enantiomeric excess and specific rotation
Distinguish between conformers and stereoisomers
Draw the family of stereoisomers for a given compound with one stereogenic center
Draw the family of stereoisomers for a given compound with two stereogenic centers
Draw the family of stereoisomers for a given compound with three or more stereogenic centers

If you have any other objectives for the "Stereochemistry" chapter that are not part of those mentioned above, please share them with us.

4. Demographic Information

What is the highest degree in chemistry offered at your institution?

- Associate's Degree
- Bachelor's Degree
- Master's Degree
- Doctorate Degree

How many years of organic chemistry teaching experience do you have?

- 1-3 years
- 4-5 years
- 6-10 years
- 11-15 years
- 16-20 years
- 21-25 years
- more than 26 years

What is the highest degree you have received?

- Master's Degree
- Doctorate Degree

What is your primary area of expertise?

- Analytical Chemistry
- Biological Chemistry
- Inorganic Chemistry
- Organic Chemistry
- Physical Chemistry

APPENDIX C

INTERVIEW GUIDE

What types of isomers do you know? Could you provide examples?

(If stereoisomers are not mentioned here, lead to examples that have a chiral center)

What are stereoisomers?

(If a student mentions a chiral atom here, we agree that we call it chiral center)

Stereoisomers differ from each other in what respect?

(Lead it to discussion about structure, not properties)

What are enantiomers? Could you provide an example of an enantiomer?

(If example with a chiral center provided, ask about possible arrangements in space)

= at this point 2 enantiomers drawn by student should be on the page =

How can you separate two enantiomers?

(If recrystallization or distillation mentioned, ask about difference in boiling points or solubility here. If a student tells about difference in boiling/melting points here, ask about difference in intermolecular forces. If recrystallization is mentioned, ask why certain compounds are soluble and some are not. Don't give the correct answer at this stage, ask the same question after discussion of diastereomers)

Do enantiomers have different melting points? Boiling points? Any other physical properties?

(Refer to the structures, if they think they are different, ask why, ask about intermolecular forces)

Do enantiomers have different biological properties? Why?

(Ask why, what does body consist of, are those molecules chiral)

(or – draw the structure of l-alanine and label that as sweet, and d-alanine as tasteless and ask why)

If in previous questions "light" was mentioned, ask:

What is the difference between "light" and "polarized light"?

How does polarimeter work?

What used as a light source in polarimeter?

What is specific rotation? What does it depend on? *(Ask about concentration)*

What are diastereomers? Could you provide an example of a diastereomer?

(Lead to the structure with multiple chiral centers if student does not know the correct answer)

How can you separate two diastereomers?

(Talk about intermolecular interactions)

How can we determine the total number of isomers for a compound with multiple chiral centers?

(Draw the structure of a compound with 3 or 5 chiral centers – or ask a student to draw it and ask what is the total number of stereoisomers)

Do diastereomers have different melting points? Boiling points?

(If previous question lead to talking about intermolecular forces, combine these two)

What is a racemate? Could you provide an example?

(Refer to the structures of a pair of enantiomers if the student has problems. If the student got the correct answer for enantiomers, ask about a racemate for compounds with multiple chiral centers)

What is a *meso*-compound? Could you provide an example?

(If correct answer was shown, ask where the plane of symmetry is, ask how if there can be any other other meso compounds)

What is a plane of symmetry? Can you draw an example of one?

If you have a compound with 4 chiral centers and a plane of symmetry, how many meso-compounds are possible?

If we have a sample of a compound, how can we tell if it is an R-isomer or an S-isomer? What methods can we use to find it out?

If we have a sample of a compound, how can we tell if it is a d-isomer or an l-isomer? What methods can we use to find it out?

Consider each of following statements. Is it true or false? Could you justify your answer? *(Ask these in other questions, where it's the most appropriate)*

> **A carbon atom with four substituents is a stereogenic center.**
>
> **Every molecule with two or more stereogenic centers is chiral.**
>
> **Compounds with an *R*-stereocenter rotate plane-polarized light clockwise.**
>
> **A racemic mixture can rotate plane-polarized light either clockwise or counterclockwise.**
>
> ***Meso*-compounds can rotate plane-polarized light either clockwise or counterclockwise.**

APPENDIX D

EXAMPLES OF CODES

Students' statements	Codes
Because the rotation is different and therefore it isn't absorbed by body the same way.	Optical rotation
The stereochemistry of ibuprofen is only recognized by the body and the other is not because its stereochemistry is different.	Configuration
The different optical rotation effects the functionality of the molecule. I forgot how/why.	Optical rotation
(-) ibuprofen does not relieve pain because it is not optically active like (+).	Optical rotation
The configuration effects how the molecule reacts withing an aques environment I really cannot remember.	Configuration
Because it attaches to molecules differently they don't "fit" the same way.	Interaction
Because + ibuprofen can fit into an enzyme when (-) ibuprofen is cannot.	Interaction
Because (-) ibuprofen rotates the opposite way then + ibuprofen and cant react the same way in the body sinse its going the opposite way.	Optical rotation
Because the + ibuprofen has the correct configuration to interact with other molecules.	Interaction
Because the orientation of the COOH is different way they have different chacteristics.	Configuration
(+) ibuprofen reliefs pain because can attach to pain signals preventing them from going to the brain. Whereass (-) ibuprofen cannot attach to pain signals in the same way. Therefore, these pain signals still go to the brain and pains felt.	Interaction
It all depends on the position of COOH group to how molecule interacts with the receptor it binds to block pain.	Configuration Interaction
Because + ibuprofen reacts with polarized light and – ibuprofin does not. So the + ibuprofen reacts with the chemicals in your body because it is in the plane of polarized light.	Optical rotation
(+) ibuprofen releaves pain because it is the enantiomer that is notices by whatever part of the body that produces pain relieving affect. The other enantiomer (-) ibuprofein is not notices by the body.	Interaction
I would think that the configuration of R asparanine allows for carboxylic acid to bind to the taste buds rather than the amino groups on the S asparagine. I am not sure.	Configuration Interaction
Your body only recognizes 1 enantiomer not the other (specific binding sites).	Interaction

