June 2018

# Survey research in postsecondary chemistry education: Measurements of faculty members' instructional practice and students' affect

Rebecca E. Gibbons
*University of South Florida*, revaclav@gmail.com

# Survey research in postsecondary chemistry education:

# Measurements of faculty members' instructional practice and students' affect

by

Rebecca E. Gibbons

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Chemistry
College of Arts & Sciences
University of South Florida

Major Professor: Jeffrey R. Raker, Ph.D.
Jennifer E. Lewis, Ph.D.
Robert F. Dedrick, Ph.D.
Theresa Evans-Nguyen, Ph.D.

Date of Approval:
June 12, 2018

Keywords: Faculty beliefs, Psychometrics, Measurement, Invariance, Affective learning

# DEDICATION

"You is kind. You is smart. You is important." – Kathryn Stockett, *The Help*.

This work is dedicated to those who enabled me to become who I am today. Most importantly, this work is dedicated to my mother and father, Jennifer and Donald Vaclav, who instilled in me a love of learning and embodied wisdom to me throughout my life.

## ACKNOWLEDGEMENTS

"I am driven by two main philosophies: Know more about the world today than I knew yesterday and lessen the suffering of others." -Neil DeGrasse Tyson.

The most important goal we can have as people, in my opinion, is to grow and gain perspective. Dr. Theresa Evans-Nguyen, thank you for always providing the perspective that I needed to make this contribution. Thank you also for listening to my responses from a million miles outside of your field during questioning and always keeping it real. I know that you are going to go on at USF to inspire the next great generation of women in chemistry! Dr. Robert Dedrick, thank you for teaching me (essentially) everything I know about how to do education research. You are a profoundly kind instructor, and I appreciate not only the classroom learning that you facilitated during my time here, but that your insights upon each committee meeting enabled me to grow as a researcher and a person. Dr. Jennifer Lewis, thank you for pushing me to my limit. Whenever I started to become complacent with my knowledge, you reminded me to stay humble and keep reading, writing, and learning.

My gratitude to my major professor, Dr. Jeffery Raker, knows no end. You taught me how to be a scholar as well as a human being, how to use a semicolon, how to not cry during my defenses, how to let it go, how to take and give feedback, how to remain kind while staying firm, and most importantly, how to contribute. This dissertation is as much yours as it is mine, and I'm both grateful to you and proud of you. I'm lucky to have been on this journey with you.

“The function of education is to teach one to think intensively and to think critically. Intelligence plus character – that is the goal of true education.” – Dr. Martin Luther King, Jr.

I cannot genuinely say that I achieved this degree without acknowledging how I got here: the Department of Chemistry, Biochemistry, and Physics at the University of Tampa, especially Dr. Eric Ballard, Dr. John Struss, Dr. Eric Werner, and Dr. Tom Jackman. The faculty and my colleagues during my time there are the reason I started on this path and I will always be indebted to them. They not only taught me chemistry, but taught me how to approach the world as a scientist.

Throughout my doctoral degree, I always had a “real job” in my teaching assistantship. To Dr. Scott Lewis and Dr. Rong Zhang, thank you for being the supervisors I did not deserve for these past two years. You’ve provided me with perspective, guidance, and inspiration, and I will always be grateful. You demonstrate grace under pressure and somehow manage to accomplish the enormous feat of facilitating and measuring learning among a large group every semester- I admire you both!

The chemical education research groups at the University of South Florida have been such an important part of my growth. I entered the program not knowing what being a chemical education researcher meant, but knowing that I care and wanted to learn. I have been blessed to have been able to become the scientist I am today by learning from a group of people who bring perspectives from every walk of life.

“Love isn’t a state of perfect caring. It is an active noun like struggle. To love someone is to strive to accept that person exactly the way he or she is, right here and now.”- Mr. Rogers.

To my husband William, I know that it was a struggle to love me for these past three years, and I appreciate you for it every day. To Peaches & Cream and Sammy, my 12 paws, thank you for always being there with a wet nose to remind me to come home, take a break, and enjoy the process.

TABLE OF CONTENTS

## LIST OF TABLES

LIST OF FIGURES

## ABSTRACT

Collection of data through survey-type measurements and analysis contributes rich, meaningful information to the chemical education research enterprise. This dissertation reports two strands of research that each contribute a “snapshot” of the state of chemical education on two different levels. The first uses survey research methods, collecting data from faculty members to learn about postsecondary chemistry education across the United States. The second uses survey instruments of student achievement emotions within the organic chemistry classroom, collecting longitudinal data to learn about the relationships of emotions with achievement over time. Both areas are of interest because chemical education research produces evidence-based instructional practices as well as survey instruments of student characteristics, many of which are ready to be used in classroom, yet there is a recognized disconnect between development of these products and enacted practices. The research in this dissertation improves upon previous methodology in both strands of research included while reporting data with implications for instructional, research, and policy matters.

A national survey of postsecondary chemistry faculty uses a stratified sampling procedure to gather information about the state of education in chemistry classrooms. The use of the teacher-centered systemic reform model of educational change enables us to use the data collected in the survey to gather empirical support of the relationship between faculty members’ beliefs about how students learn chemistry more effectively, faculty members’ self-efficacy for instruction and chemistry content, and the instructional practices that they utilize in the course for which they felt they had the most influence. This information is paramount for the developers of evidence-based instructional practices as well as parties interested in determining the methods best suited to the dissemination of these tools. Professional development activities designed to inspire the use of evidence-based instructional tools or techniques must acknowledge the belief

systems of faculty members and the need for change in these beliefs prior to the incorporation of new methods. These results present a call for reform efforts on fostering change from its core, i.e., the beliefs of those who ultimately adopt evidence-based instructional practices. Dissemination and design should incorporate training and materials that highlight the process by which faculty members interpret reformed practices within their belief system, and explore belief change in the complex context of education reform.

Another example of the use of national survey data is the determination of the niche distribution of classroom response systems, also known as clickers. It is determined in this study that clickers are used more often in large courses taught at the lower level across the United States. This niche is deemed a more suitable situation for the use of clickers than others. This information is important for researchers developing tools intended for use within the classroom. Despite the possibility for use in all contexts, the national population of faculty members will adopt tools in the contexts which are deemed most suitable; the niche markets of educational tools can provide insight in to best development practices also well as direction for the optimization of the experience for the most frequent users of these tools.

The other set of studies in this dissertation utilize the control-value theory of achievement emotions in the postsecondary organic chemistry context to explore nuanced relationships of affect with achievement. These studies utilize a longitudinal panel data collection mechanism, enhancing our ability to understand relationships. The control-value theory posits that there are a set of nine achievement emotions, dictated by control and value, which influence achievement. Two of these achievement emotions, anxiety and enjoyment, are determined in one study to fluctuate over the semester of organic chemistry and significantly influence achievement as measured by examination scores. These are supported by their theoretical interpretation as activating emotions, and when experienced, inspire students to take measures that ultimately either increase or reduce their success. A deactivating emotion, boredom, is measured in another study and found to also hold a reciprocal relationship with achievement when measured over time. In both studies, results show that the reciprocal causation model with an exam snowballing effect best fits data among the alternative models. There is a small and significant negative relationship between anxiety and performance contrasted with a positive relationship between enjoyment and performance throughout the

semester. Negative relationships were observed between boredom and examination performance across the term. In addition, relationships were observed to be stronger at the beginning of the course term. Future research should consider achievement emotions in light of educational reforms to ensure that innovative curricula or pedagogies are functioning in the classroom as intended.

# CHAPTER ONE:

# INTRODUCTION

The chemical education research (CER) enterprise, as one facet of discipline-based education research (DBER), seeks to improve the teaching and learning of chemistry in the interest of increasing both the quality and quantity of chemistry graduates (for more on DBER; see National Research Council, 2012). One success of CER is the development of evidence-based instructional practices (EBIPs, sometimes called research-based instructional strategies). These include process-oriented guided inquiry learning, in which students proceed through the learning cycle of exploring data, conceptual creation, and application (see Moog & Spencer, 2008, for more information). Another example is peer-led team learning, through which students who recently completed a course serve as instructional aides (see Wilson & Varma-Nelson, 2016, for more information). The flipped classroom is another EBIP methodology within which content delivery is conducted outside of the classroom, and classroom time is utilized primarily for students to engage in problem-solving with the assistance and supervision of the instructor (see Seery, 2015, for more information). The success of these EBIPs and others in increasing the achievement and attitudes of chemistry graduates demonstrates that CER is reaching toward the goal of improving the teaching and learning of chemistry (National Research Council, 2012).

The development and success of EBIPs is important, but, we cannot lose sight of the student experience, particularly in the context of achieving the goal of producing a greater quantity of higher quality chemistry graduates. Through CER, we have gained an understanding of the challenges facing learners of chemistry (Johnstone, 2000) as well as some of their common misconceptions (see Bodner, 1991, for a classic example of this type of research). Challenges and misconceptions are experienced in the classroom along with other non-cognitive factors, known as affective states; affect is a key piece of the theory of meaningful learning, which posits that affect is joined by cognitive and psychomotor learning in a trifold model (Bretz, 2001; Novak, 2002). Affective states have been recognized as related to achievement in many postsecondary chemistry contexts (e.g., Chan & Bauer, 2014, 2016b; Lewis, Shaw, Heitz, & Webster, 2009; Liu, Raker, & Lewis, 2018; Xu, Villafañe, & Lewis, 2013). CER has successfully developed instrumentation designed to measure common misconceptions, known as the concept inventories (e.g., Brandriet & Bretz, 2014; Bretz & Linenberger, 2012; Bretz & Murata Mayo, 2018; Luxford & Bretz, 2014; McClary & Bretz, 2012), along with instruments to measure affective states (e.g., Bauer, 2005; Bauer, 2008; Dalgety, Coll, & Jones, 2003; Liu, Ferrell, Barbera, & Lewis, 2017; Xu & Lewis, 2011). The impact of students' affective experiences on their performance in postsecondary chemistry cannot be ignored, and research in this area is important for furthering our understanding of the way learning works in our postsecondary chemistry classrooms.

While EBIP development and research on affect in chemistry have both been productive, there is a noticeable disconnect between research and enacted practice. Dissemination efforts vary, yet there is a lack of adoption of many EBIPs and instruments measuring non-cognitive factors into practice in postsecondary STEM classrooms (DeHaann, 2005; National Research Council, 2012). Recent research has indicated that despite evidence to support EBIPs as successful in increasing learning, they are not widely utilized (Stains et al., 2018). Because there is evidence both for (Lewis et al., 2009; Qureshi, Vishnumolakala, Southam, & Treagust, 2017; Vishnumolakala, Southam, Treagust, Mocerino, & Qureshi, 2017) and against (Chan & Bauer, 2015; Chase, Pakhira, & Stains, 2013) the ability of EBIPs to improve student affect, it is difficult to discern from the body of CER literature whether affect is optimized in classrooms. Based on the evidence for ineffective adoption of research into the "real world," this dissertation seeks to provide more information about the products of CER: Are they being used outside of the developers' classrooms? Can we help students learn better by understanding affective states? Production of improvements in the chemistry classroom is essential for progressing towards the goals of CER, but without a snapshot of the state of these programs and projects we cannot understand their impact. Determining how well the products of past research are functioning is an important step forward for CER.

This dissertation seeks to provide basic research for the field of CER through the application of two strands of research. The first is the evaluation of data collected through a national survey of postsecondary chemistry faculty, addressed in Chapter Two (Gibbons, Villafañe, Stains, Murphy, & Raker, 2018) and Chapter Three (Gibbons et al., 2017). The second is an evaluation of the relationship of affect with achievement in the organic chemistry classroom, addressed in Chapter Four (Gibbons, Xu, Villafañe, & Raker, 2018) and Chapter Five (unpublished work). The application of advanced methodology enables this work to glean information with implications for future research, instruction, and policy matters. Snapshot studies such as those included in this dissertation are essential for providing a baseline on which future work in CER can expand and improve.

**Methods Overview**

The studies included in this dissertation utilize quantitative research methods, providing empirical evidence for the direction and strength of relationships between variables. A survey research methodology is employed in Chapters Two and Three. This method is designed to collect data from a subset of the population of interest and determine distributions of the characteristics measured in the overall population (see Weisberg, Krosnick, & Bowen, 1996, for information on survey research methods), this research also explores relationships between observed variables within the sample population. The research reported in Chapters Four and Five employs affective survey instruments within the classroom (see American Educational Research Association, American Psychological Association, National Council on Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing, 2014, for information on instrument design and use). These studies are designed to measure student achievement emotions (Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011) and employ factor and structural analytic models to the data collected. This method allows us to explore the nuance in direction and strength of the relationships of achievement emotions with performance on examinations.

In the case of the four studies in this dissertation, a compromise is made between the desire to collect data from a wide sample of individuals and to collect the most accurate data possible by collecting self-report data from participants. While the validity and reliability of self-report data has been questioned when studying instructional practices (D'Eon, Sandownik, Harrison, & Nation, 2008; Ebert-May et al., 2011; Herrington, Yezierski, & Bancroft, 2016), this dissertation will provide evidence for the accuracy of some self-report data based on association with observational data in Chapter Two. Another potential problem arising from the self-report methodology employed is nonresponse bias (see Groves, 2006, for more), which we avoid by using weighted data analysis in Chapter Three. In self-reported measures of affect, problems such as social desirability bias, in which individuals respond according to how they believe the researcher expects them to respond, are potentially problematic (Krumpal, 2013). Our efforts to avoid social desirability bias in the studies reported in Chapters Four and Five include anonymizing data and reporting in aggregate, which have been found to limit the effect of social desirability on data compared to identified and individual-based applications of survey instruments (Nederhof, 1985). Importantly, we do not conduct our analysis in ignorance of the limitations of self-report data. We account for error in our measurement using classical test theory models in the analyses in Chapters Four and Five (see Crocker & Algina, 1968, for more information on classical test theory), and we use weighted survey data analysis in Chapter Three to account for survey sampling error.

### Survey research methods

It has been noted through systematic research that there is a disconnect between research-based recommendations for instruction and enacted practice (DeHaann, 2005; National Research Council, 2012), and studies in CER have addressed this by demonstrating interest in the national status of instruction and assessment using survey methodology. One example of this work is a needs assessment survey, in which it was determined that most faculty members recognize their department making efforts at enhanced assessment in the interest of informing accreditation bodies or for institutional improvement (Emenike, Schroeder, Murphy, & Holme, 2011, 2013). In another study, it was discovered that faculty members are generally not familiar with the terminology used by education experts in relation to assessment (Raker, Emenike, & Holme, 2013; Raker & Holme, 2014). Jargon may limit faculty members from successfully using enhanced assessment as encouraged by their departments. Ultimately, this may be a limitation on the success of such enhanced assessment efforts. Chapters Two and Three of this dissertation are an extension of this strand of research.

To successfully achieve the research goal of quantitative evaluation of the state of chemical education on the national scale, the studies included here improve on previous methodology. One example of an improvement is the incorporation of a theoretical framework to direct the creation of the survey and guide data analysis (Abraham, 2008). Theoretical frameworks found in the literature provide us with insight into the nature of the characteristics of interest in the population. A priori selecting a theoretical framework as an analysis plan guides and increases the accuracy of our quantitative research. The teacher-centered systemic reform model of educational change (TCSR; Woodbury & Gess-Newsome, 2002) was used in the development of the survey in Chapters Two and Three. The TCSR also guides the data analysis for Chapter Two. In Chapter Three, the technology adoption life cycle (TALC; Rogers, 1995) guides the data analysis. Both the TCSR and TALC are described in detail within the main body text.

Another improvement from previous national surveys in CER is the sampling technique utilized here. There are a variety of methods with which one can administer a survey on the national level, including probability and non-probability techniques. Probability techniques are defined as sampling methods in which there is a known non-zero probability that each member of the population will be selected for participation in the study and sampling error estimates can be calculated. Non-probability techniques are best described through an example: quota sampling is a non-probability sampling method in which sampling is conducted until a desired value is filled without a priori evaluation of the likelihood of selection

(Blair, Czaja, & Blair, 2014a). The survey analyzed in Chapters Two and Three uses stratified sampling, a probability procedure in which we assume that there is a natural categorizing characteristic in the population, and each individual belongs to at least one and exclusively one stratum (Blair et al., 2014a). It is from these strata that sampling is conducted, and unit response rates are calculated based on the inverse probability of selection of the respondents from each stratum.

To conduct the survey, we identified the group of faculty members who teach chemistry at the postsecondary level through data obtained from the United States Department of Education's National Center for Education Statistics Integrated Postsecondary Education Data System (IPEDS). IPEDS houses information on undergraduate and graduate degrees awarded in the United States annually. To define the population, we first collected information on each institution which had conferred at least one Bachelor's degree in chemistry in the five years preceding the survey. The websites for each institution provided information on the number and titles of faculty members within the department of chemistry, and the overall population consists of these faculty members.

The stratified sampling procedure requires that strata are established in the population before the sample is selected. Postsecondary institutions are naturally stratified in the United States; some colleges and universities are controlled by private entities and some are public. This distinction is important because cultural differences based on the locus of control of the institution have been cited as barriers or enablers of the adoption of reform like EBIPs (Cox, McIntosh, Reason, & Terenzini, 2011; DeHaann, 2005), particularly regarding the tenure and promotion reward system, which differs between public and private institutions (Shadle, Marker, & Earl, 2017). We additionally stratify the institutions based on their highest chemistry degree awarded. We added this stratification variable because an institution's highest degree awarded is associated with institutional culture as related to improved instruction (Cox et al., 2011). Therefore, the survey analyzed in Chapters Two and Three was sampled from a set of six strata: public institutions which offer Bachelor's, Master's, and Doctoral degrees as well as private institutions which offer Bachelor's, Master's, and Doctoral degrees.

To effectively calculate sampling error using known parameters after strata are established, a sample was selected based on the sample size required for a 95% confidence level and 5% confidence interval, assuming a 25% unit response rate (Blair, Czaja, & Blair, 2014b). The survey was administered via Qualtrics (online) and unit response rates were calculated upon completion of submissions. The data were analyzed using their calculated final population weights in Stata software (StataCorp, 2015). This technique ultimately allowed us to measure the precision of our measurements by reporting confidence intervals from each analysis. In this way, the improved methodology of the sampling strategy used here enables us to declare the quantitative results of our survey with greater confidence.

**Affective survey instrument methods**

Students' affective states (i.e., emotional experiences and other non-cognitive states, like self-efficacy, all of which are unobservable traits known as constructs) are influential in the determination of success. There is a recognized relationship of various affective states with achievement throughout the chemistry curriculum (e.g., Lewis et al., 2009; Villafañe, Xu, & Raker, 2016), and the study of the relationships of affect with achievement is of interest to classroom instructors (and has been for many years; see Larsen, 1986). Ideally, instructional practices designed to increase achievement can be enhanced by increasing students' affect because it has been determined that increased affect can lead to enhanced achievement (Frenzel, Goetz, Ludke, Pekrun, & Sutton, 2009; Linnenbrink, 2006).

Chapters Four and Five of this dissertation provide evidence for the nuanced relationship between affect and achievement in postsecondary organic chemistry. These studies are framed by the control-value theory (CVT) of achievement emotions (Pekrun, 2000). The CVT posits that there are nine achievement emotions which influence achievement in the classroom: enjoyment, hope, pride, relief, anxiety, anger, shame, boredom, and hopelessness. These emotions are dictated by a student's control over the content and their learning as well as the value that they place on the content. These emotions subsequently dictate the level of activation that a student experiences in their cognitive domain, affective domain, and physiological domain (see Pekrun, Hall, Goetz, & Perry, 2014 for an example of the nature of these domains referencing boredom), and therefore are influential in understanding the factors leading to achievement.

To study CVT in the context of postsecondary organic chemistry in this dissertation, affective survey instruments were administered in the classroom. In this case, a survey is a tool (much like the instruments used to discern the components of samples in benchtop chemistry) used to measure certain characteristics of interest (for an introduction to the development of such scales, see DeVelis, 2017). These instruments are often designed with response options like the traditional Likert (1932) scale, and therefore provide a quantitative picture of the construct(s) of interest. The studies in Chapters Four and Five use subscales of the Achievement Emotions Questionnaire (AEQ; Pekrun et al., 2011). We improve upon previous methodology using a relatively novel (see Villafañe et al., 2014; 2016, for another use of the study design) longitudinal panel study design, in which the affective and achievement measures were administered more than once throughout the course of the semester. The majority of previous research in postsecondary chemistry education has utilized cross-sectional designs, in which measures of non-cognitive factors have been administered at one point in the semester and achievement is later measured (e.g., Chan & Bauer, 2014; Chan & Bauer, 2016a; Ferrell & Barbera, 2015); the cross-sectional design masks changes that occur throughout the semester, and studies conducted with the cross-sectional design do not possess the ability to lend support to causal claims (Nieswandt, 2007). The longitudinal panel design further enables us to explore reciprocal relationships as they are related to changes over time, while cross-sectional studies limit the relationships to be directionally ambiguous. Chapters Four and Five include psychometric evaluations of the subscales of the AEQ followed by tests of reciprocal causation based on the study design through structural equation modeling.

Psychometric evaluation of instruments utilized in educational research requires evidence for their accuracy and precision, known in this context as validity and reliability. The conceptual basis for incorporating evidence for validity and reliability comes from the *Standards for Educational and Psychological Testing* (American Educational Research Association et al., 2014) and have been recommended for use in CER by Arjoon, Xu, and Lewis (2013). Validity is a structure for providing evidence of the appropriateness of the use of data collected via instruments. This includes determinations of the extent of internal structure of the data, typically measured by the use of factor analytic procedures (see Brown, 2015, for overview). The extent to which the measurements are related to other variables that are either synonymous or antonymous is known as relations to other variables (or external) validity. Another dimension of validity is content validity, or the extent to which the items reflect content within the universe of the construct intended to be measured. Response process validity is typically measured using cognitive interviews during which representatives of the target population explain their process of selecting a response in a think-aloud procedure. Finally, instruments are evaluated based on the validity for the consequences of scores, especially when they are used as admissions or progression criteria. Reliability, a measure of precision, is commonly measured by Cronbach's $\alpha$ (explained in detail in Cortina, 1999), which quantifies of the amount of variance in the data accounted for by the common factor between the items. The studies included here primarily provide evidence for internal structure and relations to other variables validity along with $\alpha$ of the subscales of the AEQ used.

Internal structure validity is measured using confirmatory factor analysis (CFA). All factor analytic and structural modeling was conducted in MPlus versions 7.1 and 8 (Muthén & Muthén, 1998 - 2017). CFA is a statistical analysis procedure in which parameters representing the relationships between observed variables and the construct(s) of interest are estimated according to a structure provided by the theoretical framework in addition to other studies utilizing the instrument. Models are either supported or deemed inappropriate based on a series of criteria for fit, including the chi-squared ($\chi^2$) for which statistical significance indicates poor fit, the comparative fit index (CFI) and Tucker-Lewis index (TLI), for which a value $\geq 0.95$ is considered good fit, the standardized root mean square residual (SRMR), for which a value $\leq 0.08$ is considered good fit, and the root mean square error of approximation (RMSEA), for which a value

≤ 0.05 is considered good fit, ≤ 0.08 is considered appropriate fit, and ≤ 0.10 is considered marginal fit (Hu & Bentler, 1999).

Another piece of internal structure validity essential for the studies reported here is the determination of longitudinal measurement invariance. Measurement invariance is typically considered to be an aspect of fairness, when instruments are analyzed for their consistency in measurement between individuals of different inherent groupings (i.e., race or sex characteristics). However, the measures utilized in these studies were administered to the same groups over time, therefore, longitudinal measurement invariance demonstrates consistency of measurement properties over time. This procedure accounts for whether the changes in scores on the instruments is due to alpha, beta, or gamma change. Alpha change refers to true change over time. Beta change is change in the ways in which students respond to the items after having seen them before. Gamma change occurs when the meaning of the construct changes over time (Brown, 2015). Evidence supporting invariance indicates that alpha change is observed, and that internal structure validity is upheld.

Evidence for the relations to other variables validity of the subscales of the AEQ used here is provided through structural equation modeling (SEM; see Kline, 2016, for detailed information). SEM is a maximum likelihood procedure which subsumes CFA. In SEM, parameters are estimated to explain the relationships between variables as dictated by an a priori designed model. In Chapters Four and Five, we determine which of a series of theoretically supported models fit the data the best, according to the same criteria and fit indices described above for CFA models. The SEM procedure can be compared to a linear regression procedure, in which standardized path coefficients are provided by the software program to provide us with information about the valence and direction of effects. The results of SEM, then, serve as a gauge of the relationship of the constructs of interest (in this case, the achievement emotions) to the outcomes of interest (in this case, scores on postsecondary organic chemistry examinations).

## Summary

The works described above are included as separate chapters in this dissertation. The first two studies seek to provide a snapshot of chemical education on the national scale in the United States. Chapter Two (Gibbons, Villafañe, et al., 2018) addresses the link between faculty beliefs about teaching and learning and self-efficacy in both pedagogy and chemistry content and their enacted instructional practice. Chapter Three (Gibbons et al., 2017) evaluates the niche market for classroom response systems across United States postsecondary chemistry courses. Chapters Four and Five seek to explore the relationship of affect with achievement within a single institution. Chapter Four (Gibbons, Xu, et al., 2018) explores the nuance, direction, and valence of relationships of anxiety and enjoyment with achievement in an organic chemistry context. Chapter Five (unpublished work) looks across the entire semester of organic chemistry and explores a reciprocal causation model with learning-related boredom. Each chapter demonstrates methodological improvement over past published work in CER. The type of research demonstrated in this dissertation (i.e., studies which collect data to provide empirical evidence for the state of the system of interest) are essential for making evidence-based decisions in future research, future instructional practice, and future policy (see Chapter Six).

## References

Abraham, M. R. (2008). Importance of a theoretical framework for research. In D. M. Bunce & R. Cole (Eds.), *Nuts and Bolts of Chemical Education Research* (pp. 47-66). Washington, DC: American Chemical Society.

American Educational Research Association, American Psychological Association, National Council on Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing. (2014). *Standards for Educational and Psychological Testing* (Vol. AERA). Washington, DC.

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*, 536-545. doi:10.1021/ed3002013

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. doi:10.1021/ed082p1864

Bauer, C. F. (2008). Attitude toward chemistry: A semantic differential instrument for assessing curriculum impacts. *Journal of Chemical Education, 85*, 1440-1445. doi:10.1021/ed085p1440

Blair, J., Czaja, R. F., & Blair, E. A. (2014a). Sampling I: Concepts of sample representation and sample quality In *Designing Surveys: A Guide to Decisions and Procedure* (3 ed.). Thousand Oaks, CA: Sage.

Blair, J., Czaja, R. F., & Blair, E. A. (2014b). Sampling III: Sample size and sample design. In *Designing surveys: A guide to decision and practice* (3 ed.). Thousand Oaks, CA: Sage.

Bodner, G. M. (1991). I have found you an argument: The conceptual knowledge of beginning chemistry graduate students. *Journal of Chemical Education, 68*(5), 385. doi:10.1021/ed068p385

Brandriet, A. R., & Bretz, S. L. (2014). 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. doi:10.1021/ed500051n

Bretz, S. L. (2001). Novak's Theory of Education: Human Constructivism and Meaningful Learning. *Journal of Chemical Education, 78*(8), 1107. doi:10.1021/ed078p1107.6

Bretz, S. L., & Linenberger, K. J. (2012). Development of the enzyme–substrate interactions concept inventory. *Biochemistry and Molecular Biology Education, 40*(4), 229-233. doi:10.1002/bmb.20622

Bretz, S. L., & Murata Mayo, A. V. (2018). Development of the Flame Test Concept Inventory: Measuring Student Thinking about Atomic Emission. *Journal of Chemical Education, 95*(1), 17-27. doi:10.1021/acs.jchemed.7b00594

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2 ed.). New York, NY: Guilford Press.

Chan, J. Y. K., & Bauer, C. F. (2014). Identifying At-Risk Students in General Chemistry via Cluster Analysis of Affective Characteristics. *Journal of Chemical Education, 91*(9), 1417-1425. doi:10.1021/ed500170x

Chan, J. Y. K., & Bauer, C. F. (2015). Effect of peer-led team learning (PLTL) on student achievement, attitude, and self-concept in college general chemistry in randomized and quasi experimental designs. *Journal of Research in Science Teaching, 52*(3), 319-346. doi:10.1002/tea.21197

Chan, J. Y. K., & Bauer, C. F. (2016a). Learning and studying strategies used by general chemistry students with different affective characteristics. *Chemistry Education Research and Practice, 17*, 675-684. doi:10.1039/c5rp00205b

Chan, J. Y. K., & Bauer, C. F. (2016b). Learning and studying strategies used by general chemistry students with different affective characteristics. *Chemistry Education Research and Practice, 17*(4), 675-684. doi:10.1039/C5RP00205B

Chase, A., Pakhira, D., & Stains, M. (2013). Implementing Process-Oriented, Guided Inquiry Learning for the first time: Adaptations and short-term impacts on students' attitude and performance. *Journal of Chemical Education, 90*, 409-416. doi:10.1021/ed300181t

Cortina, J. M. (1999). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*, 98-104. doi:10.1037/0021-9010.78.1.98

Cox, B. E., McIntosh, K. L., Reason, R. D., & Terenzini, P. T. (2011). A Culture of Teaching: Policy, Perception, and Practice in Higher Education. *Research in Higher Education, 52*(8), 808-829. doi:10.1007/s11162-011-9223-6

Crocker, L., & Algina, J. (1968). *Introduction to Classical and Modern Test Theory*. Orlando, FL: Holt, Rinehart and Winston.

D'Eon, M., Sandownik, L., Harrison, A., & Nation, J. (2008). Using self-assessments to detect workshop success: Do they work? *American Journal of Evaluation, 29*, 92-98.

Dalgety, J., Coll, R. K., & Jones, A. (2003). Development of chemistry attitudes and experiences questionnaire (CAEQ). *Journal of Research in Science Teaching, 40*(7), 649-668. doi:10.1002/tea.10103

DeHaann, R. L. (2005). The impending revolution in undergraduate science education. *Journal of Science Edcuation & Technology, 14*(2), 253-269.

DeVelis, R. F. (2017). *Scale development theory and applications* (4 ed.). Thousand Oaks: Sage.
Ebert-May, D., Derting, T. L., Hodder, J., Momsen, J. L., Long, T. L., & Jardeleza, S. E. (2011). What we say is not what we do: Effective evaluation of faculty professional development programs. *BioScience, 61*, 550.
Emenike, M. E., Schroeder, J., Murphy, K., & Holme, T. (2011). A snapshot of chemistry faculty members' awareness of departmental assessment efforts *Assessment Update, 23*(4), 1-2, 14-16. doi:10.1002/au.234
Emenike, M. E., Schroeder, J., Murphy, K., & Holme, T. (2013). Results from a National Needs Assessment Survey: A View of Assessment Efforts within Chemistry Departments. *Journal of Chemical Education, 90*(5), 561-567. doi:10.1021/ed200632c
Ferrell, B., & Barbera, J. (2015). Analysis of students' self-efficacy, interest, and effort beliefs in general chemistry. *Chemistry Education Research and Practice, 16*(2), 318-337. doi:10.1039/C4RP00152D
Frenzel, A. C., Goetz, T., Ludke, O., Pekrun, R., & Sutton, R. E. (2009). Emotional transmission in the classroom: Exploring the relationship between teacher and student enjoyment. *Journal of Educational Psychology, 101*, 705-716. doi:10.1037/a0014695
Gibbons, R. E., Laga, E. E., Leon, J., Villafañe, S. M., Stains, M., Murphy, K. L., & Raker, J. R. (2017). Chasm Crossed? Clicker Use in Postsecondary Chemistry Education. *Journal of Chemical Education, 94*(5), 549-557. doi:10.1021/acs.jchemed.6b00799
Gibbons, R. E., Villafañe, S. M., Stains, M., Murphy, K. L., & Raker, J. R. (2018). Beliefs about learning and enacted instructional practices: An investigation in postsecondary chemistry education. *Journal of Research in Science Teaching*, n/a-n/a. doi:10.1002/tea.21444
Gibbons, R. E., Xu, X., Villafañe, S. M., & Raker, J. R. (2018). Testing a reciprocal causation model between anxiety, enjoyment, and academic performance in postsecondary organic chemistry. *Educational Psychology*. doi:10.1080/01443410.2018.1447649
Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly, 70*(5), 646-675.
Herrington, D. G., Yezierski, E. J., & Bancroft, S. F. (2016). Tool trouble: Challenges with using self-report data to evaluate long-term chemistry teacher professional development. *Journal of Research in Science Teaching, 53*(7), 1055-1081.
Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modelling: Multidisciplinary Journal, 6*, 1-55. doi:10.1080/10705519909540118
Johnstone, A. H. (2000). TEACHING OF CHEMISTRY - LOGICAL OR PSYCHOLOGICAL? *Chemistry Education Research and Practice, 1*(1), 9-15. doi:10.1039/A9RP90001B
Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling* (4th ed.). New York, NY: Guilford Press.
Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality & Quantity, 47*(4), 2025-2047.
Larsen, R. D. (1986). Attitude reconstruction: First-day classroom advocacy. *Journal of Chemical Education, 63*(10), 864. doi:10.1021/ed063p864
Lewis, S. E., Shaw, J. L., Heitz, J. O., & Webster, G. H. (2009). Attitude Counts: Self-Concept and Success in General Chemistry. *Journal of Chemical Education, 86*(6), 744. doi:10.1021/ed086p744
Likert, R. (1932). A technique for the measurement of attitudes *Archives of Psychology, 22*, 5-55.
Linnenbrink, E. A. (2006). Emotion research in education: Theoretical and methodological perspectives on the integration of affect, motivation, and cognition. *Educational Psychology Review, 18*, 307-314. doi:10.1007/s10648-006-9028-x

Liu, Y., Ferrell, B., Barbera, J., & Lewis, J. E. (2017). Development and evaluation of a chemistry-specific version of the academic motivation scale (AMS-Chemistry). *Chemistry Education Research and Practice, 18*(1), 191-213. doi:10.1039/C6RP00200E

Liu, Y., Raker, J. R., & Lewis, J. E. (2018). Evaluating student motivation in organic chemistry courses: moving from a lecture-based to a flipped approach with peer-led team learning. *Chemistry Education Research and Practice, 19*(1), 251-264. doi:10.1039/C7RP00153C

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. doi:10.1021/ed400700q

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. doi:10.1080/09500693.2012.684433

Moog, R. S., & Spencer, J. N. (2008). POGIL: An overview. In R. S. Moog & J. N. Spencer (Eds.), *Process Oriented Guided Inquiry Learning (POGIL)* (Vol. 994, pp. 1-13). Washington, DC: American Chemical Society.

Muthén, L. K., & Muthén, B. O. (1998 - 2017). *Mplus User's Guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.

National Research Council. (2012). *Discipline-based education research: Understanding and improving learning in undergraduate science and engineering*. Washington, DC: The National Academies Press.

Nederhof, A. J. (1985). Methods of coping with social desirability bias: a review. *European Journal of Social Psychology, 15*(3), 263-280.

Nieswandt, M. (2007). Student affect and conceptual understanding in learning chemistry. *Journal of Research in Science Teaching, 44*, 908-937. doi:10.1002/tea.20169

Novak, J. D. (2002). Meaningful learning: The essential factor for conceptual change in limited or inappropriate propositional hierarchies leading to empowerment of learners. *Science Education, 86*(4), 548-571.

Pekrun, R. (2000). A social-cognitive, control-value theory of achievement emotions. In J. Heckhausen (Ed.), *Motivational psychology of human development* (pp. 143-163). Oxford: Elsevier.

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Academic Emotions Questionnaire (AEQ). *Contemporary Educational Psychology, 36*, 36-48. doi:10.1016/j.cedpsych.2010.10.002

Pekrun, R., Hall, N. C., Goetz, T., & Perry, R. P. (2014). Boredom and academic achievement: Testing a model of reciprocal causation. *Journal of Educational Psychology, 106*, 696-710. doi:10.1037/a0036006

Qureshi, S., Vishnumolakala, V. R., Southam, D. C., & Treagust, D. F. (2017). Inquiry-Based Chemistry Education in a High-Context Culture: a Qatari Case Study. *International Journal of Science and Mathematics Education, 15*(6), 1017-1038. doi:10.1007/s10763-016-9735-9

Raker, J. R., Emenike, M. E., & Holme, T. A. (2013). Using Structural Equation Modeling To Understand Chemistry Faculty Familiarity of Assessment Terminology: Results from a National Survey. *Journal of Chemical Education, 90*(8), 981-987. doi:10.1021/ed300636m

Raker, J. R., & Holme, T. A. (2014). Investigating Faculty Familiarity with Assessment Terminology by Applying Cluster Analysis To Interpret Survey Data. *Journal of Chemical Education, 91*(8), 1145-1151. doi:10.1021/ed500075e

Rogers, E. M. (1995). *Diffusion of Innovation* (4th ed.). New York: Free Press.

Seery, M. K. (2015). Flipped learning in higher education chemistry: emerging trends and potential directions. *Chemistry Education Research and Practice, 16*(4), 758-768. doi:10.1039/C5RP00136F

Shadle, S. E., Marker, A., & Earl, B. (2017). Faculty drivers and barriers: laying the groundwork for undergraduate STEM education reform in academic departments. *International Journal of STEM Education, 4*(1), 8. doi:10.1186/s40594-017-0062-7

Stains, M., Harshman, J., Barker, M. K., Chasteen, S. V., Cole, R., DeChenne-Peters, S. E., . . . Young, A. M. (2018). Anatomy of STEM teaching in North American universities. *Science, 359*(6383), 1468-1470. doi:10.1126/science.aap8892

StataCorp. (2015). *Stata Statistical Software: Release 14*. College Station, TX: StataCorp LLC.

Villafañe, S. M., Garcia, C. A., & Lewis, J. E. (2014). Exploring diverse students' trends in chemistry self-efficacy throughout a semester of college-level preparatory chemistry. *Chemistry Education Research and Practice, 15*(2), 114 -127. doi: 10.1039/c3rp00141e.

Villafañe, S. M., Xu, X., & Raker, J. R. (2016). Self-efficacy and academic performance in first-semester organic chemistry: Testing a model of reciprocal causation. *Chemistry Education Research and Practice, 17*(973-984). doi:10.1039/C6RP00119J

Vishnumolakala, V. R., Southam, D. C., Treagust, D. F., Mocerino, M., & Qureshi, S. (2017). Students' attitudes, self-efficacy and experiences in a modified process-oriented guided inquiry learning undergraduate chemistry classroom. *Chemistry Education Research and Practice, 18*(2), 340-352. doi:10.1039/C6RP00233A

Weisberg, H. F., Krosnick, J. A., & Bowen, B. D. (1996). *An introduction to survey research, polling, and data analysis* (3 ed.). Thousand Oaks, CA: Sage.

Wilson, S. B., & Varma-Nelson, P. (2016). Small groups, significant impact: A review of peer-led team learning research with implications for STEM education. *Journal of Chemical Education, 93*, 1686-1702. doi:10.1021/acs.jchemed.5b00862

Woodbury, S., & Gess-Newsome, J. (2002). Overcoming the paradox of change without difference: A model of change in the arena of fundamental school reform. *Educational Policy, 16*, 763-782.

Xu, X., & Lewis, J. (2011). Refinement of a chemistry attitude measure for college students. *Journal of Chemical Education, 88*, 561-568. doi:10.1021/ed9000071q

Xu, X., Villafañe, S. M., & Lewis, J. E. (2013). College students' attitudes toward chemistry, conceptual knoweldge and achievement: structural equation model analysis. *Chemistry Education Research and Practice, 14*(2), 188-200. doi: 10.1039/c3rp20170h

# CHAPTER TWO:

# BELIEFS ABOUT TEACHING AND LEARNING AND ENACTED INSTRUCTIONAL PRACTICE: AN INVESTIGATION IN POSTSECONDARY CHEMISTRY

## Note to Reader

This chapter is a manuscript published in the *Journal of Research in Science Teaching* (2018), https://doi.org/10.1002/tea.21444. It is reprinted here with permission of John Wiley and Sons Publishing. Permissions information can be found in Appendix B. This work was published with co-authors; the writing and data analysis and interpretation are my own, but all co-authors provided feedback on early drafts of the work. Sachel M. Villafañe was the original designer and constructor of the self-efficacy and beliefs about teaching and learning instrument (SBTL-I) used in the manuscript. Marilyne Stains was instrumental in the selection of the theoretical framework for this study. Kristen L. Murphy is the Director of the ACS Examinations Institute, and therefore provided administrative support to the study. Jeffrey R. Raker is the technical and administrative supervisor of the survey used in this report.

## Introduction

Pedagogical reform in chemistry, as in STEM education generally, is a complex endeavor with many factors that enable or inhibit efforts (Dole & Sinatra, 1998; Gess-Newsome, Southerland, Johnston, & Woodbury, 2003; Greensfeld & Elkad-Lehman, 2007; Henderson, 2008; Henderson et al., 2015; Ho, Watkins, & Kelly, 2001; Woodbury & Gess-Newsome, 2002). A reformed postsecondary chemistry classroom is defined for this study as a student-centered environment in which constructivism guides the use of pedagogical techniques such as active learning (Sawada et al., 2002). In postsecondary chemistry education, active learning is encompassed by the development and implementation of evidence-based instructional practices (EBIPs; National Research Council, 2012). EBIPs such as peer-led team learning, are based on theory and empirical research, and many have demonstrated success in increasing student content and affective learning. Despite evidence supporting the efficacy of these pedagogies, widespread EBIP adoption is lacking (Henderson & Dancy, 2009; Lund & Stains, 2015; Walczyk & Ramsey, 2003). Adoption of a reformed pedagogy is not a trivial choice, but is based on a delicate balance of complex factors including beliefs about how teaching and learning should occur (Cohen & Mehta, 2017; Henderson, Beach, & Finkelstein, 2011; Henderson et al., 2015; Lund & Stains, 2015; Walczyk, Ramsey, & Zha, 2007). This study seeks to explore the system of reform adoption in postsecondary chemistry through a survey of faculty members.

Many factors influence reform choices; it is unreasonable to exhaustively measure all factors in a single study. However, there is value in isolating and exploring relationships between such factors influencing instructional choices. Such a study allows for the exploration of generalized routes for encouraging EBIP adoption. Outside of the classroom, beliefs have been shown to promote as well as inhibit the adoption of new ideas and technologies (Moore, 2002; Rogers, 1995); understanding distribution through the lens of the beliefs and attitudes of adopters has shown to impact adoption of classroom response systems (CRS) in postsecondary chemistry education (Emenike & Holme, 2012; Gibbons et al., 2017; MacArthur, 2013; Towns, 2010). Despite acceptance that beliefs are associated with enacted pedagogies and activities, there is difficulty in operationalizing and measuring such beliefs (Pajares, 1992).

It is theorized that fundamental beliefs held by an instructor about teaching and learning, as well as their self-efficacy in enacting instructional activities, have a reciprocal influence on the adoption of a new pedagogy. Woodbury and Gess-Newsome (2002) provided theoretical support for teacher thinking as a key element to fundamental change: "teachers' knowledge and beliefs about their subject matter, or teaching and learning in their subject area, that are incompatible with reform intentions often significantly diminish the outcomes of what were meant to be fundamental reforms" (p. 771). Devlin's (2006) review of the importance of considering conceptions of teaching revealed a more nuanced view of faculty members' beliefs. Devlin argued that the current state of research on the relationship between teacher thinking and enacted instructional practice is inadequate for making recommendations for pedagogical developers and education researchers. Work by Veal, Riley Lloyd, Howell, and Peters (2016) also supports a reciprocal relationship between instructional beliefs and activities. Due to the variety of claims in the literature, an associative relationship will be explored in this study.

A lack of empirical evidence for the link between enacted instructional activities and beliefs about self-efficacy in instruction or beliefs about teaching and learning has limited the ability to establish the tenability of theories of educational change and reform adoption. We present herein the use of self-report measures with a national sample of postsecondary chemistry faculty members to better understand the quantitative link between beliefs about learning, efficacy, and enacted instructional practices. Chemistry was selected as the field in which to conduct this study as chemistry faculty represent those faculty who are traditionally well-versed in their own technical literature but unfamiliar with educational reform (e.g., Raker & Holme, 2014), despite being the instructors for the prerequisite courses for most science major fields in the undergraduate curriculum such as general and organic chemistry. These faculty members are challenged to perform research tasks while managing a load of instructional tasks, resulting in a unique context for exploring the relationship between beliefs and instruction.

**Beliefs About Teaching and Learning**

Faculty members' perspectives on teaching and learning and subsequent application of student-centered instruction are influenced by a complex set of factors (Herrington, Yezierski, & Bancroft, 2016). From a cognitive psychological approach, dissatisfaction along with social context, motivation, and self-efficacy are crucial for change (Dole & Sinatra, 1998). Empirical work suggests that change either begins with dissatisfaction with current instruction or with the belief that students learn better with different techniques then those currently used (Bauer, Libby, Scharberg, & Reider, 2013; Windschitl & Sahl, 2002). Beliefs encompass not only thoughts about how teaching and learning occur but also the level of confidence an instructor holds regarding their ability to utilize reformed pedagogies (i.e., self-efficacy; Pajares, 1992). This study seeks to explore both of these dimensions of teacher thinking.

**Previous qualitative results**

Research conducted on how faculty members' thinking corresponds to change has primarily utilized case study and qualitative methods. This includes, for example, applications of the Teacher Beliefs Interview (Luft & Roehrig, 2007; Roehrig, Kruse, & Kern, 2007).