APPENDIX E

TEST BLUEPRINT FOR SCI-30Q

Major Topic	Components of the topic	Question #s	Knowledge statements (addressed in the correct responses)
Chirality	Chiral and achiral molecules	1, 2, 8, 27	- chiral molecules are molecules that have non superimposable mirror images - achiral molecules are molecules that have superimposable mirror images
	Stereogenic centers	3, 4	- stereogenic center is an atom with four different groups attached to it
Stereoisomers	Enantiomers and diastereomers	13,15, 16, 17, 19	- enantiomers are non-superimposable mirror images - diastereomers are stereoisomers that are not mirror images
	Meso compounds	2, 8, 10, 29	- meso compounds are achiral molecule that have chirality centers
	Physical properties of stereoisomers	21, 22, 23	- enantiomers have the same physical properties - diastereomers have different physical properties
Optical activity	Levorotary and dextrorotary compounds	5, 6, 7	- levorotary compounds rotate plane polarized light to the left - dextrorotary compounds rotate plane polarized light to the right
	Relationship between structure and optical activity	28, 30	- chiral molecules are optically active - achiral molecules are not optically active - meso compounds are not optically active - racemic mixtures are not optically active
R, S nomenclature	The Cahn-Ingold-Prelog priority rules	9, 10, 14	- groups are ranked in order of precedence according to rules based on atomic numbers
	R or *S* configuration	11, 14	- system for specifying absolute configuration as R or S on the basis of the order in which groups are attached to a chirality center.
Representations	Wedge-dash projections	1, 2, 11, 12, 13, 14, 15, 16, 19, 26, 27, 29	- wedge and dash projection is a method for representing a molecule in which three types (solid, dashed, and wedge-shaped) of lines are used in order to represent the three-dimensional structure
	Fischer projections	17, 18, 20, 24, 25	- Fischer projections is a method for representing the spatial arrangement of groups around chiral carbon atoms; the four bonds to the chiral carbon are represented by a cross, with the assumption that the horizontal bonds project toward the viewer and the vertical bonds away from the viewer.
	Newman projections	17, 18	- Newman projections is a method of representation of a molecule in which the viewer's eye is considered to be sighting down a carbon-carbon bond; the front carbon is represented by a point and the back carbon by a circle

APPENDIX F

FINAL VERSION OF THE STEREOCHEMISTRY CONCEPT INVENTORY

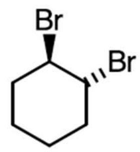
STEREOCHEMISTRY CONCEPT INVENTORY

Put the letter that corresponds to your answer in the space by question number. Each question has only one correct answer.

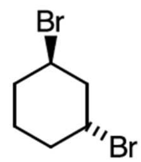
___ **0**. I agree to participate in the study and I permit the researcher to use my answers for the research purposes.
A. Yes **B.** No

___ **1.** Which of these molecules is(are) achiral?

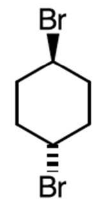
A. I.
B. II.
C. III.
D. I, II, and **III.**



I



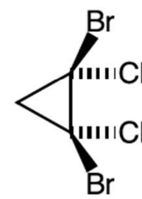
II



III

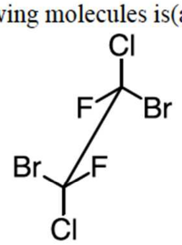
___ **2.** Is the following molecule chiral?

A. Yes, because its mirror image cannot be overlaid on itself.
B. Yes, because it has a carbon atom with four different substituents.
C. No, because it is a racemic compound.
D. No, because it has a plane of symmetry.

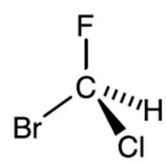


___ **3.** Which of the following molecules is(are) chiral?

A. I only.
B. II only.
C. I and **II**.



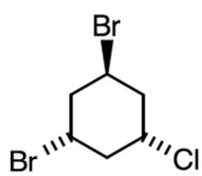
I



II

___ **4.** How many stereocenters(s) exist in the following compound?

A. Zero.
B. One.
C. Two.
D. Three.



___ **5.** A compound has *n* stereocenters. What is the maximum number of stereoisomers possible for this compound?

A. 2^n
B. n^2
C. $2n$
D. $2n + 2$

Stereochemistry Concept Inventory

___ **6.** Which of the following substituents has the highest rank according to the Cahn-Ingold-Prelog rules?

–OC(CH_3)$_3$, –F, or –Cl

A. –Cl, because it has the highest atomic mass.
B. –F, because it is the most electronegative element.
C. –OC(CH_3)$_3$, because it is the largest by size.

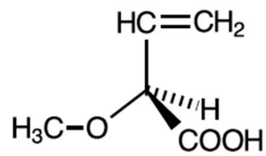
___ **7.** Which of the following substituents has the higher priority according to the Cahn-Ingold-Prelog rules?

–NH_2 or –Br

A. –NH_2 because N is more electronegative than Br.
B. –NH_2 because it has more atoms than Br.
C. –Br because it has a higher atomic mass than N.
D. –Br because it is larger than N.

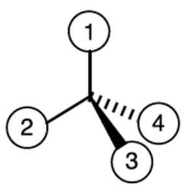
___ **8.** What are the configuration and relative priority of the substituents at the stereogenic center in the following molecule?