One dimension of thinking involves self-efficacy, that is, the perceived ability that one can complete a given task; for our study, tasks include the ability to enact instructional practices such as whole group discussion or operate classroom response systems (CRS). Feldman (2000) tracked changes in instruction, finding a connection between perceived efficacy of the method and adoption of new classroom strategies. Similarly, in a chemistry context, Orgill, Bussey, and Bodner (2015) found that faculty members who perceived particular instructional strategies (e.g., use of analogies to convey concepts and theories) to be more useful reported more frequent use.

The second dimension explored in our study is beliefs about teaching and learning. Beliefs are essential to understanding instructional choices (Harwood, Hansen, & Lotter, 2006; Lotter, Harwood, & Bonner, 2007). Lotter et al. (2007) found that what teachers believe about their students, about science, as well as their beliefs about effective teaching were influential on the impact of professional development experiences on instructional choices. Faculty members in a partnership program with practicing scientists experienced changes in their conceptions of science and self-reported implementations of reformed pedagogies (Houseal, Abd-El-Khalick, & Destefano, 2014). Community college mathematics instructors were observed to have aligned their beliefs with instruction, resulting in variation in the extent to which they used or modified reformed pedagogies (Mesa, Celis, & Lande, 2014).

Efficacy and beliefs together have shown additive impact on instruction. Sunal et al. (2001) found that faculty members who regarded their role as a facilitator of learning held high self-efficacy in teaching and were more likely to implement reformed curricula. Other research supports the finding that science teachers require both high self-efficacy and beliefs in the superiority of a reform in order to implement successful change (Haney, Lumpe, Czerniak, & Egan, 2002). Ho et al. (2001) found a dichotomy between teaching focused on transmission of knowledge and teaching focused on helping students develop their own understanding, the latter aligning with use of reformed practices.

Even when instructional reforms were required, the strength of traditionally oriented beliefs greatly inhibited teachers from implementing reforms (Gess-Newsome et al., 2003; Roehrig & Kruse, 2005; Smith & Southerland, 2007). In a department where reform was implemented in all general chemistry courses, instructors whose belief systems aligned with the reform were successful, while those who encountered negative experiences when implementing the reform reverted back to traditional methods in subsequent iterations of the courses (Gallos, van den Berg, & Treagust, 2005). A study in postsecondary biology education concluded that faculty members preferred private-empirical (i.e., anecdotal) evidence over research findings in making instructional decisions (Andrews & Lemons, 2015).

As indicated by Devlin (2006), the impact of teacher beliefs and thinking on instruction is not a direct relationship. Mutambuki and Fynewever (2012) found that chemistry faculty members, despite describing a belief that students need to extrapolate their reasoning to demonstrate learning, imposed an expert-like reasoning strategy rather than observing genuine student reasoning. Similarly, Mansour (2013) found that secondary school teachers, despite holding a constructivist philosophy of learning, do not implement constructivism-oriented practices. The multifaceted nature of reforms, and the speed at which they are disseminated, has confused teachers, resulting in a halt to the growth of the reform (Smith & Southerland, 2007).

These results from the qualitative literature described above point to the importance of self-efficacy and beliefs about teaching and learning on the adoption of more EBIPs; however, these studies fail to provide a generalizable understanding of the impact of such beliefs and confidence on instruction across the larger postsecondary chemistry curriculum (Devlin, 2006).

### Previous large-scaled results

There are challenges to measuring faculty members' beliefs and efficacy. Pajares (1992) operationalized "belief" to include teacher efficacy, self-efficacy, epistemic beliefs, and the nature of science. The challenge of measuring these constructs is in confidence in the interpretation of resultant scores (i.e., validity; e.g. DeVelis, 2017). Tools designed to measure beliefs have struggled to meet this challenge. There has been widespread use of the Approaches to Teaching Inventory (Trigwell & Prosser, 2004); however, no reliable factor structure has been determined for use with this instrument (Harshman & Stains, 2017). Similarly, use of the Science Teaching Efficacy Belief Instrument (Riggs & Enochs, 1990), Teaching of Science as Inquiry (Smolleck & Yoder, 2008; Smolleck, Zembal-Saul, & Yoder, 2006), and Inquiry Teaching Beliefs instrument (Harwood et al., 2006) have not resulted in findings that capture the extent of the relationship between beliefs and instruction (Herrington et al., 2016).

There have been several limited in scope studies that support relationships between beliefs and instruction. A study of physics faculty members revealed that the use of evidence-based instructional strategies is associated with the belief that students learn best through problem solving (Borrego, Froyd, Henderson, Cutler, & Prince, 2013). In a different study, it was found that transmission of knowledge beliefs decreased the potential for student achievement (Gow & Kember, 1993). Discursive claims were more closely related to instruction than beliefs about how teaching should be done, indicating the complex relationship between beliefs and instructional choices (Veal et al., 2016). Through a teacher training program, it was found that exposure to reform does not result in changes in instruction; preservice teachers' beliefs changed both toward and away from reform-mindedness based on their professional development experiences (Struyven, Dochy, & Janssens, 2010).

Efficacy in both content and pedagogy has been linked to how instructors conduct their courses and make pedagogical decisions (Feldman, 2000; Roehrig & Kruse, 2005). In a study of engineering faculty members, it was found that efficacy in instruction contributed to the use of active learning techniques (Colbeck, Cabrera, & Marine, 2002). Similar results have been found in elementary education, in which increased self-efficacy was found to be associated with reformed teaching in mathematics (Lakshmanan, Heath, Perlmutter, & Elder, 2011).

These findings align with qualitative findings outlined in the previous section, but further situate a need for more large-scale studies of the relationship between beliefs, self-efficacy, and enacted instructional practices. Therefore, the study presented in this report seeks to provide the support needed for the measurement tools used as well as collect information on a large scale to bolster the theoretical argument for the link between practices and beliefs.

**Enacted Instructional Practices**

The definition of a reformed postsecondary chemistry classroom invoked in this report calls for an evaluation of the practices enacted in this classroom. Instructional practices are defined as activities of the instructor or students or interaction between the instructor and student(s) that occur in the context of classroom instruction; such practices include answering student questions, asking questions utilizing CRS, and conducting whole class discussions. While these individual activities contribute to the level of reform in a classroom, the overall combination of instructional practices is a better indicator for evaluating the level of reform; such combinations of instructional practices are noted as instructional styles in our study. Measuring such instructional styles for large populations has proved challenging in previous research. Observational protocols are a key method for the measurement of reformed instruction. Two popular protocols are the Reformed Teaching Observation Protocol (RTOP; Pilburn et al., 2000; Sawada et al., 2002) and the Classroom Observation Protocol in Undergraduate STEM (COPUS; Smith, Jones, Gilbert, & Wieman, 2013). These protocols provide descriptive information about the activities of the teacher and students, and their interactions in the classroom. Roehrig et al. (2007) found a greater than 0.50 correlation between reformed instruction utilizing the RTOP and the Teacher's Beliefs Interview. Lund et al. (2015) used the RTOP and COPUS to analyze an array of classrooms. These researchers used data from 10 observational items to identify and characterize instructional profiles: *Lecture*, *Socratic*, *Peer Instruction*, and *Collaborative Learning*, listed from least to most active pedagogy. These profiles demonstrate some of the different combinations of instructional practices that are enacted in STEM education settings.

While observational studies are helpful for describing a small set of classrooms, such protocols are unreasonable for large-scale investigations aimed at capturing the national state of postsecondary chemistry reform. Self-report data are economical for large-scale studies. Self-reported data do pose an issue: self-report data do not directly correlate with observational data (D'Eon, Sandownik, Harrison, & Nation, 2008; Ebert-May et al., 2011; Herrington et al., 2016; Kane, Sandretto, & Heath, 2002; Veal et al., 2016). The self-report method used in this study reflects that of other survey instruments designed to measure instructional practice, including the Teaching Practices Inventory (Wieman & Gilbert, 2014), Science Teaching Beliefs and Practices survey (Marbach-Ad, Ziemer, Orgler, & Thompson, 2014), and Postsecondary Instructional Practices Survey (Walter, Beach, Henderson, & Williams, 2014). These instruments are lengthy and lack focus on enacted practices (see review by Williams, Walter, Henderson, & Beach, 2015). An instrument designed to measure instructional practice was therefore adapted and applied in this study. We chose the COPUS as a framework for designing a tool to capture instructional practices so that direct comparisons could be made with observational research studies conducted in similar educational contexts.

One of the goals of this study is to provide evidence in the support of the use of self-report tools as compared to observational data. To address concerns related to self-report data, we compare the results of our self-report study to the study by Lund et al. (2015), which used more small-scale, resource-intensive observational data collection strategies. Because the goal of determining level of reform in postsecondary chemistry classrooms requires exploring the multifaceted tools of instruction incorporated, rather than the individual practices utilized, this study will use the statistical method of cluster analysis to determine instructional styles. To differentiate our study from the Lund et al. observational study, the results of their study will be referred to as instructional profiles, while ours will be referred to as instructional styles.

**Theoretical Framework**

Most models of reform focus exclusively on aspects of classroom behavior, and therefore fail to capture the complex context of instruction (Henderson et al., 2015; Lotter et al., 2007). To account for variations in dissemination of reforms, Woodbury and Gess-Newsome (2002) proposed the teacher-centered systemic reform (TCSR, Figure 2.1) model, which outlines the impact of cultural context, personal contextual factors, and teacher thinking on enacted practices. In addition, interactions between cultural contexts (e.g., characteristics of the school, students, and climate), personal contexts (e.g., engagement in professional development on pedagogical innovation), and teacher thinking (e.g., beliefs about teaching and learning) are modeled in the framework. The TCSR model is practical for understanding future adoption and development of curricular reforms and has been used to frame and explain growth of reform in K-20 STEM education (Enderle, Southerland, & Grooms, 2013; Gibbons et al., 2017; Graves, Hughes, & Balgopal, 2016; Lund & Stains, 2015). The model, however, has not been widely used in postsecondary science settings (e.g., Stains, Pilarz, & Chakraverty, 2015).

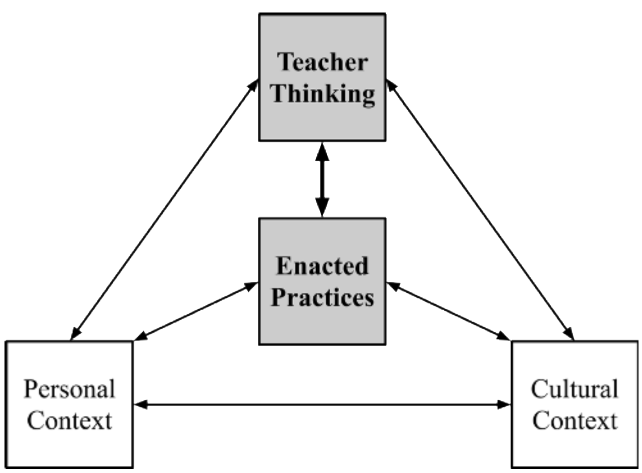


**Figure 2.1**. Teacher-centered systemic reform (TCSR) model of educational change. The bolded arrow highlights the link between teacher thinking and enacted instructional practices evaluated in this study.

One limitation that has prohibited use of the TCSR model in discipline-based education research is a lack of empirical evidence to support the relationships proposed in the model (Woodbury & Gess-

Newsome, 2002). Understanding the interaction of the theorized factors on instruction would provide validation for the relationships posited in the model (Gess-Newsome et al., 2003). To provide empirical evidence to support the theory, we evaluate the relationship between teacher thinking and enacted instructional practices. While survey studies have evaluated data framed by the TCSR model (e.g., Lund & Stains, 2015), our study is the first, large-scale validation of a relationship outlined in the TCSR model. We do not discount the importance of personal and cultural contexts as indicated by the TCSR model; however, we have focused our interest on a key relationship in the model from which further research can expand our analyses to garner a more comprehensive empirical evaluation of the model.

An additional limitation of the TCSR model is the broad characterizations of teacher thinking and enacted instructional practices. Based on the literature, beliefs about teaching and learning, and self-efficacy in enacting instructional practices were key to observed changes in or resistance to changes in practice. Based on the widespread use of the COPUS, we have chosen to operationalize enacted instructional practices as the result of a cluster analysis of self-reported instructional activities as defined in the COPUS; comparisons between our observed clusters based on self-reported data (referred to as instructional styles) and observed clusters as reported by Lund et al. (2005) based on observational data (referred to as instructional profiles) will provide validity evidence for the use of self-reported instructional practice data in our study.

### Study Goal and Research Questions

The goal of this study is to explore the relationship between beliefs about teaching and learning and enacted instructional practices in postsecondary chemistry. Our study is designed to answer two guiding research questions:

1. Do self-reported instructional practices partition into instructional styles? If so, do those instructional styles mirror instructional profiles based on observational data as reported by Lund et al. (2015)?

2. How are faculty members' beliefs about learning and efficacy in enacting pedagogies associated with self-reported instructional styles?

## Methods

### Survey

***Design.*** A national survey was administered to postsecondary chemistry faculty members via Qualtrics in February 2016. The survey asked respondents to describe a single undergraduate chemistry non-laboratory course taught over the past 3 years for which they had the most influence, including classroom practices and pedagogical techniques.

The questionnaire was framed using the TCSR model and constructed to measure personal context, teacher thinking, and cultural context in addition to enacted instructional practices. Measures included course level, number of students, institutional characteristics, number of years teaching, participating in teaching-focused workshops or positions, and beliefs about teaching and learning and self-efficacy.

***Population and sample.*** A database was built of all chemistry faculty members at institutions in the United States that conferred at least one bachelor's degree in chemistry in the years 2010-2015 as recorded by the National Center for Education Statistics Integrated Postsecondary Education Data System. Contact information was collected from institutional websites in Fall 2015 for 10,837 chemistry faculty members from 1,091 institutions. A stratified random sampling method was used to identify a sample of 6,442 faculty members. Six strata were defined by: institution control (public or private) and highest chemistry degree awarded (bachelor's, master's, or doctoral). Sample size was calculated to reach a goal 95% confidence level, 5% confidence interval, and assuming a 25% non-weighted response rate. For example, given the number of chemistry faculty members in Strata 1 and these goal parameters, 320 respondents are necessary; approximately four times that number were invited to participate in the study (i.e., 1,272). Sample selection and response rates are described in Table 2.1. In total, 1,282 chemistry faculty members responded (i.e., a 19.8% unit response rate).

**Table 2.1**. Strata and sample definition.

| Strata | Institutional Control | Highest Chemistry Degree Awarded | Number of Institutions | Number of Chemistry Faculty | Sample Size | Number of Responses |
|---|---|---|---|---|---|---|
| 1 | Public | Bachelors | 238 | 1,899 | 1,272 | 328 |
| 2 | Public | Masters | 83 | 1,080 | 1,075 | 226 |
| 3 | Public | Doctoral | 137 | 3,565 | 1,384 | 153 |
| 4 | Private | Bachelors | 551 | 2,836 | 1,335 | 403 |
| 5 | Private | Masters | 20 | 201 | 197 | 46 |
| 6 | Private | Doctoral | 62 | 1,256 | 1,179 | 126 |
| TOTALS | | | 1,091 | 10,837 | 6,442 | 1,282 |

### Measures

***Self-efficacy and beliefs about teaching and learning instrument.*** An instrument was designed as a part of this study to measure faculty members' beliefs and self-efficacy. The self-efficacy and beliefs about teaching and learning instrument (SBTL-I) was developed through a three-part instrument development process: First, 18 learning belief items and 18 self-efficacy items were constructed to parallel the practice-based items represented in the COPUS items used to design the instructional practices scale and based on published instruments (Harwood et al., 2006; Riggs & Enochs, 1990; Smolleck & Yoder, 2008; Smolleck et al., 2006; Trigwell & Prosser, 2004). Items were written to represent a teacher-centered and student-centered subscale for each set of items. The "strongly agree" at 5 to "strongly disagree" at 1 scale is typical of belief measures (e.g., Riggs & Enochs, 1990). The "completely" at 5 to "not at all" at 1 confidence scale is typical of self-efficacy measures (Bandura, 2006). The original 36 items were reviewed and revised by four education researchers familiar with measurement and psychometrics and six chemistry education practitioners.

The 36 items were pilot tested with 686 postsecondary chemistry faculty members (a population separate from the results presented herein). Initial attempts to obtain model fit based on the intended four subscales (i.e., self-efficacy and beliefs by student-centered and teacher-centered) were unsuccessful. A maximum likelihood, unweighted least squares exploratory factor analysis was conducted for each of the item sets with an oblique rotation in MPlus (Muthén & Muthén, 1998-2017). Items were retained for the measures based on standard cut-off levels (Kim & Mueller, 1978a, 1978b); both item sets yielded a two-factor solution: teacher-centered learning beliefs, student-centered learning beliefs, self-efficacy related to enacting pedagogies, and self-efficacy related to content. Intended constructs emerged with the learning beliefs measure; however, an unplanned, yet coherent set of constructs emerged with the self-efficacy measure.

The resulting 11-item learning belief measure and 12-item self-efficacy scale showed acceptable goodness-of-fit statistics for internal structure with a new sample of 1,026 postsecondary chemistry faculty members per Hu and Bentler (1999). Note: the sample utilized here is separate from the 686 faculty members in the EFA study as well as the 1,282 faculty members in the study reported herein. Fit statistics are in Appendix C (Table A1).

For the respondents of the survey analyzed in this study ($N$ = 1,282), items on both scales along with frequencies of responses by item are reported in Appendix B (Tables A2 & A3). Descriptive statistics for the four subscales are in Table 2.2. These data demonstrate acceptable goodness-of-fit statistics for internal structure per Hu and Bentler (1999), and are reported in Appendix C (Table A4).

***Enacted instructional practices.*** The instrument used in this study consists of 14 parts designed to mimic the COPUS. The self-report mechanism was developed such that respondents indicated the frequency (i.e., every class meeting, weekly, several times per semester, once, never) with which they used the 14 instructional practices. Descriptive statistics for each instructional practice are reported in Appendix C (Table A5).

**Table 2.2.** Descriptive statistics of the SBTL-I subscales.

| Factor | Mean | *SD* | Skewness | Kurtosis |
|---|---|---|---|---|
| Teacher-Centered | 3.81 | 0.50 | -0.410 | 3.675 |
| Student-Centered | 3.89 | 0.43 | -0.359 | 3.791 |
| Confidence Content | 4.14 | 0.61 | -0.423 | 2.769 |
| Confidence Pedagogy | 4.02 | 0.62 | -0.352 | 2.551 |

### Statistical procedures

***Research question 1.*** The 14 instructional practices are tedious to analyze individually; therefore, a data reduction methodology was employed to identify sets of practices that are used in conjunction during instruction. Our choice is congruent with the nature of reformed classrooms; such classrooms are characterized by a set of instructional practices rather than a singular instructional practice. Cluster analysis, therefore, is utilized to create descriptive clusters. All analyses are conducted in Stata14 (StataCorp, 2015). The cluster analysis is conducted using Ward's linkage and a matching similarity matrix (Ward, 1963). The Ward's linkage cluster analysis considers the variance in the data and generates clusters such that each observation added to a group maximizes the amount of variance accounted for by the clustering. Duda and Hart stopping rules (Duda, Hart, & Stork, 2001) are used to evaluate the cluster solution. Fisher's (1992) exact tests are used to determine cluster identities. Resultant clusters are characterized as instructional styles that can be directly compared with Lund et al.'s (2015) instructional profiles.

***Research question 2.*** A multivariate analysis of variance (MANOVA) is conducted with associated follow-up univariate analysis of variance (ANOVA) and Tukey's honestly significant difference (HSD) (Glass & Hopkins, 1984; Tukey, 1949) tests, as appropriate. These tests are used to determine differences between the SBTL-I subscales and instructional styles. We use Pillai's trace statistic for reporting MANOVA results due to its increased power when groups differ on more than two functions– in this case, the four subscales of the SBTL-I: *Teacher-Centered Learning, Student-Centered Learning, Self-Efficacy in Pedagogy,* and *Self-Efficacy in Content* (Stevens, 2009). Effect size is measured for the MANOVA test using corrected multivariate $\omega^2$ and for the univariate follow-up tests using $\eta^2$, for both of which $\leq 0.01$ is considered small, 0.06 medium, and 0.14 large (Cohen, 1973; Vacha-Hasse & Thompson, 2004; Stevens, 2009)

## Results

### Research question 1

A cluster analysis of the reported instructional practices yields a five-cluster solution [Je(2)/Je(1) = 0.8796, pseudo $T^2 = 54.73$]. Fisher's exact tests are used to determine cluster identity; 12 of the practices have significant ($p < .001$) results suggesting differences between the five clusters on that instructional practice. The results of this analysis are reported in Table 2.3 as "use," where use is defined as the combined percent of "every class meeting" and "weekly" for parsimony. The five instructional styles are classified as "Small Groups," "Interactive," "Lecture with classroom response systems/clickers (CRS)," "Lecture with Literature," and "Lecture" based on percent use and non-use of the instructional practices.

The instructional styles from our analysis align as expected with the instructional profiles found in the Lund et al. (2015) observational study. The *Lecture* and *Lecture with Literature* styles from our study are analogous to the "Lecture" profiles found by Lund et al. The *Lecture with CRS* style from our data is most similar to the "Socratic" profile from Lund et al. In both *Lecture with CRS* and Socratic, the primary activity is lecturing, but the addition of question asking to the lecture period differentiates these groupings from lecture-based methods. The *Interactive* style defined in our data incorporates a variety of techniques which do not match to the specific profiles in Lund et al.'s study; this is possibly due to the ability for observational data to detect differences in the ways and frequencies that these techniques are utilized that are not easily captured in self-report data. Finally, our *Small Groups* style is analogous to the "Collaborative Learning" profile found by Lund et al. The results of our analysis indicate that the self-reported data from our survey are adept at discerning differences between instructional styles. Our instructional profile findings, therefore, provide a means to investigate the relationships between enacted practices and the other factors of the TCSR model.