A. Configuration: *R*; Priority: HC=CH_2 > COOH > OCH_3.
B. Configuration: *R*; Priority: OCH_3 > COOH > HC=CH_2.
C. Configuration: *S*; Priority: OCH_3 > COOH > HC=CH_2.
D. Configuration: *S*; Priority: HC=CH_2 > COOH > OCH_3.

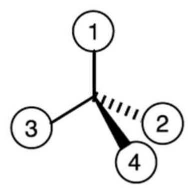


___ **9.** In the following molecules, substituents are arranged according to their priority, where "1" has the highest priority and "4" has the lowest priority. Which of the molecules has an "R" configuration?

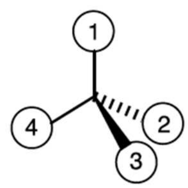
A. **I.**
B. **II.**
C. **III.**
D. **I** and **II.**



I



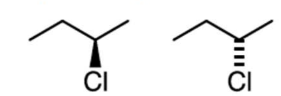
II



III

___ **10.** Indicate which of the following pair(s) of compounds represent a pair of enantiomers.

I



II

A. **I** only.
B. **II** only.
C. **I** and **II**.

___ **11.** Indicate which of the following pair(s) of compounds represent a pair of enantiomers.

I **II**

A. **I** only.
B. **II** only.
C. **I** and **II**.

___ **12.** Which of the following physical properties is/are different for compounds that are enantiomers of each other?

A. Boiling point.
B. Optical rotation.
C. Both boiling point and optical rotation.

___ **13.** Limited to the choices below, indicate which of the following compounds represent a pair of diastereomers.

I **II** **III** **IV**

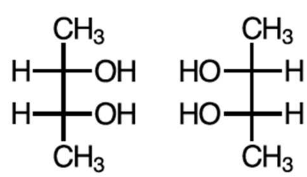
A. **I** and **II**.
B. **II** and **III**.
C. **III** and **IV**.

___ **14.** Which statement is true about the physical properties of diastereomers?

A. Boiling points are the same. Optical rotations are equal, but in opposite directions.
B. Boiling points are the same. Optical rotations are numerically different.
C. Boiling points are different. Optical rotations are numerically different.
D. Boiling points are different. Optical rotations are equal, but in opposite directions.

___ **15.** What is the relationship between the two compounds shown below?

A. Identical.
B. Enantiomers.
C. Diastereomers.



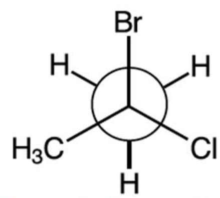
___ **16.** Which of the following can rotate plane-polarized light?

A. a 50:50 mixture of enantiomers.
B. a 50:50 mixture of diastereomers.
C. a meso compound.

___ **17.** In which direction does *(S)*-2-hexanol rotate plane polarized light?

A. Clockwise because it is dextrorotatory.
B. Counterclockwise because it has an *S*-configuration.
C. Cannot tell without experimental data.

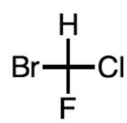
___**18.** Is the following molecule chiral?



A. Yes, because its mirror image cannot be overlaid on itself.
B. Yes, because it has a carbon atom with four different substituents.
C. No, because it is a meso compound.
D. No, because it does not have a carbon atom with four different substituents.

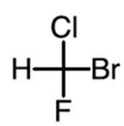
___ **19.** In the following Fischer projection, Cl is …

A. projecting toward you.
B. projecting away from you.
C. located in the plane of the paper.



___ **20.** Which atom(s) in the structure given lie(s) behind the plane of the page?

A. H and Br.
B. Cl and F.
C. Br only.
D. H only.



___ **21.** What is your major?

A. Chemistry
B. Non-chemistry (physical/natural science)
C. Non-chemistry (social sciences/humanities).
D. Other/undeclared.

___ **22.** What is your gender?

A. Male
B. Female

Appendix G

IRB APPROVALS

UNIVERSITY of
NORTHERN COLORADO

Institutional Review Board

DATE:	February 26, 2013
TO:	Alexey Leontyev, MS
FROM:	University of Northern Colorado (UNCO) IRB
PROJECT TITLE:	[428165-2] Misconceptions in Stereochemistry
SUBMISSION TYPE:	Amendment/Modification
ACTION:	APPROVAL/VERIFICATION OF EXEMPT STATUS
DECISION DATE:	February 26, 2013

Thank you for your submission of Amendment/Modification materials for this project. The University of Northern Colorado (UNCO) IRB approves this project and verifies its status as EXEMPT according to federal IRB regulations.

Alexey -

Hello and thank you for clear and concise changes to your approved IRB application for the study "Misconceptions in Stereochemistry".

Best wishes with your research.

Sincerely,

Dr. Megan Stellino, UNC IRB Co-Chair

We will retain a copy of this correspondence within our records for a duration of 4 years.

If you have any questions, please contact Sherry May at 970-351-1910 or Sherry.May@unco.edu. Please include your project title and reference number in all correspondence with this committee.

UNIVERSITY of
NORTHERN COLORADO

Institutional Review Board

DATE:	August 8, 2013
TO:	Alexey Leontyev
FROM:	University of Northern Colorado (UNCO) IRB
PROJECT TITLE:	[484510-1] Stereochemistry Instructional Survey
SUBMISSION TYPE:	New Project
ACTION:	APPROVAL/VERIFICATION OF EXEMPT STATUS
DECISION DATE:	August 8, 2013

Thank you for your submission of New Project materials for this project. The University of Northern Colorado (UNCO) IRB approves this project and verifies its status as EXEMPT according to federal IRB regulations.

Thank you for a clear concise IRB application.

Best wishes with your study.

Megan Stellino, UNC IRB Co-Chair

We will retain a copy of this correspondence within our records for a duration of 4 years.