**Table 2.3.** Percent "every class meeting" and "weekly" for enacted instructional practices by instructional styles.

| Enacted Instructional Practices | *Small Groups* | *Interactive* | *Lecture with CRS* | *Lecture with Literature* | *Lecture* | Fisher's Exact Test |
|---|---|---|---|---|---|---|
| N = | 402 | 147 | 157 | 122 | 454 | |
| Lecturing | 86.8 | 95.2 | 98.7 | 99.2 | 98.7 | *** |
| Writing on the board | 93.5 | 95.9 | 96.2 | 98.4 | 95.6 | |
| Posing questions | 98.0 | 99.3 | 94.6 | 98.4 | 93.2 | *** |
| Answering questions | 97.8 | 100.0 | 96.2 | 98.4 | 97.4 | |
| Asking clicker questions | 27.4 | 8.8 | 55.4 | 14.8 | 0.4 | *** |
| Follow-up and provide feedback after a clicker question or other activity | 62.4 | 34.0 | 100.0 | 50.8 | 9.3 | *** |
| Assigning students to work in groups | 93.5 | 82.3 | 3.2 | 17.2 | 3.3 | *** |
| Moving through the class, guiding ongoing student work | 84.6 | 95.2 | 21.1 | 18.9 | 18.7 | *** |
| Extended discussion with small groups or individuals | 74.6 | 76.9 | 8.3 | 6.6 | 8.4 | *** |
| Showing or conducting a demonstration, experiment, simulation, video, or animation | 19.9 | 57.8 | 22.9 | 30.3 | 18.3 | *** |
| Asking students to make a prediction | 16.2 | 91.2 | 22.9 | 33.6 | 18.1 | *** |
| Referencing and discussing the primary literature | 8.2 | 44.9 | 0.6 | 92.6 | 6.0 | *** |
| Discussing the process by which a model, theory, or concept was developed | 29.4 | 85.0 | 31.9 | 73.0 | 30.8 | *** |
| Initiating a whole class discussion | 30.4 | 54.4 | 10.8 | 19.7 | 11.9 | *** |

*Note*. *** p <0.001

### Research question 2

A MANOVA is used to compare the mean scores of the five instructional profiles by the four SBTL-I subscales: *Teacher-Centered Learning, Student-Centered Learning, Self-Efficacy in Pedagogy,* and *Self-Efficacy in Content*. The multivariate result is significant (Pillai's trace = 0.1371, $F$ (4, 1277) = 11.33 $p$ < .0001). As a measure of effect size, Tatsuoka's corrected multivariate $\omega^2$ is used and found to be moderate (0.12). Cohen's effect size cut-off values are used as there are no comparable large-scale empirical measures of teacher thinking with which to compare effect sizes in a meta-analytic procedure for defining effect size in context (Vacha-Hasse & Thompson, 2004).

Follow-up univariate $F$ test results are found in Table 2.4. Tests are significant ($p < .0001$) for the four subscales with corresponding small to medium effect sizes ($\eta^2$), indicating individual differences between the five clusters. Significant ($p < .005$) differences between cluster groupings via Tukey's HSD tests are reported in Table 2.4. These results suggest that the faculty members who described courses in the five instructional styles have differing views on student learning and self-efficacy in enacting instructional practices.

**Table 2.4.** Mean scores for clusters on the four SBTL-I factors.

| | | | | | | ANOVA | | |
|---|---|---|---|---|---|---|---|---|
| | C1 | C2 | C3 | C4 | C5 | $F$ (4,1277) | $\eta^2$ (size) | Tukey HSD (.005) |
| Teacher-Centered Learning | 3.72 (0.51) | 3.81 (0.54) | 3.80 (0.54) | 3.85 (0.48) | 3.88 (0.45) | 5.33 | 0.02 (s) | ----- |
| Student-Centered Learning | 3.97 (0.40) | 4.10 (0.38) | 3.83 (0.38) | 3.89 (0.47) | 3.77 (0.43) | 22.97 | 0.07 (m) | 1&5, 2&3, 2&4, and 2&5 |
| Efficacy – Pedagogy | 4.05 (0.58) | 4.27 (0.52) | 3.92 (0.62) | 4.17 (0.62) | 3.90 (0.65) | 13.77 | 0.04 (s) | 1&2, 2&3, 2&5, 3&4, and 4&5 |
| Efficacy – Content | 4.09 (0.61) | 4.31 (0.54) | 4.08 (0.63) | 4.33 (0.57) | 4.09 (0.61) | 7.82 | 0.02 (s) | 1&2, 1&4, 2&3, 2&5, 3&4, and 4&5 |

*Note.* C1 = *Small Groups*; C2 = *Interactive*; C3 = *Lecture with Clickers*: C4 = *Lecture with Literature*; C5 = *Lecture*; $\eta^2$ (size) cut-off values: ≤ 0.01 small, ≤ 0.06 medium, ≤ 0.14 large (Cohen, 1973).

The two learning beliefs subscales provide insight to the differences between styles. Faculty members in the *Lecture with CRS* style have scores between the other clusters on both *Student-Centered* and *Teacher-Centered Learning* factors; faculty members in this cluster report mixed beliefs about the ways in which students learn best. Faculty members in the *Interactive* and *Small Groups* styles score the highest on the *Student-Centered Learning* factor. Differences on the *Student-Centered Learning* factor produced the highest effect size, indicating that differences between faculty members in this grouping are the largest in our sample.

In terms of the two self-efficacy subscales, faculty members in the *Lecture* and *Lecture with CRS* styles report the least confidence in their ability in terms of pedagogy, which is confirmed in their choice to use more traditional pedagogies. Those most confident in their pedagogical ability are those in the *Interactive* and *Lecture with Literature* styles. Faculty members in the *Lecture with Literature* style as well as the *Interactive* style feel the most strongly about their content ability. Those in the *Lecture* and *Small Groups* styles report the same, indicating that instructional style may be related to factors other than efficacy in content.

**Discussion**

This study is designed to elicit evidence of the link between enacted instructional practices and instructor thinking posited in the TCSR model of educational reform (Woodbury & Gess-Newsome, 2002). The results indicate that the way faculty members teach can be described by a coherent set of instructional styles (which align with observational findings), and that there is a significant difference of scores on a thinking instrument involving both learning beliefs and self-efficacy beliefs between faculty members who have differing instructional styles.

In the development of the TCSR model, qualitative research was conducted to identify potential links between cultural context, personal context, and teacher thinking (Gess-Newsome et al. 2003). One of the goals of this study is to apply the TCSR model to postsecondary chemistry and find an empirical link in a larger sample than past studies to support the robust nature of the model; such work responds to issues raised in the educational reform literature of avoiding over interpretation of non-generalizable qualitative results (Devlin, 2006). The data presented support a result similar to qualitative studies in which the beliefs and self-efficacy of instructors were associated with enacted instructional practices (Andrews & Lemons, 2015; Orgill et al., 2015; Roehrig & Kruse, 2005). Descriptive information about differences between faculty members who operate their classrooms in different ways is supportive of the TCSR model as a framework with which to structure future understanding of chemistry education. This result is similar to those found by others utilizing the TCSR model (Enderle et al., 2014; Enderle et al., 2013; Gibbons et al., 2017; Guerrero, 2010; Lund & Stains, 2015; Stains et al., 2015).

**Research question 1**

Our first research question consisted of determining the degree to which our data clustered into instructional styles. The resultant cluster groupings differed in their use of instructional practices as evidenced by significant Fisher's exact tests. Three of the five instructional styles utilized lecturing as the primary instructional technique; there were varying levels of incorporating student engagement along with the lecture in these styles including discussing the primary literature (*Lecture with Literature* style) and using CRS (*Lecture with CRS* style). Faculty members employing an *Interactive* style reported using demonstrations, small group work, and whole class discussions more than other respondents, indicating that the classes taught by these faculty members experienced a variety of activities that were incorporated in the course. Faculty members utilizing a *Small Groups* style assigned students to work in small groups more than other respondents.

These instructional styles align with the profiles found using observation in a variety of undergraduate science classrooms (Lund et al., 2015). The profiles defined in the Lund et al. (2015) study discerned between *Lecturing*, *Socratic* method (i.e., frequently asking questions), *Peer Instruction*, and *Collaborative Learning*. The alignment of the *Lecture* and *Lecture with Literature* styles found in this study with the *Lecture* profile demonstrate the continued use of didactic teaching methods in chemistry. The use of CRS as a tool to ask student questions and typically to encourage student interaction in the *Lecture with CRS* style is similar to the *Socratic* profile, in which faculty members begin to use more cooperative methods in the classroom. The *Interactive* and *Small Groups* styles found in this study represent faculty members who have adopted more active learning pedagogies in their classroom in a similar way to the *Peer Instruction* and *Collaborative Learning* profiles.

Being able to define instructional styles using a self-report instrument is an important step in furthering research on faculty members' use of pedagogical reforms; the results of this study support the use of a COPUS-based self-report measure in discerning instructional styles that can enable researchers to determine the state of classroom instruction. The problems faced by previous self-report instruments are not noticeable in this analysis, because the alignment with observational data indicates that we did not experience "social desirability bias" in our survey- many respondents still responded that they primarily lectured in their courses. This is an important finding in this context, because self-report instruments have been frequently criticized for not reflecting the reality of the educational environment. We suspect that this survey produced such a result because it was of low stakes to the respondents i.e., it was not conducted by their own institution and explored other areas as well as instructional practice.

**Research question 2**

Our second research question considered differences between faculty members in the resultant cluster groupings of instructional styles. The SBTL-I was developed to yield discernable scores between faculty members on their impressions about the ways that students learn best and their ability to facilitate student learning on four subscales: *Teacher-Centered Learning, Student-Centered Learning, Self-Efficacy in Pedagogy,* and *Self-Efficacy in Content*. Identifiable differences were observed between the instructional styles by mean scores on the four SBTL-I subscales.

Group differences on Self-Efficacy scores indicate meaningful information about faculty members who adopt different styles. Faculty members who are more confident in their ability in pedagogy use a wide variety of instructional techniques in the *Interactive* and *Lecture with Literature* style. However, there is no indication of a trend in the responses to this study between instructional style and efficacy in content or pedagogy across the increasing use of reformed classroom practices. For example, faculty members in the *Lecture* style and faculty members in the *Small Groups* style score similarly on the Self-Efficacy subscale even though these two styles are opposite with respect to pedagogical techniques. This indicates that the incorporation of chemistry literature into the classroom environment is conducted by instructors who are more confident in their ability to perform instruction. The incorporation of chemistry literature is challenging because these texts are not written in a format interpretable to the layperson or novice, especially an undergraduate student. A faculty member, then, must be confident in their ability to help students interpret the texts when used in classroom contexts. This is demonstrated by significant differences seen in the results of this study. Similarly, instructors who utilize a variety of classroom instructional methods in the *Interactive* style face challenges that require a higher level of confidence in their instructional ability; these include their use of demonstrations, which require a consideration for safety and preparation of the classroom.

Scores on the *Teacher-Centered Learning* and *Student-Centered Learning* subscales were significantly and importantly different. Faculty members in the *Interactive* and *Small Groups* styles hold beliefs in the arena of *Student-Centered Learning*, while those in the *Lecture*-based styles report stronger beliefs that students learn best in a teacher-centered environment. These findings support the initial understanding found in the literature that instructors' beliefs will align with their instructional choices. Based on the definition used in this study, the instructors whose classrooms reflect a reformed environment have more strongly held reform-minded beliefs about teaching and learning.

This link between beliefs about self-efficacy as an instructor and beliefs about how students learn best supports the posited relationship between teacher thinking and enacted instructional practices in the TCSR model. This finding is relevant because this study was conducted on a nationwide scale and used self-reported data, while confirming the theoretical link between beliefs and practice. This result provides support to the wide array of literature on the link between thinking and practice, but contributes a larger empirical base and opposes criticism of previous qualitative and small-scale quantitative studies. Our study also incorporates information found in the population of postsecondary chemistry faculty members; these individuals play a significant role in instructing prerequisite courses required of a variety of science fields at the postsecondary level.

**Implications for Research**

The results of this study inform those working on the development of reformed practices and curricula. Support for the TCSR model provided in this report should encourage others to adopt this framework in the design of innovative pedagogical techniques. Of importance is the consideration of beliefs in the construction of reformed pedagogy and curricula, and efficacy in implementing reform initiatives. While evidence may support the value of a new curriculum or pedagogy to increase student learning, many faculty members will not adopt a new technique because of their previously established belief systems (Addy & Blanchard, 2010). This is due to a complex array of factors, including strongly held beliefs about teaching and learning (Pajares, 1992). If faculty members believe that the best way for students to learn is through didactic teaching methods, those faculty members will continue to use such methods until they have a personal experience which indicates otherwise, as demonstrated in Andrews and Lemons (2015).

Developers of EBIPs must recognize the challenge that reform efforts hold for instructors who have become accustomed to traditional methods of instruction (Henderson et al., 2015). To disseminate evidence to support the impact of EBIPs, professional development programs are continually designed and offered to encourage adoption (Bauer et al., 2013; Enderle et al., 2014; Hutchins & Friedrichsen, 2012; Lakshmanan et al., 2011; Llawrenz, Huffman, & Gravely, 2007; Richards-Babb, Penn, & Withers, 2014; Stains et al., 2015). Professional development programs are best equipped to demonstrate the utility of EBIPs (Boz & Uzuntiryaki, 2006; Struyven et al., 2010). A portion of each reform effort aligned with such professional development must include an appeal to changing the beliefs held by faculty members, that is, the agents of instructional change (Sunal et al., 2001). Without a belief in the importance of student-centered learning and improved self-efficacy, teacher-centered instruction is more likely to continue to occur as evidenced by this study (Gess-Newsome et al., 2003; Lakshmanan et al., 2011).

Ultimately, these findings cast hope over the state of faculty members' thinking in postsecondary chemistry in the United States: *Student-Centered Learning* beliefs were overall more prevalent than *Teacher-Centered Learning* beliefs. This indicates some level of the cognitive dissonance from which to leverage reform efforts (Bauer et al., 2013; Greensfeld & Elkad-Lehman, 2007; Kane et al., 2002; Sandi-Urena, Cooper, & Gatlin, 2011; Windschitl & Sahl, 2002). Despite the slow spread of EBIPs in postsecondary chemistry, the characterization of instructional practices into styles provide chemistry education researchers with a framework to describe student-centered teaching in a context for future faculty members and to potentially identify transitional pedagogical techniques.

### Implications for Faculty Members

For chemistry faculty members, these results present a call for a focus of reform efforts on fostering change from its core, that is, the beliefs of those who will ultimately adopt the change in their daily experiences. Using the TCSR model as a framework, we can better understand what happens in the classrooms that we seek to improve. Based on the empirical link between thinking and practice demonstrated here and in previous literature, we encourage faculty members to consider how they believe students learn best and how their beliefs align with their practices, that is, we encourage reflective practice (Kane, Sandretto, & Heath, 2004). One crucial aspect of faculty members' change highlighted in the TCSR model and in other studies on teacher change is dissatisfaction (Bauer et al., 2013; Windschitl & Sahl, 2002). Faculty members are unlikely to change their classroom style without feeling unhappy with current practice. Reflection during adoption of a reform is imperative to nurturing the sense of dissatisfaction that leads faculty members to embrace change and encourage understanding of reformed instruction (Greensfeld & Elkad-Lehman, 2007; Kane et al., 2004; Sandi-Urena et al., 2011).

**Limitations**

While the results of this study provide information to inform and support use of the TCSR model in postsecondary science education research and reform efforts, there are areas for improvement. Primarily, the TCSR model includes other important factors to consider when considering systemic change in the classroom outside of the teacher thinking as evaluated in our work. For example, cultural context is influenced by extra-institutional, institutional, and departmental factors that enable or disable faculty members' participation in pedagogical reform (Henderson et al., 2015; Woodbury & Gess-Newsome, 2002). The personal context of a faculty member including the way that they learned the content, their participation in teaching professional development workshops, and their content knowledge influences the way that they think about teaching and learning as well as their instructional style (Lakshmanan et al., 2011; Veal, 2004). These factors are outlined in our theoretical framework and confound our results. While the study described in this report was designed to evaluate only one aspect of the TCSR model, future studies should incorporate measures of cultural factors, personal factors, and thinking factors to understand better faculty members' practices.

Secondly, the mechanism for capturing enacted instructional practices used in this study loses some empirical strength as a result of self-report (D'Eon et al., 2008; Ebert-May et al., 2011; Herrington et al., 2016); however, the ability to gain understanding from a larger subset of the population is essential to identify relationships that exist between thinking and practice (Henderson & Dancy, 2009; Williams et al., 2015). There are significant resource barriers to conducting an observational study of a subset of the entire population of interest, and such studies must account for increased measurement error across multiple raters. Therefore, our instrument was intentional to minimize self-report error. Primarily, the inclusion of five response categories for the use of each pedagogical technique (rather than a binary use/non-use answer option) allowed faculty members to report some use while avoiding "social desirability bias" (Krumpal, 2013). If a faculty member felt that the inclusion of small group work or whole class discussions, for example, are preferred by those who administered this survey, they might be more likely to select the use of techniques, even though they do not incorporate them into their enacted instructional practices. While we cannot know if this occurred, we accommodated for such a possibility by offering a "several times during the semester" and "rarely" option. These options allowed faculty member to respond in a manner that better reflected their real classroom practices. Another advantage of the self-report instrument used is its alignment with the COPUS observation protocol; this allows for future work in which observations of classes can be associated with self-report data to provide validity information for the instrument itself as well as findings from similar studies. The results of this study should not be discounted due to the inclusion of a self-report variable; the strength of instrument design and the alignment of instructional styles with a robust observational study support the validity of our results.

**Future Work**

Our survey asked faculty members to describe one course for which they had taught over the past three years and in which they had the most perceived influence. This provides information about the environment in which, logically, a faculty member can enact ideal and desired practices; thus, the course chosen best reflects the relationship of interest in this study. Because of the perceived control over the course, we begin to understand the way that a faculty member would construct a course if allowed to do so. In a different context, a faculty member's beliefs might not align with the instructional practices used, and future work should address such instances. It has been found that faculty members' thoughts about teaching and learning do not always align with their claims about what occurs in their idealized classroom (Veal et al., 2016); this finding should also be considered in future research regarding instructional styles and their relationships with beliefs. Comparisons between courses can provide insight to the differences between course contexts and instructional styles used (Walczyk & Ramsey, 2003). This will be helpful in future analyses of the differences between courses taught at varying levels, and to those who do not major in science fields compared to science majors.

## Conclusion

The results presented in this report of a national survey of postsecondary chemistry faculty members provide empirical evidence for the link between faculty members' thinking and enacted instructional practice. Discernment of instructional styles from a set of instructional practices allows us to understand more about the techniques used by faculty members in courses for which they have the highest level of perceived influence. These instructional styles align with those found in a study of observational data, indicating a common pattern of instructional activities. This study found a difference in the scores on an instrument designed to measure faculty members' thinking about teaching and learning as well as efficacy in pedagogy and content between faculty members who teach using different instructional styles. This result is important for understanding the spread of curricular reform; significant differences support the often empirically unsupported claim that faculty members' thinking and practice are related. Our results support the use of the TCSR model as a framework to develop instructional reforms and encourage us to further consider the multifaceted nature of reform when working with faculty members. Our results should empower developers of pedagogical and curricular reforms to consider how beliefs and efficacy are influencing the growth of educational reforms. The results presented inform our perspective of reforms inside our own classrooms.

## References

Addy, T. M., & Blanchard, M. R. (2010). The problem with reform from the bottom up: Instructional practices and teacher beliefs of graduate teaching assistants following a reform-minded university teacher certificate program. *International Journal of Science Education*, *32*(8), 1045-1071. doi:10.1080/09500690902948060

Andrews, T. C., & Lemons, P. P. (2015). It's personal: Biology instructors prioritize personal evidence over empirical evidence in teaching decisions. *CBE- Life Sciences Education*, *14*(1), ar7. doi:10.1187/cbe.14-05-0084

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), Self-Efficacy Beliefs of Adolescents (pp. 307-337). Greenwich, CT: Information Age Publishing.

Bauer, C., Libby, R. D., Scharberg, M., & Reider, D. (2013). Transformative research-based pedagogy workshops for chemistry graduate students and postdocs. *Journal of College Science Teaching*, *43*(2), 36-43.

Borrego, M., Froyd, J. E., Henderson, C., Cutler, S., & Prince, M. (2013). Influence on engineering instructors' teaching and learning beliefs on pedagogies in engineering science courses. *International Journal of Engineering Education, 29*(6), 1456-1471.

Boz, Y., & Uzuntiryaki, E. (2006). Turkish prospective chemistry teachers' beliefs about chemistry teaching. *International Journal of Science Education, 28*(14), 1647-1667. doi: 10.1080/09500690500439132

Cohen, D. K., & Mehta, J. D. (2017). Why reform sometimes succeeds: Understanding the conditions that produce reforms that last. *American Educational Research Journal*, *54*(4), 644-690. doi:10.3102/0002831217700078

Cohen, J. (1973). Eta-squred and partial eta-squared in fixed factor ANOVA designs. E*ducational and Psychological Measurement, 33*(1), 107-112. doi: 10.1177/001316447303300111

Colbeck, C., Cabrera, A. F., & Marine, R. (2002). Faculty motivation to use alternative teaching methods Paper presented at the Annual Conference of the American Educational Research Association, New Orleans, LA.

D'Eon, M., Sandownik, L., Harrison, A., & Nation, J. (2008). Using self-assessments to detect workshop success: Do they work? *American Journal of Evaluation, 29*(1), 92-98. doi: 10.1177/1098214007312630

DeVelis, R. F. (2017). Scale development theory and applications (4 ed.). Thousand Oaks: Sage.

Devlin, M. (2006). Challenging accepted wisdom about the place of conceptions of teaching in university teaching improvement. International *Journal of Teaching and Learning in Higher Education, 18*(2), 112-119.

Dole, J. A., & Sinatra, G. M. (1998). Reconceptualizing change in the cognitive construction of knowledge. *Educational Psychologist*, *33*(2/3), 109-128. doi:10.1080/00461520.1998.9653294

Duda, R. O., Hart, P. E., & Stork, D. G. (2001). Pattern Classification (2 ed.). New York, NY: Wiley.

Ebert-May, D., Derting, T. L., Hodder, J., Momsen, J. L., Long, T. L., & Jardeleza, S. E. (2011). What we say is not what we do: Effective evaluation of faculty professional development programs. *BioScience, 61*(7), 550. doi:10.1525/bio.2011.61.7.9

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. doi:10.1021/ed200207p

Enderle, P. J., Dentzau, M., Roseler, K., Southerland, S., Granger, E., Hughes, R., . . . Saka, Y. (2014). Examining the influence of RETs on science teacher beliefs and practice. *Science Education*, *98*(6), 1077-1108. doi:10.1002/sce.21127

Enderle, P. J., Southerland, S. A., & Grooms, J. A. (2013). Exploring the context of change: Understanding the kinetics of a studio physics implementation effort. *Physical Review Special Topics- Physics Education Research*, *9*(1), 010114. doi: 10.1103/PhysRevSTPER.9.010114

Feldman, A. (2000). Decision making in the practical domain: A model of practical conceptual change. *Science Education, 84*(5), 606. doi:10.1002/1098-237X(200009)84:5<606::AID-SCE4>3.0.CO;2-R

Fisher, R. A. (1922). On the interpretation of Chi-squared from contingency tables, and the calculation of P. *Journal of the Royal Statistical Society, 85*(1), 87-94. doi:10.2307/2340521

Gallos, M. R., van den Berg, E., & Treagust, D. F. (2005). The effect of integrated course and faculty development: Experiences of a university chemistry department in the Philippines. *International Journal of Science Education*, *27*(8), 985-1006. doi: 10.1080/09500690500038447

Gess-Newsome, J., Southerland, S., Johnston, A., & Woodbury, S. (2003). Educational reform, personal practical theories, and dissatisfaction: The anatomy of change in college science teaching. *American Educational Research Journal, 40*(3), 731-767. doi: 10.3102/00028312040003731

Gibbons, R. E., Laga, E. E., Leon, J., Villafañe, S. M., Stains, M., Murphy, K., & Raker, J. R. (2017). Chasm crossed? Clicker use in postsecondary chemistry education. *Journal of Chemical Education*, *94*(5), 549-557. doi:10.1021/acs.jchemed.6b00799

Glass, G. C., & Hopkins, K. (1984). Statistical methods in education and psychology. Englewood Cliffs, NJ: Prentice-Hall.

Gow, L., & Kember, D. (1993). Conceptions of teaching and their relationship to student learning. *British Journal of Educational Psychology*, *63*(1), 20-23. doi:10.1111/j.2044-8279.1993.tb01039.x

Graves, L. A., Hughes, H., & Balgopal, M. M. (2016). Teaching STEM through horticulture: Implementing an edible plant curriculum at a STEM-centric elementary school. J*ournal of Agricultural Education, 57*(3), 192-207. doi:10.5032/jae.2016.03192

Greensfeld, H., & Elkad-Lehman, I. (2007). An analysis of the processes of change in two science teachers educators' thinking. *Journal of Research in Science Teaching, 44*(8), 1219-1245. doi:10.1022/tea.20185

Guerrero, S. (2010). The role of instructor thinking in technology-based reform: A multiple case study. *Journal of the Research Center for Educational Technology, 6*(2), 18-30.

Haney, J., Lumpe, A., Czerniak, C., & Egan, V. (2002). From beliefs to actions: The beliefs and actions of teachers implementing change. *Journal of Science Teacher Education, 13*(3), 171-187. doi:10.1023/A:1016565016116

Harshman, J., & Stains, M. (2017). A review and evaluation of the internal structure and consistency of the Approaches to Teaching Inventory. *International Journal of Science Education, 39*(7), 918-936. doi:10.1080/09500693.2017.1310411

Harwood, W. S., Hansen, J., & Lotter, C. (2006). Measuring teacher beliefs about inquiry: The development of a blended qualitative/quantitative instrument. *Journal of Science Education and Technology, 15*(1), 69-79.

Henderson, C. (2008). Promoting instructional change in new faculty: An evaluation of the physics and astronomy new faculty workshop. *American Journal of Physics, 76*(2), 179-187. doi:https://doi.org/10.1119/1.2820393

Henderson, C., Beach, A. L., & Finkelstein, N. D. (2011). Facilitating change in undergraduate STEM instructional practices: An analytic review of the literature *Journal of Research in Science Teaching, 48(*8), 952-984. doi:10.1002/tea.20439

Henderson, C., Cole, R., Froyd, J. E., Friedrichsen, D. G., Khatri, R., & Stanford, C. (2015). Designing educational innovations for sustained adoption: A how-to guide for education developers who want to increase the impact of their work.

Henderson, C., & Dancy, M. H. (2009). Impact of physics education research on the teaching of introductory quantitative physics in the United States. *Physical Review Special Topics- Physics Education Research, 5*(2). doi:10.1103/PhysRevSTPER.5.020107

Herrington, D. G., Yezierski, E. J., & Bancroft, S. F. (2016). Tool trouble: Challenges with using self-report data to evaluate long-term chemistry teacher professional development. *Journal of Research in Science Teaching, 53*(7), 1055-1081. doi:10.1002/tea.21323

Ho, A., Watkins, D., & Kelly, M. (2001). The conceptual change approach to improving teaching and learning: An evaluation of a Hong Kong staff program. *Higher Education, 42*(2), 143-169. doi:10.1023/A:1017546216800

Houseal, A. K., Abd-El-Khalick, F., & Destefano, L. (2014). Impact of a student-teacher-scientist partnership on students' and teachers' content knowledge, attitudes toward science, and pedagogical practices *Journal of Research in Science Teaching, 51*(1), 84-115. doi:10.1002/tea.21126

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modelling: Multidisciplinary Journal, 6,* 1-55. doi:10.1080/10705519909540118

Hutchins, K. L., & Friedrichsen, P. J. (2012). Science faculty belief systems in a professional development program: Inquiry in college laboratories. *Journal of Science Instructor Education, 23*(8), 867-887. doi:10.1007/s10972-012-9294-z

Kane, R., Sandretto, S., & Heath, C. (2002). Telling half the story: A critical review of research on the teaching beliefs and practices of university academics. *Review of Educational Research, 72*(2), 177-228. doi:10.3102/00346543072002177

Kane, R., Sandretto, S., & Heath, C. (2004). An investigation into excellent tertiary teaching: Emphasising reflective practice. *Higher Education, 47*(3), 283-310. doi:10.1023/B:HIGH.0000016442.55338.24

Kim, J.-O., & Mueller, C. W. (1978a). Factor Analysis: Sage.

Kim, J.-O., & Mueller, C. W. (1978b). Introduction to Factor Analysis: Sage.

Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality & Quantity, 47*(4), 2025-2047. doi:https://doi.org/10.1007/s11135-011-9640-9

Lakshmanan, A., Heath, B. P., Perlmutter, A., & Elder, M. (2011). The impact of science content and professional learning communities on science teaching efficacy and standards-based instruction. *Journal of Research in Science Teaching, 48*(5), 534-551. doi:10.1002/tea.20404

Llawrenz, F., Huffman, D., & Gravely, A. (2007). Impact of the collaboratives for excellence in teacher preparation program. *Journal of Research in Science Teaching, 44*(9), 1348-1369. doi:10.1002/tea.20207

Lotter, C., Harwood, W. S., & Bonner, J. J. (2007). The influence of core teaching conceptions on teachers' use of inquiry teaching practices *Journal of Research in Science Teaching, 44*(9), 1318-1347. doi:10.1002/tea.20191

Luft, J., & Roehrig, G. (2007). Capturing science instructors' epistemological beliefs: The development of the instructor beliefs interview. *Electronic Journal of Science Education, 11*(2), 38-63.

Lund, T., Pilarz, M., Velasco, J. B., Chakraverty, D., Rosploch, K., Undersander, M., & Stains, M. (2015). The best of both worlds: Building on the COPUS and RTOP observation protocols to easily and reliably measure various levels of reformed instructional practice. *CBE- Life Sciences Education, 14(*2). doi:10.1187/cbe.14-10-0168

Lund, T., & Stains, M. (2015). The importance of context: An exploration of factors influencing the adoption of student-centered teaching among chemistry, biology, and physics faculty. *International Journal of STEM Education, 2*(13). doi:10.1186/s40594-015-0026-8

MacArthur, J. (2013). How will classroom response systems "cross the chasm"? *Journal of Chemical Education, 90*(3), 273-275. doi:10.1021/ed300215d

Mansour, N. (2013). Consistencies and inconsistencies between science teachers' beliefs and practices *International Journal of Science Education, 35*(7), 1230-1275. doi:10.1080/09500693.2012.743196

Marbach-Ad, G., Ziemer, K. S., Orgler, M., & Thompson, K. V. (2014). Science teaching beliefs and reported approaches within a research university: Perspectives from faculty, graduate students, and undergraduates *International Journal of Teaching and Learning in Higher Education, 26*(2), 232-250.

Mesa, V., Celis, S., & Lande, E. (2014). Teaching approaches of community college mathematics faculty: Do they relate to classroom practices? *American Educational Research Journal, 51*(1), 117-151. doi:10.3102/0002831213505759

Moore, G. A. (2002). Crossing the chasm. New York: Collins.

Mutambuki, J., & Fynewever, H. (2012). Comparing chemistry faculty beliefs about grading with grading practices. *Journal of Chemical Education, 89*(3), 326-334. doi:10.1021/ed1000284

Muthén, L. K. & Muthén, B. D. (1998-2017). *Mplus users guide. Seventh edition.* Los Angeles, CA: Muthén & Muthén.

National Research Council. (2012). Discipline-based education research: Understanding and improving learning in undergraduate science and engineering. Washington, DC: The National Academies Press.

Orgill, M., Bussey, T. J., & Bodner, G. M. (2015). Biochemistry instructors' perceptions of analogies and their classroom use. *Chemistry Education Research and Practice, 16*(4), 731-746. doi:10.1039/C4RP00256C

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research, 62(*3), 307-332. doi:10.3102/00346543062003307

Pilburn, M., Sawada, D., Turley, J., Falconer, K., Benford, R., I., B., & Judson, E. (2000). Reformed Teaching Observation Protocol (RTOP) reference manual. Tempe, AZ: Arizona Collaborative for Excellence in the Preparation of Teachers

Raker, J. R., & Holme, T. A. (2014). Investigating faculty familiarity with assessment terminology by applying cluster analysis to interpret survey data. *Journal of Chemical Education, 91*(8), 1145-1151.

Richards-Babb, M., Penn, J. H., & Withers, M. (2014). Results of a practicum offering teaching-focused graduate student professional development. *Journal of Chemical Education, 91*(11), 1867-1873. doi:10.1021/ed500134d

Riggs, I., & Enochs, L. (1990). Towards the development of an elementary teacher's science teaching efficacy belief instrument. *Science Education, 74*(6), 625-637. doi:10.1002/sce.3730740605

Roehrig, G., Kruse, R. A., & Kern, A. (2007). Teacher and school characteristics and their impact on curriculum and instruction. *Journal of Research in Science Teaching, 44*(7), 887-907. doi:10.1002/tea.20180

Roehrig, G. H., & Kruse, R. A. (2005). The role of teachers' beliefs and knowledge in the adoption of a reform-based curriculum. *School Science and Mathematics, 105*(8), 412-422. doi:10.1111/j.1949-8594.2005.tb18061.x

Rogers, E. M. (1995). Diffusion of Innovation (4 ed.). New York, NY: Free Press.

Sandi-Urena, S., Cooper, M., & Gatlin, T. A. (2011). Graduate teaching assistants' epistemological and metacognitive development. *Chemistry Education Research and Practice, 12*(1), 92-100. doi:10.1039/C1RP90012A

Sawada, D., Pilburn, M., Judson, E., Turley, J., Falconer, K., Benford, R., & Bloom, I. (2002). Measuring reform practices in science and mathematics classrooms: The reformed teaching observation protocol. *School Science and Mathematics, 102*(6), 245-253. doi:10.1111/j.1949-8594.2002.tb17883.x

Smith, L. K., & Southerland, S. (2007). Reforming practice or modifying reforms?: Elementary teachers' response to tools of reform. *Journal of Research in Science Teaching, 44*(3), 396-423. doi:10.1002/tea.20165

Smith, M. K., Jones, F. H. M., Gilbert, S. L., & Wieman, C. E. (2013). The Classroom Observation Protocol for Undergraduate STEM (COPUS): A new instrument to characterize university STEM classroom practices. *CBE- Life Sciences Education, 12*(4), 618-627. doi:10.1187/cbe.13-08-0154

Smolleck, L. A., & Yoder, E. P. (2008). Further development and validation of the teaching science as inquiry (TSI) instrument. *School Science and Mathematics, 108*(7), 291-297.

Smolleck, L. A., Zembal-Saul, C., & Yoder, E. P. (2006). The development and validation of an instrument to measure preservice teachers' self-efficacy in regard to the teaching of science as inquiry. *Journal of Science Teacher Education, 17*(2), 137-163. doi:10.1007/s10972-006-9015-6

Stains, M., Pilarz, M., & Chakraverty, D. (2015). Short and long-term impacts of the Cottrell Scholars collaborative new faculty workshop. *Journal of Chemical Education, 92*(9), 1466-1476. doi:10.1021/acs.jchemed.5b00324

StataCorp. (2015). Stata Statistical Software: Release 14. College Station, TX: StataCorp, LP.

Stevens, J. P. (2009). Applied multivariate statistics for the social sciences. New York, NY: Routledge.

Struyven, K., Dochy, F., & Janssens, S. (2010). 'Teach as you preach': The effects of student-centered versus lecture-based teaching on student teachers' approaches to teaching. *International Journal of Science Education, 33*(1), 43-64. doi:10.1080/02619760903457818

Sunal, D. W., Hodges, J., Sunal, C. S., Whitaker, K. W., Freeman, L. M., Edwards, L., . . . Odell, M. (2001). Teaching science in higher education: Faculty professional development and barriers to change. *School Science and Mathematics, 101*(5), 246. doi:10.1021/ed9000624

Towns, M. (2010). Crossing the chasm with classroom response systems. *Journal of Chemical Education, 87*(12), 1317-1319.

Trigwell, K., & Prosser, M. (2004). Development and use of the approaches to teaching inventory. *Educational Psychology Review, 16*(4), 409-424. doi:10.1007/s10648-004-0007-9

Tukey, J. (1949). Comparing individual means in the analysis of variance. *Biometrics, 5*(2), 99-114. doi:10.2307/3001913

Vacha-Hasse, T., & Thompson, B. (2004). How to estimate and interpret various effect sizes. *Journal of Counseling Psychology, 51(*4), 473-481. doi:10.1037/0022-0167.51.4.473

Veal, W. R. (2004). Beliefs and knowledge in chemistry teacher development. *International Journal of Science Education, 26*(3), 329-351. doi:10.1080/0950069032000097389

Veal, W. R., Riley Lloyd, M. E., Howell, M. R., & Peters, J. (2016). Normative beliefs, discursive claims, and implementation of reform-based science standards. *Journal of Research in Science Teaching, 53*(9), 1419-1443. doi:10.1002/tea.21265

Walczyk, J. J., & Ramsey, L. L. (2003). Use of learner-centered instruction in college science and mathematics classrooms. *Journal of Research in Science Teaching, 40*(6), 566-584. doi:10.1002/tea.10098

Walczyk, J. J., Ramsey, L. L., & Zha, P. (2007). Obstacles to instructional innovation according to college science and mathematics faculty. *Journal of Research in Science Teaching, 44*(1), 85-106. doi:10.1002/tea.20119

Walter, E. M., Beach, A. L., Henderson, C., & Williams, C. T. (2014). Describing instructional practice and climate: Two new instruments. Paper presented at the Transforming Institutions: 21st Century Undergraduate STEM Education Conference, Indianapolis, IN.

Ward, J. H. J. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association, 58(*301), 236-244. doi:10.2307/2282967

Wieman, C. E., & Gilbert, S. L. (2014). The teaching practices inventory: A new tool for characterizing college and university teaching in mathematics and science. *CBE- Life Sciences Education, 13*(3), 552-569. doi:10.1187/cbe.14-02-0023

Williams, C. T., Walter, E. M., Henderson, C., & Beach, A. L. (2015). Describing undergraduate STEM teaching practices: A comparison of instructor self-report instruments. *International Journal of STEM Education, 2*(18), 1-14. doi:10.1186/s40594-015-0031-y

Windschitl, M., & Sahl, K. (2002). Tracing teachers' use of technology in a laptop computer school: The interplay of teacher beliefs, social dynamics, and institutional culture. *American Educational Research Journal, 39*(1), 165-205. doi:10.3102/00028312039001165

Woodbury, S., & Gess-Newsome, J. (2002). Overcoming the paradox of change without difference: A model of change in the arena of fundamental school reform. *Educational Policy, 16,* 763-782. doi:10.1177/089590402237312

CHAPTER THREE:

CHASM CROSSED? CLICKER USE IN POSTSECONDARY CHEMISTRY EDUCATION

**Note to Reader**

This chapter is a published manuscript in the *Journal of Chemical Education*. Reprinted with permission from Gibbons, R. E., Laga, E. E., Leon, J., Villafañe, S. M., Stains, M., Murphy, K., & Raker, J. R. (2017) Chasm crossed? Clicker use in postsecondary chemistry education. *Journal of Chemical Education*, 94(5), 549-557. DOI: 10.1021/acs.jchemed.6b00799. Copyright 2017 American Chemical Society. Permissions information can be found in Appendix B. This work was published with co-authors. Emily E. Laga and Jessica Leon were undergraduate researchers and contributed to the collection of the sample frame for the survey. Sachel M. Villafañe was involved in the development of the instrument for the survey. Marilyne Stains was instrumental in the selection of the theoretical framework for the survey development. Kristen Murphy is the Director of the ACS Examinations Institute, and therefore provided administrative support to the study. Jeffrey R. Raker is the technical and administrative supervisor of the survey used in this report.

**Introduction**

The purpose of this study is to examine the adoption of classroom response systems (CRSs) in undergraduate chemistry classrooms in the United States. CRSs have been described as applicable for all educational contexts in chemistry (Sevian & Robinson, 2011). In order to achieve the broadest understanding of our community's adoption of CRS, we analyze data from a national survey of chemistry faculty, and we consider CRS use based on factors that contribute to adoption of new pedagogies and technologies. This work addresses questions left unanswered by prior work about the contexts where CRS are most prevalently used. In addition, we have greatly improved upon prior survey methodologies used in chemical education research. Through a rigorous stratified sampling strategy and by weighting our survey data, we are able to provide confidence intervals that account for sampling error and the presence of nonresponse bias in our data.

**Classroom Response Systems**

Classroom response systems, i.e., clickers and personal device response systems (e.g., smartphones and tablet-style computers), are one of many technology-based systemic reform tools (MacArthur & Jones, 2008). While CRSs were originally marketed as a tool with the ability to change the way technology is used in *all* classroom settings, in fields other than postsecondary chemistry education, a leveling out of CRS adoption has been reported (Henderson, Dancy, & Niewiadomska-Bugaj, 2012). The technology goes by many names (MacArthur & Jones 2008); "classroom response systems" is used throughout this paper to refer to a technology in which students have an individual device with which they answer questions in real time via the Internet, radio or infrared frequencies, but we note that the term "clicker" is used synonymously with CRS in the literature and education communities. CRS technology now includes software that harnesses cell phones, laptops, and hard-wired systems. CRSs are ultimately tools designed to provide immediate feedback on student learning and encourage student collaboration. CRSs have been used in an array of classrooms for several purposes, including formative and summative assessment, collaborative learning, and taking attendance (MacArhtur & Jones, 2008).

**Chemical Education Reform**

The development of CRSs is linked to the reform movement in postsecondary chemistry classrooms. When student-centered instruction and curricula supported by experimental evidence are implemented, greater student learning occurs (Childs, 2009). Such pedagogies are broadly defined as evidence-based instructional practices (EBIPs, Stains, Pilarz, & Chakraverty, 2015). While the use of a

CRS alone is not an EBIP, some EBIPs incorporate the use of CRSs as a strategy to improve assessment and increase student collaboration (MacArthur & Jones, 2008). For example, CRSs have been incorporated into the Process-Oriented Guided Inquiry Learning pedagogical approach (MacArthur & Jones, 2008), PhET simulation approach (MacArthur & Jones, 2013), and in adapted versions of the flipped classroom approach (Chen, Stelzer, & Gladding, 2010; Phillis, Brewer, Hoogendyk, Goodwin, & Carter). Sevian and Robinson (2011) argued for the applicability of CRSs to all educational situations, providing evidence of the effectiveness of CRS in enhancing student learning.

CRSs have been used as a tool to create more active learning environments, but not at the rate developers expected. Despite growing evidence of the benefits of CRSs, faculty members report obstacles to CRS adoption. Roadblocks include lack of support and challenges due to the demographics of the institutional environment (Woodbury & Gess-Newsome, 2002). A CRS requires time to learn and implement; CRSs are often shelved in favor of more traditional methods (Koenig, 2010).

Despite evidence of effectiveness, most reforms do not create systemic change without sustained adopter support (Henderson et al., 2015). Integrated support is noted by the Increase the Impact research team who cite a “lack of dynamic development of techniques after initial interest” as a reason for the drop-off in implementation in many EBIPs (Henderson et al., 2015). Khatri et al. (2016) published a guide for developers to better disseminate and propagate reform initiatives; their recommendation is for disseminators to take into account the individuals who will be changing their classroom practices along with the departmental, institutional, and extra-institutional contexts for change when developing and propagating reforms. Support and tools for implementing CRSs exist; however, widespread adoption has not occurred. Considering the bulk of research on the effectiveness of CRSs has been done in large lecture courses, we hypothesize that faculty have determined that CRSs are only useful in a limited context.

**Theoretical Framework**

Chemical education researchers have been interested in the way CRSs have been implemented into classrooms and whether adoption will "take off" as expected (Towns, 2010; Emenike & Holme, 2012; MacArthur, 2013). Rogers' (1995) technology adoption life cycle (TALC) has been used to understand CRS adoption in chemistry classrooms (Towns, 2010). The TALC was developed by Rogers through a review of technological advances in multiple fields; the model broadly describes the growth of prior technologies in order to better understand future technological developments (Rogers, 1995). Technology adoption involves features of the adopters and the innovation along five stages (see Figure 3.1).

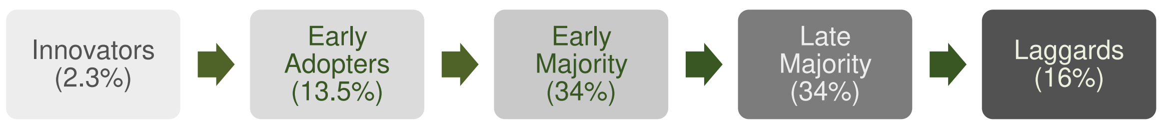


**Figure 3.1**. Five-stage technology adoption life cycle (TALC).

Each stage includes a profile of who is likely to adopt the technology. The first adopters are the *Innovators*; they are excited about new technology and serve as *β*-testing agents to determine viability in the field. The *Early Adopters* are more likely to implement a new technology once the most prominent errors have been corrected; this group's membership is frequently considered as change agents who have leadership capabilities in their institution allowing them to adopt new technologies to demonstrate to colleagues. The more pragmatic *Early Majority* waits to see evidence that the new technology supports a desirable outcome such as learning; their confidence in the technology is a necessity before adoption. The *Late Majority* are only likely to implement a new technology after it has become the norm; *Laggards* are not likely to adopt if another option is left in the market (Towns, 2010; Rogers, 1995; Moore, 2002). These profiles help determine how to promote technology adoption.

Rogers (1995) assigned a percentage to each stage on the basis of what has been seen with technological innovations (see Figure 3.1). Marketing researcher Moore (2002) noted a gap in the TALC (i.e., a chasm) between the Early Adopters and the larger Early Majority of users, where total technology use jumps from 16% to 50% of a population. This chasm has been referenced when considering CRS adoption in chemical education. Emenike and Holme (2012) reported the current adoption of CRS at 18.6%. On the basis of percent adoption, these results indicated that CRS adoption fits between the Early Adopters and Early Majority stages; the authors declared that the chasm had not yet been crossed. Researchers have argued that crossing the chasm will be difficult because of faculty unwillingness to welcome new technologies into teaching practices (Towns, 2010; MacArthur, 2013). We argue herein that faculty members have implemented CRSs in learning environments where CRSs are believed to be useful (i.e., large lecture courses) and thus that the alignment of the CRS characteristics with instructional context is the main driver of adoption. We hypothesize little movement in the overall percent adoption of CRSs; however, we expect to observe that the chasm has been crossed with greater than 50% adoption when considering contextual factors (e.g., course size).

**Features of the Innovation**

Rogers' (1995) theory of diffusion of innovations includes characteristics of the technology that are key to the decision to adopt. An innovation's relative advantage, compatibility, complexity, divisibility, and communicability influence its rate of adoption. Emenike and Holme (2012) hypothesized that there are theoretical differences in adoption based on course size and institution type that may determine the utility of CRSs in specific environments. Exploring the characteristics of the innovation through the lens of differing contexts will give the analysis of this survey data perspective.