If you have any questions, please contact Sherry May at 970-351-1910 or Sherry.May@unco.edu. Please include your project title and reference number in all correspondence with this committee.

This letter has been electronically signed in accordance with all applicable regulations, and a copy is retained within University of Northern Colorado (UNCO) IRB's records.

UNIVERSITY of
NORTHERN COLORADO

Institutional Review Board

DATE:	May 21, 2014
TO:	Alexey Leontyev
FROM:	University of Northern Colorado (UNCO) IRB
PROJECT TITLE:	[538339-2] Experts' Validation of the Stereochemistry Concept Inventory
SUBMISSION TYPE:	Amendment/Modification
ACTION:	APPROVAL/VERIFICATION OF EXEMPT STATUS
DECISION DATE:	May 19, 2014

Thank you for your submission of Amendment/Modification materials for this project. The University of Northern Colorado (UNCO) IRB approves this project and verifies its status as EXEMPT according to federal IRB regulations.

Hello Alex,

Thank you for making the requested modifications. I am recommending the approval of your proposal. I wish you well in your research.

Sincerely,

Wendy Highby

We will retain a copy of this correspondence within our records for a duration of 4 years.

If you have any questions, please contact Sherry May at 970-351-1910 or Sherry.May@unco.edu. Please include your project title and reference number in all correspondence with this committee.

This letter has been electronically signed in accordance with all applicable regulations, and a copy is retained within University of Northern Colorado (UNCO) IRB's records.

UNIVERSITY of
NORTHERN COLORADO

Institutional Review Board

DATE:	September 3, 2014
TO:	Alexey Leontyev
FROM:	University of Northern Colorado (UNCO) IRB
PROJECT TITLE:	[647768-1] Stereochemistry Concept Inventory Data Collection
SUBMISSION TYPE:	New Project
ACTION:	APPROVAL/VERIFICATION OF EXEMPT STATUS
DECISION DATE:	September 3, 2014

Thank you for your submission of New Project materials for this project. The University of Northern Colorado (UNCO) IRB approves this project and verifies its status as EXEMPT according to federal IRB regulations.

Hello Alexey,

Thank you for your carefully prepared IRB application. You are approved with no issues. Good luck on this important research.

Sincerely,

Nancy White, PhD, IRB Co-Chair

We will retain a copy of this correspondence within our records for a duration of 4 years.

If you have any questions, please contact Sherry May at 970-351-1910 or Sherry.May@unco.edu. Please include your project title and reference number in all correspondence with this committee.

This letter has been electronically signed in accordance with all applicable regulations, and a copy is retained within University of Northern Colorado (UNCO) IRB's records.

The University Of
TAMPA

To: Mr. Alexey Leontyev
From: Stephen Blessing, IRB
Re: 14-77 IRB Submission
Date: October 28, 2014

Your proposal 14-77, "Stereochemistry Concept Inventory Data Collection" has been approved via exempt review by the IRB for one calendar year.

Sincerely,

Stephen Blessing

Dr. Stephen Blessing
IRB Committee Chair

SYMBOL OF EDUCATIONAL EXCELLENCE
The University of Tampa | 401 W. Kennedy Blvd. | Tampa, FL 33606-1490 | (813) 253-3333 | www.ut.edu

NORTHERN ARIZONA UNIVERSITY

Institutional Review Board for the Protection of Human Subjects in Research

Northern Arizona University
PO Box 4087
Flagstaff, AZ 86011-4087

928-523-4236
928-523-1075 fax
www.research.nau.edu/vpr/IRB

October 4, 2014

Mr. Alexey Leontyev
Doctoral Student
Chemical Education Program
Dept. of Chemistry and Biochemistry
University of Northern Colorado
Box 98
Greeley, CO 80639

Re: Permission to conduct research at Northern Arizona University

Dear Mr. Leontyev:

I have reviewed your request regarding your research and I am pleased to support your research entitled "Stereochemistry Concept Inventory."

I understand that your research involves recruiting and surveying the stereochemistry knowledge of students at Northern Arizona University. This research will take place during the Fall 2014 semester in the classroom of Dr. Cindy Browder and has been reviewed and approved by the University of Northern Colorado Institutional Review Board.

Good luck with your research!

Regards,

Cynthia L. Johnson, BS, CRC, CRA
IRB Coordinator

BALL STATE UNIVERSITY

ACADEMIC AFFAIRS
OFFICE OF RESEARCH INTEGRITY

Muncie, Indiana 47306-0155
Phone: 765-285-5070/5034
Fax: 765-285-1624/1328

October 20, 2014

Alexey Leontyev
University of Northern Colorado

RE: Letter of Support from Ball State University Institutional Review Board

Dear Mr. Leontyev,

The Ball State University IRB has received your IRB approval letter from the University of Northern Colorado and your protocol for your study, *"Stereochemistry Concept Inventory Data Collection."*

Based on your IRB approval letter and letter of support by the BSU faculty, you have permission to recruit students at Ball State University.

If you have any further questions or concerns, please contact the Ball State University IRB.

Sincerely,

John M. Mulcahy, Jr.
Associate Director,
Office of Research Integrity
Ball State University
Muncie, IN 47306
765.285.5106
765.285-5052

Mount Saint Mary College
330 Powell Avenue, Newburgh, New York 12550 • 845-561-0800 • Fax: 845-562-6762
www.msmc.edu

To whom it might concern,

My name is Dr. Fahey from the Division of Natural Science at Mount Saint Mary College. This letter is to confirm that I am willing to work with Alexey Leontyev and Richard Hyslop from the University of Northern Colorado and assist them in data collection for their study titled "Development of a Stereochemistry Concept Inventory." The data collection will be conducted in an ethical manner and participants' confidentiality of responses will be assured.

If you have any questions regarding my involvement in the research project, feel free to contact either myself at

Dr. Jodie Fahey
330 Powell Ave.
Newburgh, NY 12550
845-569-3555

Sincerely,

Jodie Fahey

Dr. Jodie Fahey

NOTICE OF ACTION – PROTOCOL APPROVAL
Institutional Review Board (IRB) for the Protection of Human Subjects

University of Wisconsin – River Falls

Principal Investigator: Leontyev, Alex
Sponsor/Support: Peterson, Karl
Protocol Title: Stereochemistry Concept Inventory
Protocol Number: W2014 – T187
Committee Action: Approved on November 26, 2014, Expires on November 25, 2015

Dear Alex and Karl,

The Institutional Review Board of the University of Wisconsin – River Falls has reviewed your proposal and **approved** your study for a period of one year. Approval is based on identification that the study has met federal regulations set forth in 45 CFR 46.111.

Review of the project has identified that:

Top of Form

- ☑ Risks to subjects are minimized.
- ☑ Risks to subjects are reasonable in relation to anticipated benefits, if any, to subjects, and the importance of the knowledge that may reasonably be expected to result.
- ☑ Selection of subjects is equitable.
- ☑ Informed consent will be sought from each prospective subject or the subject's legally authorized representative, in accordance with, and to the extent required by §46.116.
- ☑ Informed consent will be appropriately documented, in accordance with, and to the extent required by §46.117.
- ☑ The research plan makes adequate provision for monitoring the data collected to ensure the safety of subjects.
- ☑ There are adequate provisions to protect the privacy of subjects and to maintain the confidentiality of data.

Bottom of Form

Conditions for approval:

- Any changes to the protocol, consent, recruitment, or data collection materials must be approved by the IRB before they are implemented.
- The IRB should be contacted immediately if there are any problems, adverse events, or new information that affect the risk to subjects.

Office of Grants and Research, 101 North Hall, 410 South Third Street
University of Wisconsin – River Falls, River Falls, WI 54022
Phone: (715) 425-3843, Fax: (715) 425-0649

- The study is approved for one year (365 days), starting from the date of approval. If the study is to continue beyond the expiration date, a request for renewal must be submitted to the IRB. This approval needs to be acquired before the current expiration date.
- Subjects taking part in the research should be provided with a copy of the consent form unless the requirement for a consent form has expressly been waived by the IRB.
- Consent documentation and IRB records should be retained by the researcher for at least 3 years after the completion of the project.

Good luck with your study.

Todd Wilkinson, Ph.D.
Chair, Institutional Review Board for Human Subjects

cc: IRB File

To:Leontyev, Alexey

Cc:Caroline Clower [CarolineClower@clayton.edu]

Proposal #: 20141029001 A. Leontyev,
Title of Study: Stereochemistry Concept Inventory Data Collection
Review: Initial Review
Completed By: Fran Norflus
Date: 10/29/2014
Dear Primary Investigator:
Your proposal entitled "Stereochemistry Concept Inventory Data Collection" has been contingently approved. The study will be about the development of a diagnostic instrument to study stereochemistry concept inventory. This study meets the criteria for exemption under 45 CFR 46.101(b)1.
Pending the following revisions, the proposal may be approved:
1. The investigator needs to submit his certificate for NIH training of human subjects.
2. A full copy of the instrument is needed (we only have question 1 as indicated on the last page of the full study).
3. The researcher is responsible for contacting and recruiting individuals for potential participation and cannot use a mass email.
This determination is based upon the following documents:
1. IRB letter #1
2. IRB full study
Research cannot begin until the proposal is fully approved.
Please submit the revised study along with any needed revisions. Your response to this letter is required by November 28, 2014 or your proposal will be considered withdrawn. Thank you, and please feel free to contact me at 678.466.4852 or at Fnorflus@clayton.edu if you have any questions.
Sincerely,
Fran Norflus
IRB Chair

THE UNIVERSITY OF
SOUTHERN
MISSISSIPPI

Matthew G. Donahue, Ph.D.
Assistant Professor of Chemistry and Biochemistry
The University of Southern Mississippi
118 College Drive #5043
Hattiesburg, MS 39406-0001
601.266.4166 (Voice)
614-203-1123 (Cell)
601.266.6075 (Fax)
matthew.donahue@usm.edu
www.usm.edu/chem

Mr. Alexey Leontyev
Department of Chemistry and Biochemistry
University of Northern Colorado
Campus Box 98
Greeley, CO 80639
419-308-9472
Alexey.leontyev@unco.edu

20-October-2014

Dear Mr. Leontyev,

This letter is to confirm that I am willing to work with you and Richard Hyslop from the University of Northern Colorado on study titled "Development of a Stereochemistry Concept Inventory." Having taken the online version of the study myself, I am confident that it will not cause any undue stress on my Fall 2014 organic chemistry 2 CHE 256 section. Since you have ensured that the data collection will be conducted in an ethical manner and participants' confidentiality of responses is assured, I endorse the use of this tool to test the student's understanding of stereochemistry.

If you have any questions regarding my involvement in the research project, feel free to contact me at the information listed above.

Best Regards,

Matthew G. Donahue 21.OCT.2014

Matthew G. Donahue

BUTLER UNIVERSITY

Institute for Research
and Scholarship

4600 Sunset Avenue
Indianapolis, Indiana 46208-3485
(317) 940-9766
Fax: (317) 940-9074
E-mail: birs@butler.edu
Web: http://www.butler.edu/birs

INSTITUTIONAL REVIEW BOARD

DATE:	December 5, 2014
TO:	Alexey Leontyev, Prof. Adam Azman
FROM:	Joel Martin Chair, Institutional Review Board
RE:	IRB Protocol
TITLE:	Stereochemistry Concept Inventory
SUBMISSION TYPE:	Cooperative Research Review

On behalf of Butler's Institutional Review Board, I am pleased to inform you that your application to accept the approval of another institution's IRB has been approved as of September 3, 2014, the date of the other IRB's approval. Since your study was approved as exempt, there will be no further review of your protocol, and you are cleared to begin the procedures outlined in your protocol. Although your study is exempt from a continuing review, you and your research team are not exempt from ethical research practices and should therefore employ all protections for your participants and their data which are appropriate to your project.