Relative advantage is a description of how much better the new innovation is at achieving the goals of the adopters than current technologies (Rogers, 1995). The relative advantage of CRSs with respect to other methods of assessment has been evaluated along with the ability of CRSs to assist in the facilitation of group work and collaboration. Despite the variety of devices supported, clickers are the most studied. Mazur introduced this technology in his Harvard physics classrooms in 1991 (Crouch & Mazur, 2001); since then, clickers have been adopted and studied across the academy. Vickrey et al. (2015) found that when incorporated with the evidence-based practice of peer instruction, a CRS can increase performance outcomes. Smith et al. (2009) and Asirvatham (2005) both found similar positive impacts from CRS use when exploring in-class collaboration. MacArthur and Jones (2008) found in general chemistry courses positive increases in pass rates in courses with regular clicker use (Poulis, Massen, Robens, & Gilbert, 1998; Hall, Collier, Thomas, & Hilgers, 2005); and attributed this to formative assessment techniques facilitated by clickers. MacArthur and Jones (2008) subsequently adopted the technology in a large-enrollment general chemistry course of their own, finding significant success in learning outcomes. Addison, Wight, and Milner (2009), however, found neither improved nor decreased content learning between classrooms using and not using clickers, but student reports of perceived learning indicated that students felt clickers increased their course involvement and learning. MacArthur and Jones (2008) found that eight of 12 studies regarding CRS use were in courses with a large number of students (i.e., 75+ students). In comparison with other formative assessment techniques and tools to foster active learning, clickers have been shown to improve the educational experience for students. A manual for CRS use in chemistry from Asirvatham (2010) also highlighted the use of these technologies in classrooms of 100 or more.

The compatibility of an innovation relates the technology to the norm of the environment in which it is being adopted. The accessibility of CRSs for formative and summative assessment and encouraging student collaboration has changed over time. The price and inconvenience for students and instructors to learn how to use certain technologies has been noted as a reason many refuse to adopt clickers. In response, a range of technologies (e.g., WebCT) have been developed that accomplish the same task without the additional cost of a clicker device. Bunce et al. (2006) compared clickers and WebCT, a program using cell phones, giving more clout to web-based student response systems. Lee et al. (2013) used student cell phones as clickers and found that the technology can be implemented successfully. Shea (2016) established that there are a wide range of techniques for incorporating CRSs in the classroom, with software being continually updated. Both the device type and question type are variable within the use of CRSs, despite most users' assumption that only multiple-choice questions are available for use with the technology. As noted by Seery (2013), not only are multiple-choice questions appropriate with CRSs, but most of the technology can also be used with a variety of questioning options. As these aspects of CRSs have developed, they have become more compatible with tools that faculty are comfortable with.

The complexity of an innovation is a description of how challenging a new technology is to learn. A CRS is not necessarily more complex than other assessment forms; part of the relative advantage of a CRS is how quickly assessments can be analyzed compared to paper-based assessments. New software enhances this ability. A common inhibition is the time to develop appropriate CRS questions to ask during lecture and the time to incorporate the use of such questions in to preexisting lectures (Koenig, 2010). Bruck and Towns (2009) evaluated the types of clicker questions asked in a general chemistry course and found that students are more successful at answering lower-level cognitive-function-based questions when in collaboration with others. A manual on CRS use in large lecture courses from Asirvatham (2010) provides example questions that encourage faculty to incorporate higher-order thinking skills in to their in-class CRS questions, including visualization and problem-solving skills. Woelk (2008) developed a taxonomy of CRS questions to alleviate the question creation concern; most textbooks also now come with CRS questions (Towns, 2010).

The divisibility of an innovation is how likely it is for an individual to trial test the technology. In regard to CRSs, divisibility is sometimes cost-prohibitive. Increases in adoption costs are seen by many as too high to be worthwhile. Koenig (2010) outlined barriers with physics instructor colleagues for whom the clicker technology was provided and found that colleagues were not likely to implement the technology if their institution did not provide the required devices; when devices were purchased by the department, CRSs were seen as a useful tool. The cost of incorporating a CRS for the first time is high without department funding and encouragement.

Communicability is how easily results can be shared with others. Social interaction between faculty members allows for communication about CRS success and failure. CRSs have been used for over 20 years; information about their use is widespread. There is support for the successful adoption of CRSs as a pedagogical tool; more pragmatic educators in the *Early Majority* have many resources with which to understand the applicability of CRSs in the classroom.

As indicated, the profiles associated with population proportions describe the growth of a new technique or technology along these characteristics of the innovation itself (Rogers, 1995). In pedagogical reform, not only are personal factors important, but consideration of all aspects of the complex higher education system is essential (Henderson et al., 2015). In order to consider CRS adoption using the TALC, we must also consider the characteristics of the departmental, institutional, and extra-institutional level factors that impact incorporation of CRSs, including specific course types.

**Classroom response systems and the TALC**

For this study, we interpret results from a national survey of postsecondary chemistry faculty on the current state of CRSs. While an outcome like percent adoption is an efficient way of charting use, we argue that a broader understanding of the contexts in which CRSs are being implemented is more beneficial to understanding CRS adoption because contextual factors help to determine utility in practice, especially as the contexts align with the characteristics of the innovation described above (Woodbury & Gess-Newsome, 2002; Henderson et al. 2015; Moore, 2002).

Emenike and Holme (2012) noted that chemistry faculty perceive the utility of CRSs and yet do not adopt their use. MacArthur (2013) claimed that *Early Adopters* must use their status as change agents to encourage their colleagues to adopt. Emenike and Holme (2012) predicted that the common environment for the use of clickers is at doctoral granting institutions where large-enrollment introductory courses are found; the authors did not collect data to support this claim. Along with situational factors such as institution type, course level, and course size, it is vital, now six years later, to better understand the contexts in which faculty use CRSs.

In this report, CRS adoption is considered in light of faculty rank, public or private institutional control, course level, and number of students in the course. We conclude that faculty members have determined the context in which CRSs are most applicable in the chemistry classroom.

## Research Question

Our study is guided by the following question: In what contexts are U.S. faculty members utilizing classroom response systems (CRS) as a component of postsecondary chemistry education?

## Methodology

### Survey

A survey of postsecondary chemistry faculty was conducted via Qualtrics in February 2016. The University of South Florida's Institutional Review Board approved the study: #Pro00025183. Participants responded to survey items in relation to a self-selected undergraduate chemistry course taught in the past three years for which they had the most influence.

The survey included items about classroom practices and pedagogical techniques used in the respondent's articulated course, respondent demographics, departmental and institutional environment demographics, and respondent's beliefs about teaching and learning. Respondents were asked to report the frequency with which they utilized a CRS in their classroom.

Answers to three survey items (i.e., CRS use, who decided CRS use, and confidence using a CRS) were analyzed for the study reported herein; these items were considered by the five factors hypothesized to be limiting or enabling in regards to CRS use.

- *Total course enrollment size*. Course size has been found as a critical factor in understanding adoption of pedagogical and curricular reforms (Cheung, 2011).
- *Course level*. Factors surrounding the course itself are listed as the primary reasons for adopting or rejecting a certain new curriculum or pedagogy (Mack & Towns, 2016). We consider the level at which the articulated course was taught. Respondents had the choice between four course levels (descriptions were provided to the respondent as outlined below) congruent with the American Chemical Society's Committee on Professional Training (CPT) Guidelines for Undergraduate Bachelor's Degree Programs (CPT, 2015).
    1. *Introductory* – remedial or general chemistry
    2. *Foundation* – the first course in a subdisciplinary area; the course builds on the introductory coursework typically taught in general chemistry and has a general chemistry prerequisite
    3. *In-Depth* – the prerequisite is the foundation course in the subdisciplinary area(s)
    4. *Advanced or Special Topics*
- *Public or private control*. Control is used an indicator of institution size; private institutions are typically smaller than publicly controlled institutions (DeHaan, 2005). Institutional control information was obtained from the Integrated Postsecondary Education Data System (IPEDS).
- *Respondent title*. This information was determined from the respondent's departmental website. Title is used as a notation of tenure status and a proxy for number of years of experience in teaching at the undergraduate level; faculty experience has been previously found to be a indicator of choices made in the classroom (Davidovitch & Soen, 2006; Barlow & Antoniou, 2007).

**Sample**

The survey sample was selected from a database of chemistry faculty at postsecondary institutions awarding at least one bachelor's degree in chemistry in the past five years (Institutions $n$ = 1,128); this was done via an analysis of IPEDS data. University websites were referenced to compile the list of chemistry faculty. Faculty lists were unavailable for 37 institutions. In the end, 10,837 chemistry faculty members were identified at 1,091 institutions (i.e., the defined population). A stratified random sampling method was used to select 6,442 faculty from six strata defined by institutional control and highest chemistry degree awarded (Neyman, 1934). Sample size was determined for each stratum on the basis of a desired 95% confidence level, a 5% confidence interval, and the assumption of an aggressive 25% non-weighted response rate (see Table 3.1).

**Table 3.1.** Strata and sample definition.

| Strata | Institutional Control | Highest Chemistry Degree Awarded | Number of Institutions | Number of Chemistry Faculty | Sample Size |
|---|---|---|---|---|---|
| 1 | Public | Bachelors | 238 | 1,899 | 1,272 |
| 2 | Public | Masters | 83 | 1,080 | 1,075 |
| 3 | Public | Doctoral | 137 | 3,565 | 1,384 |
| 4 | Private | Bachelors | 551 | 2,836 | 1,335 |
| 5 | Private | Masters | 20 | 201 | 197 |
| 6 | Private | Doctoral | 62 | 1,256 | 1,179 |
| TOTALS | | | 1,091 | 10,837 | 6,442 |

In total, 1,282 chemistry faculty members responded to the survey; this represents a 33.3% unit response rate (see Table 3.2). This unit response rate falls below National Center for Education Statistics recommended guidelines; therefore, a nonresponse bias analysis was conducted. Upon comparison of the response rates on institutional characteristics included in the strata definition, potential for nonresponse bias was found when considering the unit response rates of faculty from institutions with differing highest chemistry degrees. Therefore, probability weights were used in all statistical analyses (Groves, 2006); a probability weight is the inverse of the ratio of number of respondents to the total number of chemistry faculty in each stratum (see *Final Weight* in Table 3.2). This is the first instance of probability weights in a survey research study published in this *Journal* to account for sampling error and response bias.

**Table 3.2**. Respondents and response rates.

| Strata | Sample Size | Initial Weight | Number of Responses | Unit Response Rate (%) | Final Weight |
|---|---|---|---|---|---|
| 1 | 1,272 | 1.49 | 328 | 38.5 | 5.79 |
| 2 | 1,075 | 1.01 | 226 | 21.1 | 4.78 |
| 3 | 1,384 | 2.58 | 153 | 28.5 | 23.30 |
| 4 | 1,335 | 2.12 | 403 | 64.1 | 7.04 |
| 5 | 197 | 1.02 | 46 | 23.8 | 4.37 |
| 6 | 1,179 | 1.07 | 126 | 11.4 | 9.97 |
| TOTALS | 6,442 | | 1,282 | 33.3 | |

### Statistical analysis

Descriptive and inferential statistics were conducted using Stata 13 with probability weights, stratification, and a finite population correction (StataCorp, 2013). Upper- and lower-bound 95% confidence intervals (CIs) are reported. Two-way cross tabulations with tests of independence were used to determine response differences. Odds ratios from logistic regressions with weighted survey data were used as measures of unstandardized effect sizes in instances of statistically significant $\chi^2$ results ($\alpha = 0.01$, Chen, Cohen, & Chen, 2010)

For continuous data, weighted-means analyses of variance (ANOVAs) were used to determine response differences from the survey items. Corresponding $\eta^2$ effect size values are reported in instances of significant $F$-statistics ($\alpha = 0.01$): $\eta^2 > 0.01$ = small; $\eta^2 > 0.06$ = medium; $\eta^2 > 0.14$ = large (Cohen, 1988; Smithson, 2001).

## Results

We address the three survey items by presenting overall descriptive statistics followed by inferential and effect-size statistics based on contextual variables. Respondents were asked to answer each item in reference to a specific course identified earlier in the survey; the title of that course was input into subsequent items in the survey where "[your course]" appears herein.

**Q1. The following methods can be used when teaching. Please indicate how often you used these methods when you last taught [your course].**

- *Asking clicker questions.*

Respondents were asked to note the frequency of asking clicker questions in their course (see Table 3.3). To best understand CRS use, a "Use" category was created including those who answered "Every Class Meeting," "Weekly," and "Several Times Per Semester."

**Table 3.3**. Weighted frequencies of asking clicker questions.

| Response Option | Total (%) | Lower Bound (95% ci) | Upper Bound (95% ci) |
|---|---|---|---|
| Every Class Meeting | 16.76 | 14.29 | 19.55 |
| Weekly | 4.31 | 3.12 | 5.94 |
| Several Times Per Semester | 1.87 | 1.21 | 2.88 |
| Rarely | 3.61 | 2.68 | 4.84 |
| Never | 73.45 | 70.38 | 76.31 |
| "Use" | 21.07 | 18.39 | 24.02 |
| "Nonuse" | 78.93 | 75.98 | 81.61 |

Compared with the 18.6% adoption reported by Emenike and Holme (2012), CRS use has only slightly increased over the past six years. We recognize that the survey item in our survey was somewhat different (i.e., Emenike and Holme asked more broadly about CRS use across all courses taught by the respondent, while our survey pertained to one specific course); however, we report that 21.07% of respondents note asking clicker questions as part of their regular classroom practices in the course for which they had the most influence. Although a direct comparison between the Emenike and Holme finding and our result is not entirely appropriate given the differing survey items, a small increase in CRS use has occurred.

**Q2. The last time you taught [your course], who were the primary decision makers for [classroom response system]?**

The results in Table 3.4 indicate that approximately three-quarters of respondents had the primary decision maker role regarding a classroom response system in their specified course. (CRS use was one of several aspects, including textbooks and curricular materials, for which the respondent was asked who were the primary decision makers). A sum total of 90.11% of respondents had some personal involvement in the decision.

**Table 3.4**. Weighted frequencies of primary decision makers for use of classroom response systems.

| Response Option | Total (%) | Lower Bound (95% ci) | Upper Bound (95% ci) |
|---|---|---|---|
| Yourself | 76.82 | 72.87 | 80.34 |
| Yourself and one other person | 4.55 | 3.1 | 6.63 |
| Yourself and several other people | 8.74 | 6.55 | 11.59 |
| Someone else or several other people | 9.89 | 7.44 | 13.03 |

There is a significant association between the primary decision makers for classroom response systems and use ($\chi^2(3) = 40.39$, design-based $F(2.97, 2251.90) = 8.71$, $p < .001$). A weighted logistic regression analysis predicting CRS use by primary decision maker revealed the following:

- Faculty who responded "Someone else or several other people" were 2.62 times ($p = .003$) more likely to report CRS use compared with faculty who responded "Yourself."
- Faculty who responded "Yourself and several other people" were 3.85 times ($p < .001$) more likely to report CRS use compared with faculty who responded "Yourself."

(Note: Nonsignificant odds ratios are not reported for logistic regression analyses throughout this article.) While these results suggest that when groups of individuals make decisions about classroom response systems, CRSs are more likely to be implemented, we must analyze other contextual factors prior to interpretation given that large lecture courses are most likely coordinated and taught by a group of faculty.

**Q3. How confident are you in [using student response systems (e.g., clickers, TopHat)]?**

Per the TALC framework, the *Early Majority* has no use for a technology for which they are not comfortable, whereas *Early Adopters* often take on a new technology in a "learn-as-you-go" fashion. Table 3.5 shows a relatively even distribution of confidence in using classroom response systems. The majority of respondents are at minimum "moderately confident" in their ability to use CRS (total of 55.39%).

**Table 3.5**. Weighted frequencies of confidence in using student response systems.

| Response Option | Total (%) | Lower Bound (95% ci) | Upper Bound (95% ci) |
|---|---|---|---|
| Completely Confident | 18.1 | 15.65 | 20.84 |
| Very Confident | 17.19 | 14.88 | 19.77 |
| Moderately Confident | 20.1 | 17.65 | 22.79 |
| Somewhat Confident | 16.14 | 14.03 | 18.51 |
| Not at all Confident | 28.47 | 25.56 | 31.57 |

There is a significant association between confidence in using a CRS and CRS use ($\chi^2(1) = 246.71$, design-based $F(1, 1276) = 156.04$, $p < .001$). A weighted logistic regression analysis predicting CRS use by confidence in using CRS revealed the following:

- Faculty who were at minimum "Moderately confident" in their ability to use a CRS were 7.16 times ($p < .001$) more likely to use a CRS than faculty who were "Somewhat confident" and "Not at all confident."

This finding supports the claim that confidence in a technology is directly associated with adoption of that technology.

**CRS adoption by context**

***Course size (total number of students).*** Our central hypothesis is that CRS use is more prevalent in large lecture courses. We begin our targeted analysis of this hypothesis by considering the use of CRSs (Q1) and the total number of students enrolled in the respondent's course. Faculty respondents who regularly use clickers report an average total course size of 423 students (330 to 514 students, lower and upper bound), whereas faculty respondents who do not regularly use clickers report total course sizes of 144 students (119 to 168 students, lower and upper bound). A weighted-means ANOVA yielded a significant result with a medium effect size ($\eta^2 = 0.061$, $F(1, 1280) = 83.33$, $p < .001$).

Similarly, there is a statistically significant difference in the total number of students enrolled in the courses taught by respondents and who was the primary decision maker of a CRS use (Q2) as determined by a weighted-means ANOVA with a medium effect size ($\eta^2 = 0.072$, $F(3, 760) = 19.62$, $p < .001$). Additionally, larger total course sizes are associated with higher levels of confidence in using CRSs (Q3): a weighted-means ANOVA yielded a significant result with a small effect size ($\eta^2 = 0.015$, $F(4, 1277) = 4.77$, $p < .001$).

These results are the logical extension of the association between CRS use, primary decision maker of CRS use, and confidence using CRS. Larger total course enrollments are associated with the primary decision-making role being shared with or entirely made by others, and use of a CRS is associated with confidence using a CRS.

***Course level.*** Significant differences are observed between CRS use and course level, i.e., introductory or remedial, foundation, in-depth, or advanced ($\chi^2(12) = 100.21$, design-based $F(11.30, 14,1415.29) = 5.89$, $p < .001$). Faculty teaching introductory courses are 4.16 times ($p < .001$) more likely than faculty teaching in-depth courses, 3.81 times ($p < .01$) more likely than faculty teaching advanced courses, and 3.13 times ($p < .001$) more like than faculty teaching foundation courses to report regular clicker use in the course for which they had the most influence.

There is a statistically significant difference in primary decision-making role (Q2) and course level ($\chi^2(9) = 36.34$, design-based $F(8.08, 6124.74) = 3.06$ $p < .01$). A weighted logistic regression ($F(3,756) = 4.61$, $p < .01$) yielded the following:

- Faculty teaching in-depth courses are 1.47 times ($p < .01$) more likely than faculty teaching introductory courses to report "Yourself" versus all other options combined.
- Faculty teaching foundation courses are 1.41 times ($p < .001$) more likely than faculty teaching introductory courses to report "Yourself" versus all other options combined.

There is a statistically significant difference between the level of the course for which the respondent had the most influence and confidence in using a CRS in the classroom (Q3) ($\chi^2(12) = 61.50$, design-based $F(11.71, 14942.20) = 3.73$, $p < .001$).

These results suggest a connection to course level; however, it is important to recall the significant association between course level and course size. A weighted-means ANOVA between course level and course size yields a significant result with a large effect size ($\eta^2 = 0.171$, $F(3, 1278) = 87.69$, $p < .001$). While there may be an independent association between CRS use and course level, we recognize that course level and course size are confounding.

***Public versus private control.*** There is a statistically significant difference between CRS use (Q1) and institutional control, i.e., public or private ($\chi^2(4) = 28.53$, design-based $F(3.97, 5071.72) = 5.84$, $p < .001$). A weighted logistic regression analysis predicting clicker use by institutional control revealed that faculty at public institutions are 1.76 times ($p < .001$) more likely to report regular clicker use in the course for which they had the most influence compared with faculty at private institutions, ($F(1,1276) = 12.91$, $p < .001$).

Institutional control had a statistically significant difference with primary decision maker of CRS use (Q2) ($\chi^2(3) = 13.69$, design-based $F(2.96, 2241.61) = 3.67$, $p < .05$), and the following results were obtained:

- Faculty at public institutions are 2.85 times ($p < .05$) more likely to report "Yourself and several other people" than "Yourself and one other person" compared with faculty at private institutions.

- Faculty at public institutions are 3.94 times ($p < .01$) more likely to report "Someone else or several other people" than "Yourself and one other person" compared with faculty at private institutions.
- Faculty at public institutions are 2.17 times ($p < .01$) more likely to report "someone else or several other people" than "Yourself" compared to faculty at private institutions.

We recognize that in smaller departments, which are more prevalent at private institutions, there might not be more than one individual to serve as the primary decision maker for a course; therefore, these results are confounded by the small number of faculty at private institutions.

There is no statistically significant difference in confidence level with CRS use (Q3) and institutional control ($\chi^2(4) = 8.87$, design-based $F(4.00, 5100.53) = 1.81$, $p > .05$).

A weighted means ANOVA between institutional control and course size yielded a significant result with a small effect size ($\eta^2 = 0.035$, $F(1, 1280) = 47.07$, $p < .001$). Given the small effect size between institutional control and course size, we conclude that considering CRS use by institutional control provides an added understanding of the contexts in which CRS are adopted.

***Faculty rank (proxy for teaching experience).*** Faculty rank has been used as a proxy for teaching experience (Davidovitch & Soen, 2006; Barlow & Antoniou, 2007). When CRS use (Q1), primary decision maker for CRS use (Q2), and confidence in using CRS (Q3) are considered by faculty rank, no statistical differences between groups are observed.

- Q1: $\chi^2(8) = 16.74$, design-based $F(7.46, 9520.62) = 1.64$, $p > .05$
- Q2: $\chi^2(6) = 4.01$, design-based $F(5.90, 4475.90) = 0.45$, $p > .05$
- Q3: $\chi^2(8) = 20.87$ design-based $F(7.91, 10096.82) = 1.95$, $p > .05$

While faculty rank has been shown as an important personal context in other work, we fail to find evidence to support the association between faculty rank and our study measures.

***Cross tabulation of course size, course level, institutional control with CRS use.*** Our analyses conclude that course size, course level, and institutional control are separately associated with CRS use.

We are thus interested in knowing the level of CRS use in these contexts combined. Therefore, each context was divided into binary categories: (A) course size was considered “Large” for enrollments larger than the median course size (i.e., 55 students) and “Small” for enrollments smaller than the median course size; (B) course level was considered “Lower” for introductory and foundation level courses and “Upper” for in-depth and advanced level courses; (C) institutional control was either “Public” or “Private.” Eight contexts resulted from these binary combinations (see Table 3.6). The percent CRS is reported for each of these contexts.