Please note the following conditions apply to all IRB approvals:
1. No subjects may be involved in any study procedure prior to the IRB approval date.

2. All unanticipated or serious adverse events must be reported to both the Butler IRB and the primary IRB within 5 days.

3. All protocol modifications must be IRB approved prior to implementation unless they are intended to reduce risk. This includes any change of investigator, or site address.

4. All protocol deviations must be reported to both the Butler IRB and the primary IRB within 5 days.

5. All recruitment materials and methods must be approved by the IRB prior to being used.

I offer my congratulations on your approval and wish you success on your research. Should you desire additional assistance or clarification, please call me at 940-9971 or email jmmarti1@butler.edu.

BYU
IDAHO

October 22, 2014

Dear Alex,

Your request to BYU-Idaho online students for the study entitled *Stereochemistry Concept Inventory Data Collection* is approved for 12 months from the date of this letter. The survey should not be administered to students under 18 years of age. Also, please ensure that your survey sample accurately reflects the target population.

Please notify the IRB if you intend to make any significant modifications to the study's design or implementation.

As always, good luck with your study and thanks for the continued notification about changes and modifications to your study.

Regards,

Scott J. Bergstrom

Scott J. Bergstrom, Ph.D.
Chair, BYU-Idaho Institutional Review Board

To: Alexey Leontyev

From: KatieAnn Skogsberg IRB Chair

Date: Oct 20, 2014

Re: Conditional exempt status for
Proposal Name: Stereochemistry concept inventory data collection
Proposal Number: University of Northern Colorado [647768-1]

Greetings Alexey,

Thank you for submitting your information regarding the study "*Stereochemistry concept inventory data collection.*" I have received the documentation verifying that the IRB committee at your institution (UNCO) has reviewed your proposal and that it meets their "except" status requirements. My review of your proposal determines that this project is EXEMPT from further regulatory oversight by Centre College as long as it continues to meet the following conditions:

1. The data collected does not pose a risk to your participants legally (e.g. revealing drug use or other violations of the law), economically (e.g. revealing anything that may prevent them from obtaining, or keeping a job), or socially (e.g. revealing sexual preferences or other socially sensitive information without their permission).
2. The data you collect remains confidential, is stored in a secure location, and is destroyed after the requisite period of time (3 years).
3. You may not modify your project significantly. Changes may trigger the need for reevaluation of this exemption by the IRB. We advise, therefore, that you consult with me or a member of the IRB Staff in planning any changes.
4. You are obligated to notify the IRB immediately if you significantly modify your protocol, or if adverse events occur during the conduct of this research project.
5. This approval is valid only for a period of one calendar year. Should your project extend beyond that time period, you will need to obtain an approval for continuation.

Thank you for your cooperation in the Human Research Review Process. Please accept our best wishes for successful completion of this research project. If you have any questions, please contact KatieAnn Skogsberg (katieann.skogsberg@centre.edu), (859) 238-5238.

KatieAnn Skogsberg
Centre College IRB Chair

Centre College IRB
Assurance #FWA00017871
IRB approval code UNCO 647768-1
Approval date 10-20-14
Valid through 10-14-15
IRB Chair

To: Alexey Leontyev & Jennifer Muzyka

From: KatieAnn Skogsberg IRB Chair

Date: May 7, 2015

Re: Conditional exempt status for a REVISED proposal
Proposal Name: Stereochemistry concept inventory data collection
Proposal Number: University of Northern Colorado [647768-1]

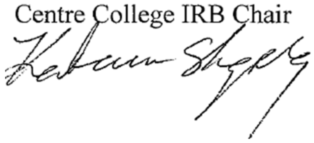
Greetings Alexey,

Thank you for submitting your information regarding revisions to the study "*Stereochemistry concept inventory data collection.*" Specifically, you have indicated that you have changed to an online versus a paper version of your survey. Our review of this change indicates that it does not alter the risk level to the participants and that your study continues to be EXEMPT from further regulatory oversight by Centre College as long as it continues to meet the following conditions:

1. The data collected does not pose a risk to your participants legally (e.g. revealing drug use or other violations of the law), economically (e.g. revealing anything that may prevent them from obtaining, or keeping a job), or socially (e.g. revealing sexual preferences or other socially sensitive information without their permission).
2. The data you collect remains confidential, is stored in a secure location, and is destroyed after the requisite period of time (3 years).
3. You may not modify your project significantly. Changes may trigger the need for reevaluation of this exemption by the IRB. We advise, therefore, that you consult with me or a member of the IRB Staff in planning any changes.
4. You are obligated to notify the IRB immediately if you significantly modify your protocol, or if adverse events occur during the conduct of this research project.
5. This approval is valid only for a period of one calendar year. Should your project extend beyond that time period, you will need to obtain an approval for continuation.

Thank you for your cooperation in the Human Research Review Process. Please accept our best wishes for successful completion of this research project. If you have any questions, please contact KatieAnn Skogsberg (katieann.skogsberg@centre.edu), (859) 238-5238.

KatieAnn Skogsberg
Centre College IRB Chair



May, Sherry

From: May, Sherry
Sent: Monday, October 27, 2014 1:19 PM
To: 'Cynthia Johnson'
Cc: White, Nancy
Subject: RE: Leontyev (647768-1) - Request for Confirmation of IRB Approval

Dear Cynthia,

Yes, Alexey Leontyev's project entitled, *Stereochemistry Concept Inventory Data Collection*, has been approved/verified Exempt according the federal IRB regulations by the University of Northern Colorado (UNCO) IRB.