**Table 3.6**. Cross tabulation of course size, course level, institutional control with CRS use.

| Size | Level | Control | *n* | % CRS Use | Lower Bound (95% ci) | Upper Bound (95% ci) |
|---|---|---|---|---|---|---|
| Large | Lower | Public | 351 | 36.97 | 30.71 | 43.70 |
| Large | Upper | Public | 59 | 28.58 | 16.73 | 44.35 |
| Large | Lower | Private | 210 | 24.32 | 18.93 | 30.67 |
| Small | Lower | Private | 198 | 15.31 | 11.01 | 20.89 |
| Small | Lower | Public | 148 | 11.21 | 6.20 | 20.89 |
| Large | Upper | Private | 21 | 10.59 | 2.65 | 34.05 |
| Small | Upper | Private | 146 | 10.34 | 6.46 | 16.14 |
| Small | Upper | Public | 149 | 9.55 | 5.02 | 17.41 |

Our results support the claim that CRS use is more prevalent in large courses taught at the introductory and foundation levels. However, CRS use is not exclusive to this context, we observe that CRSs are being used in all course sizes, course levels, and institutional types (We causation against over interpretation of CRS use in large upper-level courses given the small *n*-values and subsequent large confidence intervals.)

**Discussion**

**In what postsecondary chemistry education contexts are U.S. faculty utilizing classroom response systems?**

The utility of CRSs in postsecondary chemistry education can be understood through the contexts in which CRSs are adopted. Despite promotion and research, CRS use appears to have settled in to a niche. In summary, CRSs are more prevalently used at public institutions, in classrooms with a large (>55) number of students, and in introductory/foundation courses. Faculty using CRSs report personal involvement in the

decision-making process for determining CRS use; these respondents also feel confident in their ability to use a CRS. We found no association between CRS use and faculty title, a proxy for teaching experience (Davidovitch & Soen, 2006; Barlow & Antoniou, 2007). This implies that faculty are making a utility-oriented choice when adopting a CRS as an instructional practice; the contextual factors studied help us understand the courses and classrooms in which faculty find CRSs most useful. The importance of these factors is clear through the lens of the theoretical framework and the characteristics of the innovation. In trying to make sense of the growth and use of CRSs as a proxy for understanding the growth of reform movements in the field of chemical education, our study gives a broader insight into a spectrum of the factors necessary to consider adoption of instructional technology.

The framework used to guide this study, Rogers' (1995) technology adoption life cycle (TALC), considers population percentages along with psychographic factors. The population percentage required to consider a technology in the *Early Adopters* stage is at minimum 15.8%; in order to fill the *Early Majority*, a total of 50% of the population must use a new technology. Emenike and Holme (2012) found that 18.6% of respondents indicated using clickers in the classroom, suggesting that only *Early Adopters* are using the technology; we found similar results with a 21.07% adoption rate. When considering the population of chemistry faculty in the U.S., the adoption percentage has not grown to any considerable degree. CRS adoption is in a position that could be considered within the "chasm;" in other words, some members of the

*Early Majority* has taken on the new technology, but there are not enough individuals to deem the technology as being adopted across the population.

Despite the low incidence of CRS use, associated personality profiles from the TALC tell a more nuanced story. The TALC defines *Early Adopters* as enthusiastic to take on the challenge of adopting a new technology. The *Early Majority*, on the other hand, will consider using a new technology only when they have seen sufficient evidence for its success. Because CRSs have been shown to be an effective tool in enhancing student learning, there is no reason to expect that only the *Early Adopters* are using CRSs in the classroom. The results from Q2 indicate that 90.11% of faculty who are personally involved in the selection of CRSs teach at private institutions, teach courses with fewer students, and teach courses at the foundation or in-depth level. Faculty using CRSs report being on a team or outside of the selection of such technology; these individuals are at public institutions, teaching courses with large numbers of students and courses at the introductory level. The ability to choose which, if any, CRS system to use in the classroom is a key power invested in *Early Adopters*, but those who use CRSs are not those for whom this power is singularly invested. Since those using CRSs are not necessarily in control of their CRS system, it is clear that the *Early Adopters* are not the only group that uses CRSs in their classrooms.

Those using CRSs in the classrooms of large-enrollment and introductory/foundation-level courses teach in an environment in which the technology has been shown to work. While the TALC is useful in describing the growth of popular technology, Rogers indicates in his work that the utility of the technology must be considered when interpreting such growth. The population in which a new technology spreads is a population for which the technology is always applicable; the results of this study indicate that postsecondary chemistry faculty have determined that the applicability of CRSs is not widespread but instead is limited to specific situations. Understanding the rich nature of CRS use requires a framework with a multifaceted approach. Using Towns' report on the TALC as a tool to understand the growth of CRSs (Towns, 2010), Emenike and Holme (2012) called for more than just percentage adoption metrics, specifically research on course size and CRS use and the association between institution type and CRS use.

We found that CRSs are less likely to be used in classrooms with small numbers of students. In addition, we found that institutional control (i.e., public or private) and course level are associated with CRS use more strongly than confidence in using the technology. These results indicate that the situational context is more important than the individual instructor in the adoption of a CRS.

There is an array of application techniques for adoption of CRSs; however, we propose that limited adoption is due to an applied utility of the technology for large introductory and foundation level courses. We propose that faculty members teaching courses in this niche have chosen to adopt a CRS because the context is appropriate (Woelk, 2008). Our findings are helpful for considering how developers and researchers can disseminate and promote CRSs to appeal to this niche market. In education, active learning techniques such as CRSs are not viewed as applicable in all situations. Because our results indicate that situational and personal contextual factors are crucial to the use and nonuse of a particular technology, it is reasonable to expect that the same multifaceted problems of adoption apply in to other pedagogical decisions.

**Limitations**

Our analysis is limited by the response rate of our sample. Unequal response rates among the six strata required the use of adjusted probability weights in our statistical procedures, which introduced larger confidence intervals. Our sample size, then, limited our ability to consider an association between CRS use and collective contextual profiles; several *n*-values in Table 3.6 are extremely low. The strength of stratified sampling, weighted descriptive and inferential statistical analyses, and interpretation of results through a theoretical framework outweigh such a limitation.

Another limitation is our self-reporting protocol for collecting data about instructional practices. Self-reported data do not capture instructional practices as accurately as resource-intensive observational protocols. It would be too costly to conduct a study of the magnitude reported herein using such protocols.

Observation protocols, such as COPUS (Smith et al., 2013), could be used in more limited ways to validate self-reported classroom practice data, e.g., clicker use for a subset of the sample.

The personal and situational contexts analyzed in this report provide us with information regarding CRS use nationally. However, we hypothesize other factors that might additionally show a difference between faculty and institutions in which CRSs are or are not adopted. The situational context analyzed in this report could have been more thorough with consideration of the subdisciplines of more advanced courses, as this may impact CRS use. In addition, asking respondents about an adopted or recommended CRS system in their institution would have provided another dimension to consider CRS adoption.

**Conclusion**

While advocates of CRS use, frequent CRS users, and researchers of CRS use collectively purport the applicability of CRSs in *all* classroom situations, the results of this study suggest a niche in which the use of CRSs has flourished and outlines a contextual profile for that niche. In light of this information, we claim that the chasm has not yet been crossed, although our data do suggest that CRS use in large courses at an introductory or foundation level is closer to the 50% adoption threshold than other contexts.

Our results are the first instance in this *Journal* of a stratified sampling strategy for a survey research study; additionally, our results are the first instance in this *Journal* of the use of probability weights for determining confidence intervals to account for sampling error and response bias. The adoption of these rigorous survey research methods have enabled us to make more certain claims about the population of chemistry faculty then previous survey research studies on the instances of CRSs and instructional practice in postsecondary chemistry education.

Frameworks that consider adoption of education reforms must include contextual factors including number of students, course level, and institutional control. Personal factors, such as faculty rank and confidence in using a CRS, were not found to be as important as contextual factors in this study; however, we do not suggest that such factors are not important for understanding the adoption of other technological tools designed to assist instruction. Our findings suggest that chemistry faculty do indeed have the opportunity to choose to use a CRS in the classroom; however, in the end, 78.93% choose not to do so; we argue that this is because CRS technology has been deemed useful in a certain type of chemistry classroom. We propose that further research be conducted to understand the rationales put forward by postsecondary chemistry faculty for why they ultimately decided to use or not use a CRS and, more broadly, to adopt or not adopt evidence-based instructional pedagogies.

**References**

Addison, S.; Wright, A.; Milner, R. (2009), Using clickers to improve student engagement and performance in an ntroductory biochemistry class. *Biochemistry and Molecular Biology Education, 37* (2), 84-91.

Asirvatham, M. R. (2005), IR Clickers and ConcepTests: Engaging Students in the Classroom. Winter ConfChem: Trends and New Ideas in Chemical Education. http://confchem.ccce.divched.org/2005WinterConfChemP5 (accessed Feb 2017).

Asirvatham, M. R. (2010), *Clickers in action: Increasing student participation in general chemistry*. Norton: New York.

Barlow, J.; Antoniou, M. (2007), Room for improvement: The experiences of new lecturers in higher education. *Innovations in Educaction and Teaching International, 44* (1), 67-77.

Bruck, A. D.; Towns, M. H. (2009), Analysis of classroom response system questions via four lenses in a general chemistry course. *Chemistry Education Research and Practice***,** *10*, 291-295.

Bunce, D. M.; Vandenplas, J. R.; Havanki, K. L. (2006), Comparing the effectiveness on student achievement of a student response system versus online WebCT quizzes. *Journal of Chemical Education***,** *83*(3), 488-493.

Chen, H.; Cohen, P.; Chen, S. (2010), How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Communication in Statistics-Simulation and Computation, 39*, 860-864.

Chen, Z. Z.; Stelzer, T.; Gladding, G. (2010), Using multimedia modules to better prepare students in an introductory physics lecture. *Physics Education Research***,** *6*, 101018-010108-5.

Cheung, D. (2011), Instructor beliefs about implementing guided-inquiry laboratory experiments for secondary school chemistry. *Journal of Chemical Education 88*, 1462-1468.

Childs, P. E. (2008), Improving chemical education: Turning research into effective practice. *Chemistry Education Research and Practice 10*, 189-203.

Cohen, J. (1988), *Statistical power analysis for the behavioral sciences*. 2nd ed.; Lawrence Earlbaum Associates: Hillsdale, NJ.

Committee on Professional Training (2015), *Undergraduate professional education in chemistry: ACS guidelines and evaluation procedures for bachelor's degree programs*. American Chemical Society: Washington, D.C.

Crouch, C. H.; Mazur, E. (2001), Peer instruction: Ten years of experience and results. *American Journal of Physics, 69*, 970-977.

Davidovich, N.; Soen, D. (2006), Using students' assessments to improve instructors' quality of teaching. *Journal of Further & Higher Education 30*, 351-376.

DeHaann, R. L. (2005), The impending revolution in undergraduate science education. *Jouranl of Science and Educational Technology, 14* (2), 253-269.
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.
Groves, R. M. (2006), Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly, 70* (5), 646-675.
Hall, R. H.; Collier, H. L.; Thomas, M. L.; Hilgers, M. G. (2005), In *A student response system for Increasing engagement, motivation, and learning in high enrollment lectures*, Eleventh Americas Conference on Information Systems, Omaha, NE, August 11-14; Omaha, NE.
Henderson, C.; Cole, R.; Froyd, J.; Friedrichsen, D.; Khatri, R.; Stanford, C. (2015), *Designing Educational innovations for sustained adoption: A how-to guide for education developers who want to increase the impact of their work*. Increase the Impact: Kalamazoo, MI.
Henderson, C.; Dancy, M.; Niewiadomska-Bugai, M. (2012), Use of research-based instructional strategies in introductory physics: Where do faculty leave the innovation-decision process? *Physical Review: Physics Education Research, 8*, 020104- 020119.
Khatri, R.; Henderson, C.; Cole, R.; Froyd, J.; Friedrichsen, D.; Stanford, C. (2016), Designing for sustained adoption: A model of development education innovations for successful propagation. *Physical Review: Physics Education Research, 12* (010112).
Koenig, K. (2010), Building acceptance for pedagogical reform through wide-scale implementation of clickers. *Journal of College Science Teaching, 39*, 46-50.
Lee, A. W. M.; Ng, J. K. Y.; Wong, E. Y. W.; Tan, A.; Lay, A. K. Y.; Lai, S. F. Y. (2013), Lecture rule no. 1: Cell phone ON, Please! A low-cost personal response system for learning and teaching. *Journal of Chemical Education, 90* (4), 388-389.
MacArthur, J. (2013), How will classroom response systems "cross the chasm"? *Journal of Chemical Education, 90* (3), 273-275.
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*, 187-195.
MacArthur, J. R.; Jones, L. L. (2013), Self-assembled student interactions in undergraudate general chemistry clicker classrooms. *Journal of Chemical Education, 90* (12), 1586-1589.
Mack, M. R.; Towns, M. H. (2016), Faculty beliefs about the purposes for teaching undergraduate physical chemistry courses. *Chemistry Education Research and Practice, 17*, 80-99.
Moore, G. A. (2002), *Crossing the chasm*. Collins: New York, NY.
Neyman, J. (1934), On the two different asepcts of the representative methods: The method of stratified sampling and the method of purposive selection. *Journal of the Royal Statistical Society A, 97*, 558-606.
Phillis, R.; Brewer, S.; Hoogendyk, T.; Goodwin, S.; Connor, E. University of Massachusetts Biology 100 Course Redesign Using Technology to Facilitate Active Learning in the Large Lecture Hall. http://bcrc.bio.umass.edu/pew/phillis.pdf (accessed Feb 2017).
Poulis, J.; Massen, C.; Robens, E.; Gilbert, M. (1998), Physics lecturing with audience paced feedback. *American Journal of Physics, 66*, 439-441.
Rogers, E. M. (1995), *Diffusion of Innovation*. 4th ed.; Free Press: New York, NY.
Seery, M. K. (2013), Harnessing technology in chemistry education. *New Directions for Institutional Research, 9* (1), 77-86.
Sevian, H.; Robinson, W. E. (2011), Clickers promote learning in all kinds of classes: Small and large, graduate and undergraduate. *Journal of College Science Teaching, 40* (3), 14-18.
Shea, K. M. (2016), Beyond clickers: Next generation classroom response systems for organic chemistry. *Journal of Chemical Education, 93* (5), 971-974.
Smith, M. K.; Jones, F. H. M.; Gilbert, S. L.; Wieman, C. E. (2013), The Classroom Observation Protocol

for Undergraduate STEM (COPUS): A new instrument to characterize university STEM classroom practices. *CBE - Life Sciences Education, 12*, 618-627.
Smith, M. K.; Wood, W. B.; Adams, W. K.; Wieman, C.; Knight, J. K.; Guild, N.; Su, T. T. (2009), Why peer discussion improves student performance on in-class concept questions. *Science*, *323* (5910), 122-124.
Smithson, M. (2001), Correct confidence intervals for various regression effect sizes and parameters: The importance of noncentral distributions in computing intervals. *Educational and Psychological Measurement, 61*, 605-632.
Stains, M.; Pilarz, M.; Charkraverty, D. (2015), Short- and long-term impacts of the Cottrell Scholars Collaborative New Faculty Workshop. *Journal of Chemical Education, 92*, 1466-1476.
StataCorp (2013), *Stata 13 Survey Data Reference Manual*. Stata Press: College Station, TX.
Towns, M. (2010), Crossing the chasm with classroom response systems. *Journal of Chemical Education 87* (12), 1317-1319.
Vickrey, T.; Rosploch, K.; Rahmanian, R.; Pilarz, M.; Stains, M. (2015), Research-based implementation of peer instruction: A literature review. *CBE - Life Science Education 14*, 1-11.
Woelk, K. (2008), Optimzing the use of personal response devices (clickers) in large-enrollment introductory courses. *Journal of Chemical Education, 85* (10), 1400-1405.
Woodbury, S.; Gess-Newsome, J. (2002), Overcoming the paradox of change without difference: A model of change in the arena of fundamental school reform. *Educational Policy, 16*, 763-782.

CHAPTER FOUR:

TESTING A RECIPROCAL CAUSATION MODEL BETWEEN ANXIETY, ENJOYMENT, AND PERFORMANCE IN ORGANIC CHEMISTRY

## Note to Reader

This is the author's accepted manuscript of an article published as the version of record in *Educational Psychology* © 07 Mar 2018 - https://www.tandfonline.com/10.1080/01443410.2018.1447649. Permissions information can be found in Appendix B, and the Journal can be found at http://www.tandfonline.com. This work was also published with the contributions of co-authors. Xiaoying Xu and Sachel M. Villafañe both contributed as postdoctoral research associates in the collection and initial data analysis, although all analysis reported in the manuscript is my own. Jeffrey R. Raker provided administrative support to the research.

## Introduction

Learning is an emotional experience (Mandler, 1990; Pekrun, 2000, 2006). Achievement emotions hold predictive ability towards final scores and performance outcome measures (Daniels et al., 2009; Huang, 2011; Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011; Pekrun, Goetz, Titz, & Perry, 2002; Pekrun, J., & Maier, 2009; Putwain, Larkin, & Sander, 2013). This has been demonstrated in the field of chemistry at the undergraduate level (Nieswandt, 2007) and has been well accepted in educational psychology (Richardson, Abraham, & Bond, 2012). Studies on affect in chemistry courses and the effect of affect on performance can be found in the literature (e.g., Teo, Goh, & Yeo, 2014); however, achievement emotions have been understudied in the chemistry education context. This study evaluates a reciprocal causation relationship between enjoyment and anxiety, two achievement emotions and performance in a postsecondary organic chemistry course.

While achievement emotions as a set of constructs have not received a great deal of attention in chemistry education research, anxiety towards chemistry has been well studied (Bowen, 1999; Eddy, 2000; Kurbanoglu & Akim, 2010; McCarthy & Widanski, 2009). Expanding this work to other achievement emotions is essential to building a well-rounded understanding of the student experience (Abendroth & Friedman, 1983). Organic chemistry students are often distressed in the environment of the course (Grove & Bretz, 2010), which holds a large proportion of weight in applications for graduate and professional school (Muller, 2013). High assigned value, as demonstrated by motivation to learn organic chemistry, comes from both intrinsic and extrinsic sources; extrinsic motivation, like scoring high on professional school exams, has been found to be negatively correlated with performance (Lynch & Trujillo, 2011). Understanding the relationship between affect and performance in this context will enable us to better understand learning in high-stakes courses.

Predictors of success in postsecondary organic chemistry courses include prior math and chemistry performance as well as non-cognitive attributes such as goals and attitudes (Steiner & Sullivan, 1984; Tien, Roth, & Kampmeier, 2022; Zoller & Pushkin, 2007). However, previous mathematics experience and performance is less related to organic chemistry achievement than general chemistry achievement because organic chemistry approaches problems from a structure-function epistemology (Grove & Bretz, 2010). Faculty have reported students entering organic chemistry courses with unsatisfactory prior knowledge despite acceptable performance in prerequisite courses (Duis, 2011). One challenge to learning organic chemistry is that problems in organic chemistry are not solved through algorithmic thinking as they are in lower level chemistry courses, but through higher order cognitive skills and in-depth content knowledge (Anderson & Bodner, 2008; Grove & Bretz, 2010; Grove, Cooper, & Cox, 2012). Similarly, content such as hydrogen bonding is frequently defined differently in organic chemistry than in general chemistry

(Henderlierer, Smart, Anderson, & Elian, 2001). Ultimately, the organic chemistry course challenges students to perform despite success in previous math and science courses (Anderson & Bodner, 2008).

**Control-Value Theory**

This study, focused on the analysis of the impact of both anxiety and enjoyment in organic chemistry, is framed by the control-value theory of achievement emotions (CVT; Pekrun, 2006). CVT posits the influence of evaluations of control and value on the affective domain and the direct impact on affect on performance (Pekrun, 2000, 2006). CVT outlines nine achievement emotions based on prevalence and perceived importance in groups of undergraduate students (see Table 4.1; Pekrun et al., 2002). An achievement emotion is a set of domain-specific processes experienced by a student in an academic environment (Goetz, Pekrun, Hall, & Haag, 2006). Achievement emotions are defined by activation and valence, as influenced by a student's control over the content and value placed on the achievement activity. The 216-item, 24-subscale Achievement Emotions Questionnaire (AEQ; Pekrun, 2000, 2006; Pekrun et al., 2002) serves as the key assessment tool for measuring achievement emotions.

**Table 4.1.** Nine emotions (including valence and activation) of the control-value theory of achievement emotions.

| | Positive Valence | Negative Valence |
|---|---|---|
| Activating | *Enjoyment*, Hope, Pride | *Anxiety*, Anger, Shame |
| Deactivating | Relief | Boredom, Hopelessness |

*Note*. Italic emotions are the focus of the study reported in this manuscript.

Control and value are antecedents to the achievement emotions (Goetz et al., 2006). Control can be described as a student's level of understanding of the content and is evaluated based on prior knowledge along with self- and instructor-appraisal (Pekrun, 2006). Value refers to the value placed on the content or activity by a student; domain specificity requires that this be separated into outcomes and activities (Goetz et al., 2006). The use of CVT as the theoretical framework for this study is justified by demonstrations of the importance of control and value in our context from the literature. In a study of organic chemistry students, Lynch and Trujillo (2011) found a positive correlation between academic performance and student autonomy, or control over their learning. Similarly, research has demonstrated that organic chemistry students highly value the outcomes and activities because the importance of the subject for their career aspirations, made evident through content requirements on entrance exams for graduate and professional school (Grove, Hershberger, & Bretz, 2008; Muller, 2013). It has been noted that high value placed on organic chemistry content is associated with increased student attitudes (Dwyer & Childs, 2014).

In this study, two of the achievement emotions highlighted in the CVT are studied: anxiety and enjoyment. Anxiety was selected because it is frequently studied in the academic context, allowing comparisons to other research. Enjoyment falls on the opposite side of the valence scale in CVT (i.e. positive valence) and is also considered an activating emotion; therefore, we are interested in its differential impact on achievement. The potential for measuring all nine of the achievement emotions is reduced in this study when considering power; many studies of affect focus on a small number of the achievement emotions because sample size limits the number of factors included in the statistical analyses (Ahmed, van der Werf, Kuyper, & Minnaert, 2013).

**Anxiety**

Anxiety is a negative, activating emotion. Anxiety experienced in the academic context is known to be domain-specific and experienced differently in different courses as well as different contexts, such as studying or during class (Pekrun, 2006). Negative valence indicates that high anxiety levels will result in decreased performance. The activation designation is important to learning because activation can lead to decreased attention to the material at hand (Pekrun et al., 2009). A student experiencing anxiety is distracted by experiencing apprehension regarding success; the student is motivated to escape the situation and suffers through worry and tension as their heart rate and body temperature rise, all reducing attention (Spielberger, 1972). Reduced attention provides the theoretical link to reduced performance; this theoretical understanding is also well understood empirically, which contributes to its inclusion in this study.