Best,

Sherry May
Electronic Research Specialist
IRB Administrator

University of Northern Colorado
Campus Box 143
Kepner #25
Greeley, CO 80639
Office: 970-351-1910
Fax: 970-351-1934
http://www.unco.edu/osp

UNIVERSITY of NORTHERN COLORADO Bringing education to life.

From: Cynthia Johnson [mailto:C.Johnson@nau.edu]
Sent: Saturday, October 25, 2014 12:41 PM
To: May, Sherry
Cc: Institutional Review Board
Subject: Leontyev (647768-1) - Request for Confirmation of IRB Approval

Dear Ms. May,

One of UNC's doctoral students, Alexey Leontyev, has contacted our IRB for permission to conduct research activities on our campus. Mr. Leontyev has emailed a copy of your IRB's Exempt approval letter, but we would prefer confirmation of approval directly from your IRB. Could you verify that the referenced study is still approved as Exempt? An email would be fine.

Thank you for your kind assistance.

Regards,

Cynthia Johnson
IRB Coordinator
Northern Arizona University
Box 4137, Flagstaff, AZ 86011
Office: 928-523-2895
c.johnson@nau.edu

Sent:Thursday, October 30, 2014 1:36 PM

To: Leontyev, Alexey

Cc: Hyslop, Richard

October 30, 2014

Alexey Leontyev

c/o Dr. Richard Hyslop

Department of Chemistry & Biochemistry

University of Northern Colorado

Protocol Number: 647768-1

Project Title: Stereochemistry Concept Inventory

Dear Mr. Leontyev and Dr. Hyslop:

The University of Nebraska at Kearney's IRB has reviewed your request to honor the University of Northern Colorado's Institutional Review Board's review and approval of the aforementioned project. We will honor this request with the understanding that primary jurisdiction for oversight of any data collection on the UNK campus will be with the UNK IRB.

It is understood this project will be conducted in full accordance with all applicable sections of the IRB Guidelines. It is also understood that the IRB will be immediately notified of any proposed changes that may affect the status of your research project.

Good luck with your study,

Kathy Zuckweiler

Kathryn M. Zuckweiler, Ph.D., IRB Director

Janna Shanno, IRB Coordinator

University of Nebraska at Kearney

Institutional Review Board

Founders Hall Suite 1000

Kearney, NE 68849

(308) 865-8843

unkirb@unk.edu

To whom it might concern,

My name is Professor Pete Golden from the Department of Chemistry at Sandhills Community College in Pinhurst, NC. This letter is to confirm that I am willing to work with Alexey Leontyev and Richard Hyslop from the University of Northern Colorado and assist them in data collection for their study titled "Development of a Stereochemistry Concept Inventory." The data collection will be conducted in an ethical manner and participants' confidentiality of responses will be assured.

If you have any questions regarding my involvement in the research project, feel free to contact either myself at

Professor Pete Golden

Sandhills Community College

3395 Airport Rd.

Pinehurst, NC 28374

Sincerely,

ST. MARY'S UNIVERSITY

To whom it might concern,

My name is Dr. Dmitry Khon from the Department of Chemistry at St. Mary University. This letter is to confirm that I am willing to work with Alexey Leontyev and Richard Hyslop from the University of Northern Colorado and assist them in data collection for their study titled "Development of a Stereochemistry Concept Inventory." The data collection will be conducted in an ethical manner and participants' confidentiality of responses will be assured.

If you have any questions regarding my involvement in the research project, feel free to contact either myself at

Dmitriy Khon, Ph.D.
Assistant Professor
Department of Chemistry and Biochemistry
St Mary's University
Garni Science Hall, Room 105
One Camino Santa Maria
San Antonio, TX 78228
210-436-3740
dkhon@stmarytx.edu

Sincerely,

Dr. Dmitry Khon

DEPARTMENT OF CHEMISTRY AND BIOCHEMISTRY ■ ONE CAMINO SANTA MARIA ■ SAN ANTONIO, TEXAS 78228-8521
OFFICE (210) 436-3237

Armstrong
STATE UNIVERSITY

Department of Chemistry & Physics | College of Science & Technology
11935 Abercorn Street | Savannah, Georgia 31419
912.344.3219 | Fax 912.344.3432 | armstrong.edu

10/24/14

To Whom it May Concern,

My name is Suzanne Carpenter in the Department of Chemistry and Physics at Armstrong State University. This letter is to confirm that I am willing to work with Alexey Leontyev and Richard Hyslop from the University of Northern Colorado and assist them in data collection for their study entitled, "Development of a Stereochemistry Concept Inventory." The data collection will be conducted in an ethical manner and participants' confidentiality of responses will be assured.

If you have any questions regarding my involvement in the research project, feel free to contact me.

Sincerely,

Suzanne Carpenter
Associate Professor of Chemistry

Bakersfield College
1801 Panorama Drive
Bakersfield, CA 93305

Bakersfield COLLEGE

December 3, 2014

To whom it might concern,

My name is Dr. Kenward Vaughan from the Department of Chemistry at Bakersfield College. This letter is to confirm that I am willing to work with Alexey Leontyev and Richard Hyslop from the University of Northern Colorado and assist them in data collection for their study titled "Development of a Stereochemistry Concept Inventory." The data collection will be conducted in an ethical manner and participants' confidentiality of responses will be assured.

If you have any questions regarding my involvement in the research project, feel free to contact me via the address, phone, or email noted herein.

Sincerely,

Kenward S. Vaughan

Kenward Vaughan, Ph.D.
Professor of Chemistry
Chair, Department of Physical Science

(661) 395-4243
kvaughan@bakersfieldcollege.edu

BALL STATE UNIVERSITY

College of Sciences and Humanities
Department of Chemistry

Muncie, Indiana 47306-0445
Phone: 765-285-8060

To whom it may concern:

My name is Dr. Ryan Jeske from the Department of Chemistry at Ball State University. This letter is to confirm that I am willing to work with Alexey Leontyev and Richard Hyslop from the University of Northern Colorado and assist them in data collection for their study titled "Development of a Stereochemistry Concept Inventory." The data collection will be conducted in an ethical manner and participants' confidentiality of responses will be assured.

If you have any questions regarding my involvement in the research project please contact me at the address below.

Sincerely,
Ryan Jeske

Assistant Professor,
Ball State University Department of Chemistry
CP 305E
765-285-8078
rjeske@bsu.edu

RE: IRB Approval (stereochemitry test)
JoAnn Johnson [joann.johnson@usm.edu]
You replied on 10/22/2014 11:50 AM.
Sent:Wednesday, October 22, 2014 10:24 AM
To: Leontyev, Alexey
Hi Alexey,

I received the permission letter from Dr. Donahue, you're good to begin your survey.

Thank you,
Jo Ann Johnson

-----Original Message-----
From: Leontyev, Alexey [mailto:Alexey.Leontyev@unco.edu]
Sent: Tuesday, October 21, 2014 3:32 PM
To: Matthew Donahue; JoAnn Johnson
Subject: RE: IRB Approval (stereochemitry test)

Jo Ann,

Do I need to provide you with anything else? I have attached my IRB and approval. Dr. Donahue will provide students with a link to the electronic version of the survey.

Thanks.
Alexey
Alexey Leontyev
Doctoral student,
Chemical Education Program
Department of Chemistry and Biochemistry University of Northern Colorado Campus Box 98 Greeley, CO 80639
Phone: 419-308-9472
Email: alexey.leontyev@unco.edu

FW: Stereochemistry test
Joe Galusha [Joe.Galusha@wallawalla.edu]
Sent:Wednesday, December 03, 2014 9:39 AM
To: Leontyev, Alexey
Cc: Kyle Craig [Kyle.Craig@wallawalla.edu]
Dear Alexey,

I am pleased to inform you that the WWU EIRC has approved your research project to be done on our campus. Please make the arrangements necessary for this to happen. I believe Dr. Craig will be your point of contact.

Again, congratulations on this successful step in your program.

Joe Galusha, Chair
WWU EIRC
-----Original Message-----
From: Leontyev, Alexey [mailto:Alexey.Leontyev@unco.edu]
Sent: Friday, November 21, 2014 10:15 AM
To: Kyle Craig
Cc: Joe Galusha
Subject: RE: Stereochemistry test

Please find my application packet attached. I have also included original IRB from the University of Northern Colorado along with copy of the instrument, consent form, and approval letter.

The electronic copy of the instrument is available here:
https://unco.co1.qualtrics.com/SE/?SID=SV_1NBCWh2PjdbffxP

Feel free to browse but please put in the last question that you are not a student so I can remove this entry from a dataset.

Thanks,

Alexey Leontyev
Doctoral student,
Chemical Education Program
Department of Chemistry and Biochemistry University of Northern Colorado Campus Box 98 Greeley, CO 80639
Phone: 419-308-9472
Email: alexey.leontyev@unco.edu

WITTENBERG UNIVERSITY INSTITUTIONAL REVIEW BOARD PETITION

Date:	October 17, 2014
Principal Investigator(s):	Justin Houseknecht
	Pete Hanson
Phone Number(s):	937-327-6437
	937-327-6455
Email Address(es):	jhouseknecht@wittenberg.edu
	phanson@wittenberg.edu
If the PI is a student, identify the faculty member supervising the project:	
Title of Research Project:	Stereochemistry Concept Inventory pilot

Please answer all questions:

1. Summarize the objects of your research:

   "The purpose of this study is to validate the Stereochemistry Concept Inventory" developed by Alexey Leontyev in the Hyslop group at the University of Northern Colorado. Specifically, the question addressed will be: "Can the Stereochemistry Concept Inventory prodce reliable data and valid inferences for the assessment of important concepts of stereochemistry?" (quotes from IRB application of Alexey Leontyev, University of Northern Colorado)

2. Describe the subjects to be used, the selection process by which the subjects were chosen and the procedures to be followed. (This should include a summary of instructions to the subjects, and a description of their tasks).

   All students in Chem 201 will be asked to complete a paper version of the Stereochemistry Concept Inventory developed by the Hyslop group at the University of Northern Colorado. This should take 15-20 minutes during a laboratory session. Completed surveys will be faxed to the Hyslop group where they will be scored by Alexey Leontyev. We will be provided with summary scores and topics that the instrument identified as most challenging for my students.

From: Ralph Lenz
Sent: Sunday, October 26, 2014 4:26 PM
To: Justin B. Houseknecht; Peter E. Hanson; Jeff A. Ankrom; Nancy S. Woehrle; Regina A. Post; feltz@deltapsychologycenter.com; Ralph Lenz
Cc: June A. Viers
Subject: irb houseknecht-hanson

Hi Justin and Pete,

Your materials are in order and I see no problems

with expediting your request. So on behalf of the

Witt IRB I am communicating our approval of your

petition.

Good luck with this project.

Ralph

Witt IRB chair

From: Justin B. Houseknecht
Sent: Wednesday, October 22, 2014 9:46 AM
To: Ralph Lenz
Cc: Peter E. Hanson
Subject: Chem 201 IRB petition

Ralph,

Pete and I would like to pilot a concept inventory in Chem 201 this Fall. If there are concerns or omissions in the attached documents please let us know.

Thanks,

Justin