Anxiety was selected for analysis in this study because this emotion is explored with increasing frequency in the context of chemistry education. Most studies indicate a negative relationship between anxiety and performance (Eddy, 2000). The Attitude towards the Subject of Chemistry Instrument (ASCI; Bauer, 2008), a chemistry-specific measure of affect, has an anxiety subscale. In general chemistry, students in a low affective group, as defined by measures including the ASCI(v2) (Xu & Lewis, 2011), indicated not understanding their notes, rote transcribing during class time, heavy reliance on TAs and peers for information and lower performance (Chan & Bauer, 2016). Anxiety has been found to have a negative relationship to both self-efficacy and metacognitive self-regulation in a study of introductory chemistry students (Aydin, Uzuntiryaki, & Demirdogen, 2011). In the organic chemistry classroom context, a negative relationship between test anxiety and performance has been found (Lynch & Trujillo, 2011), along with perceived competence and anxiety to be opposite and equal predictors of success on performance (Black & Deci, 2000). The results of this study contribute to the body of literature on affect in chemistry education by providing a measurement of anxiety in the organic undergraduate course.

**Enjoyment**

Enjoyment is a positive, activating emotion. Positive valence indicates that a student experiencing enjoyment will have increased performance. Activation indicates that the student will engage with the content and feel motivated to maintain attention (Pekrun, 2006). The activation of the cognitive domain implies an increase in attention and the use of learning strategies; enjoyment has been found to be a positive predictor of the use of learning strategies such as elaboration and metacognition in undergraduates as well as elementary students (Artino & Jones, 2012; Pekrun, 2000). Increased attention, then, is associated with an increased potential for achievement. Positive emotions are frequently ignored in research on achievement emotions; however, it has been argued that the impact of positive emotions should be considered and enhanced to maximise educational effectiveness (Fredrickson, 2001; Linnenbrink-Garcia &

Pekrun, 2011). Because enjoyment is associated with increased achievement through the activation mechanism, which contributes to decreased achievement when anxiety is experienced, it represents an opposite impact which should be measured to achieve the goals of attitudinal research in education.

Enjoyment was selected for inclusion in this study based on empirical evidence for its impact on achievement. Despite few studies in chemistry education, the support for enjoyment is established in other settings. A study of elementary students found that enjoyment was associated with increased use of deep and metacognitive learning strategies as well as increased performance (Ahmed et al., 2013). In a study of university mathematics students, academic enjoyment served as a vehicle through which self-regulation and achievement interact; students with higher enjoyment had increased regulation and achievement, while students with low enjoyment had a negative interaction between regulation and achievement (Villavicencio & Bernardo, 2013). Enjoyment for medical students contributed to higher course grades and board examination scores (Artino, La Rochelle, & Durning, 2010). Enjoyment along with the closely related construct interest have been found to have a positive relationship with performance and conceptual understanding in postsecondary chemistry, mathematics and science courses (Ainley & Ainley, 2011; Goetz, Frenzel, Hall, & Pekrun, 2007; Nieswandt, 2007; Singh, Granville, & Dika, 2002). Because enjoyment has been understudied in the chemistry education context, this study will utilise a measure of enjoyment to analyse its reciprocal relationship with achievement. Our study will provide a framework for conducting future studies to explore the relationship of the positive emotions with achievement. Ultimately, this work will deepen our understanding of the effect of affect on achievement in chemistry classrooms.

### Reciprocal Causation

Studying anxiety and enjoyment through the lens of the control-value theory (CVT) of achievement emotions enables this study to explore a reciprocal causation model. Reciprocal causation follows from CVT in that antecedents and outcomes are linked through the continuous reappraisal of control and value (Pekrun, Hall, Goetz, & Perry, 2014). The dynamic nature of affect in educational attainment is reflected in studies which have demonstrated reciprocal causation over time; affect has been shown to fluctuate at moments where re-evaluation is possible, i.e. after performance measures (Pekrun et al., 2009). Reciprocal causation relationships have been shown between achievement goals and learning-related emotions in college students (Putwain et al., 2013), academic performance and boredom in a first-year psychology course (Pekrun et al., 2014) and self-efficacy and organic chemistry performance (Villafañe, Xu, & Raker, 2016).

Reciprocal causation mechanisms do not indicate a 'causal' relationship in the colloquial use of the term. The relationship between scores on performance outcome measures such as exam score and measures of affect or motivation is correlative and complex. The longitudinal process by which data are collected and analysed lends credence to theorised reciprocal relationships (Anderson & Evans, 1974). The pre/post-application of affective measures along with performance fail to capture the development of emotional learning, which has been demonstrated to reflect changes though an academic term, indicating that longitudinal reciprocal models are an improvement upon pre/post-educational research designs (Nieswandt, 2007). Therefore, the study described herein seeks to explore the relationship of anxiety and enjoyment to achievement through the lens of a methodologically rigorous reciprocal causation mechanism.

**The Current Study**

The impacts of anxiety and enjoyment are separately understood in the realm of education, although not specifically in chemistry. This study seeks to understand their interrelationships and continue exploration in the affective domain of organic chemistry, and is guided by two research questions:

1. What are the interrelationships between anxiety, enjoyment and examination performance?
2. What are the strength and valence of these relationships?

## Methods

### Participants and procedure

Participants were students (*N* = 907) enrolled in four sections of the first semester of a year-long postsecondary organic chemistry course sequence. This research was conducted at the University of South Florida (USF), a large, public, research-intensive institution in the Southeast United States. The student population was 52% female, 46% white, 34% Hispanic, 18% Asian and 10% black, with an average raw SAT mathematics score of 589.

Students in the course completed the study measures via Qualtrics (online). Questionnaires were made available three days prior to the scheduled course examinations, and two reminder emails were sent to students to complete the questionnaires within the 72 h before the examination. Questionnaires were closed for submission upon the beginning of each examination, upon which time data were collected from Qualtrics for analysis. Students were offered extra credit for the completion of the questionnaires. Data were collected in accordance with USF Institutional Review Board application Pro#00017861 approved on 18 June 2014.

### Statistical analysis software

Data are prepared and descriptive and correlative statistics are calculated in Stata 14 (StataCorp, 2015). Structural analyses including confirmatory factor and structural equal model analyses are conducted in Mplus 7.14 (Muthén & Muthén, 1998 - 2017).

### Study measures

***Academic performance.*** Performance is measured by three term examinations written by the course instructors. These exams include eleven to fourteen open-response items rubric-graded by the instructors and teaching assistants. Each grader is assigned a single exam item or page of items to grade, providing consistency in scoring across each examination. Exam grading is spot checked by the course instructors, and students in the course have the opportunity to request regrades (most commonly for incorrect total score determination). Any grades that changed due to regrades are reflected in the data used in this study.

***Achievement emotions questionnaire (Pekrun et al., 2011).*** The AEQ was developed to measure achievement emotions. While there are several measures of chemistry anxiety (i.e. ASCI), these measures do not have a comparable enjoyment measure. For consistency, we choose to use the anxiety and enjoyment measures from the AEQ. Each emotion has a subscale designed to measure class-related, study-related and exam-related emotions; emotions are observed to be distinct and separable in these contexts (Goetz et al., 2006).

The enjoyment and anxiety subscales of the AEQ, as utilized in this study, are found in Table 4.2. The exam-related subscales were not administered at the first data collection period; students must have experience in the exam context to evaluate their emotional experiences. Item response options are 'strongly disagree' to 'strongly agree' across a five-point Likert scale. In this study, initial scales include all items on the original AEQ subscales. In the interest of reducing student burden when filling out the questionnaires, item reduction is conducted using modification indices provided by confirmatory factor analysis (Brown, 2015). Items which do not contribute a unique explanation of the constructs of interest are eliminated in a sequential method like the item reduction technique reported by Xu and Lewis (2011).

**Psychometric evidence of AEQ**

Prior to evaluation of the predicted reciprocal causation model, psychometric properties of the measurements are evaluated. Cronbach's alpha is used as a reliability estimate; a value greater than .7 is considered satisfactory (Cortina, 1999). A confirmatory factor analysis approach is used to establish the structure of latent constructs (Brown, 2015). The confirmatory factor model is shown in Figure 4.1. The factors of interest are anxiety and enjoyment at each time point as measured by class-, study- and exam-related subscales.

**Table 4.2.** Anxiety and enjoyment subscale items.

| Subscale | Item |
|---|---|
| Class Related Anxiety | I feel nervous in organic chemistry class.<br>I worry the demands in organic chemistry class might be too great.<br>When I think about organic chemistry class, I feel sick.<br>I worry whether I'm sufficiently prepared for organic chemistry class.<br>I feel scared about organic chemistry class.<br>I get scared that I might say something wrong in organic chemistry class, so I'd rather not say anything. |
| Study Related Anxiety | When I look at the organic chemistry textbook, I get anxious.<br>When I have to study for organic chemistry class, I start to feel sick.<br>While studying for organic chemistry class, I feel like distracting myself in order to reduce my anxiety.<br>I get tense and nervous when studying for organic chemistry.<br>I worry whether I'm able to cope with all the work necessary for organic chemistry class. |
| Exam Related Anxiety | During organic chemistry exams, I feel nervous and uneasy.<br>I get so nervous I wish I could just skip organic chemistry exams.<br>I get so nervous I can't wait for organic chemistry exams to be over.<br>I feel sick to my stomach when thinking about taking organic chemistry exams.<br>I am so anxious that I'd rather be anywhere else than taking organic chemistry exams.<br>I worry whether I will pass organic chemistry exams.<br>My hands get shaky when taking organic chemistry exams. |
| Class Related Enjoyment | I am looking forward to learning a lot in organic chemistry class.<br>I enjoy participating in organic chemistry class so much that I get energized thinking about it.<br>My enjoyment of organic chemistry class makes me want to participate in the class.<br>I enjoy being in organic chemistry class.<br>I get excited about going to organic chemistry class.<br>After organic chemistry class I start looking forward to the next organic chemistry class.<br>I am motivated to go to organic chemistry class because it's exciting. |
| Study Related Enjoyment | When my studies in organic chemistry are going well, it gives me a rush of excitement.<br>I look forward to studying for organic chemistry.<br>I am so happy about the progress I have made in organic chemistry that I am motivated to continue studying for the course.<br>I enjoy the challenge of learning the material for organic chemistry class.<br>Reflecting on my progress in coursework for organic chemistry makes me happy. |
| Exam Related Enjoyment | I look forward to organic chemistry exams.<br>I am happy that I can cope with organic chemistry exams.<br>Before taking organic chemistry exams, I sense a feeling of readiness.<br>Because I enjoy preparing for organic chemistry exams, I'm motivated to do more than is necessary.<br>I look forward to demonstrating my knowledge of organic chemistry on exams. |

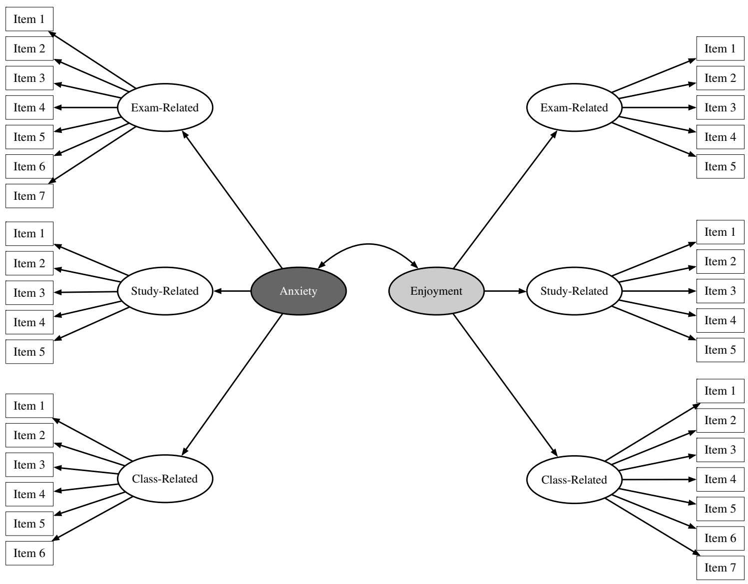


**Figure 4.1**. Confirmatory factor analysis for administrations 2 and 3 of the anxiety and enjoyment subscales.

The longitudinal nature of this study requires that invariance be assessed over time to ensure that the structure of the items and latent constructs is consistent (Widaman, Ferrer, & Conger, 2010); this type of invariance informs us if the instrument used is measuring changes in the intended constructs over time or if variance in the data might be related to changes in other characteristics (Brown, 2015).

The process of longitudinal measurement invariance requires that a series of confirmatory factor analyses be run, with all data fit simultaneously. The configural model allows all parameters (i.e. factor loadings, item intercepts and error variance) to be estimated freely across administrations of the measures. The metric model constrains factor loadings to be equal. The scalar model additionally constrains item intercepts. The strict model constrains factor loadings, item intercepts and individual item error variances to be identical at each time of data collection. Ideal invariance would be represented by identical fit between each model; acceptable, but not identical, levels of fit indices in a more constrained model are also acceptable (Widaman & Thompson, 2003). Decreases in fit are measured by the Satorra-Bentler-adjusted $\chi^2$ comparison (Satorra & Bentler, 2001) along with traditional confirmatory factor analysis fit indices.

### Model evaluation: structural equation modelling

A structural equation modelling approach using MPlus version 7.14 (Muthén & Muthén, 1998 - 2017) is conducted to evaluate the relationships in research questions (1) and (2). Models with predicted relationships between anxiety, enjoyment and performance at each time are developed and tested based on previous reciprocal causation studies (Pekrun et al., 2014; Villafañe et al., 2016). Model 1 (Figure 4.2) is the hypothesised reciprocal causation model, as predicted by the control-value theory. Model 2 (Figure 4.3a) demonstrates performance as a predictor of future enjoyment and anxiety measures. Model 3 (Figure 4.3b) holds that affect measures serve as predictors of performance. Model 4 (Figure 4.3c) holds that while longitudinal effects of each affect and exam performance are predictive, there is no relationship between the two. The 'snowball' effect in which scores on Exam 1 have a predictive ability on both of the subsequent exams is unique to the context of organic chemistry; Villafañe et al. (2016) found the impact of this effect in a study relating self-efficacy to organic chemistry performance, and is conceptually supported by other studies of the highly cumulative organic chemistry curriculum (Grove & Bretz, 2010; Lynch & Trujillo, 2011). We test the reciprocal model without this snowball effect to determine its validity in Model 5 (Figure 3d). Composite scores of averages of all items for each subscale of each achievement emotion are used rather than the measurement model (Figure 1) as the confirmatory factor analyses supports that the items assigned to each subscale measure a single construct and can be represented as a single value (Little, Cunningham, Shahar, & Widaman, 2002; Rushton, Brainerd, & Pressley, 1983)

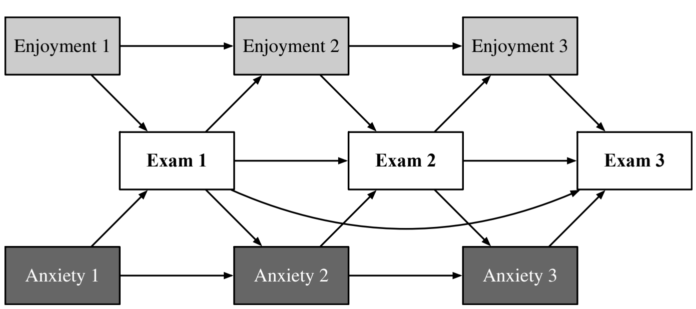


**Figure 4.2**. Proposed reciprocal causation model including "snowball" examination effect.

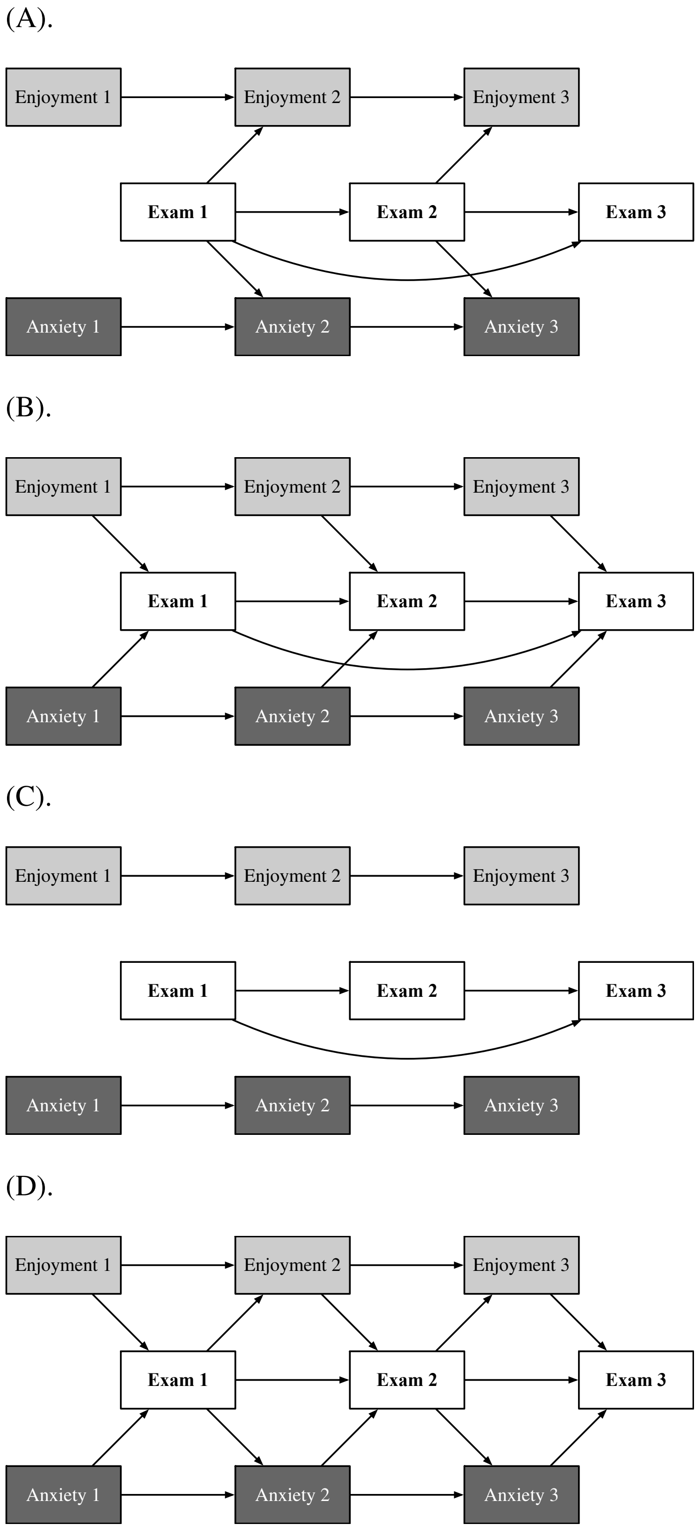


**Figure 4.3.** Alternative models evaluated. (A). Model 2: Examination effects. (B). Model 3: Achievement Emotion effects. (C). Model 4: Autoregressive. (D). Model 5: Reciprocal causation model without “snowball” examination effect.

## Results

In order to address the research questions guiding this study, psychometric evidence for the scales used was first established. The results supporting this evidence are evaluated using descriptive statistics, followed by confirmatory factor analysis and longitudinal measurement invariance. The research questions are addressed in the structural equation modelling procedure, which is discussed after psychometric evidence is provided.

### Descriptive statistics

Response rates for each administration of the instruments used in this research are found in Table 4.3. Descriptive and correlative data are reported in Table 4.4. These data demonstrate that the students in the course have a normal distribution of exam scores and scores on the study measures. Negative correlations between exam performance and anxiety along with positive correlations between enjoyment and exam performance align with the theory-based valence. Cronbach's alpha values range from .94 to .96, indicating high and consistent reliability across time (Cortina, 1999). The gradual increase in anxiety over time as indicated by mean values in Table 4.4 is accompanied by a lack of pattern for enjoyment; this might indicate that while anxiety increases consistently, enjoyment varies across a course.

**Table 4.3**. Response rates for survey and examination administrations.

| Measure | | *N* | Administration | | |
|---|---|---|---|---|---|
| | | | 1 | 2 | 3 |
| Examination | Number | 907 | 880 | 828 | 783 |
| | Rate (%) | 100 | 97 | 91 | 86 |
| Enjoyment | Number | 907 | 827 | 781 | 758 |
| | Rate (%) | 100 | 91 | 86 | 84 |
| Anxiety | Number | 907 | 830 | 776 | 762 |
| | Rate (%) | 100 | 92 | 86 | 84 |

**Table 4.4.** Descriptive statistics and Pearson correlations for study measures.

| | Exam | | | Enjoyment | | | Anxiety | | |
|---|---|---|---|---|---|---|---|---|---|
| Time | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 |
| *n* | 880 | 828 | 783 | 827 | 781 | 758 | 830 | 776 | 762 |
| *M* | 0 | 0 | 0 | 3.34 | 3.09 | 3.31 | 3.17 | 3.19 | 3.25 |
| *SD* | 1 | 1 | 1 | .67 | .69 | .82 | .79 | .83 | .87 |
| Min | -5.4 | -3.34 | -2.58 | 1 | 1 | 1 | 1 | 1 | 1 |
| Max | 1.86 | 1.78 | 2.25 | 5 | 5 | 5 | 5 | 5 | 5 |
| Skew | -.6 | -.63 | -.03 | -.3 | -.07 | -.35 | -.08 | -.05 | .02 |
| Kurt | .35 | -.24 | -.68 | .51 | .16 | .29 | -.41 | -.30 | -.44 |
| Alpha | | | | .94 | .95 | .95 | .94 | .96 | .96 |
| exam1 | 1.00 | .71 | .65 | .17 | .25 | .24 | -.22 | -.34 | -.32 |
| exam2 | | 1.00 | .64 | .16 | .25 | .31 | -.18 | -.28 | -.32 |
| exam3 | | | 1.00 | .12 | .21 | .26 | -.20 | -.27 | -.32 |
| enj1 | | | | 1.00 | .73 | .62 | -.47 | -.39 | -.33 |
| enj2 | | | | | 1.00 | .77 | -.46 | -.56 | -.45 |
| enj3 | | | | | | 1.00 | -.39 | -.46 | -.49 |
| anx1 | | | | | | | 1.00 | .76 | .68 |
| anx2 | | | | | | | | 1.00 | .79 |
| anx3 | | | | | | | | | 1.00 |

**Psychometric evidence**

Confirmatory factor analyses at each time for the subscales of anxiety and enjoyment provided good fit criteria for each iteration as defined by the $\chi^2$, comparative fit index (CFI), standardised root-mean square residual (SRMR) and root mean square standard error of approximation (RMSEA) comparable to other applications AEQ subscales in college classrooms (see Table 4.5; Hoyle & Panter, 1995; Hu & Bentler, 1999; Pekrun et al., 2014). Good fit at each time provides psychometric support for the structural validity of the subscales used (Pekrun et al., 2011). Composite scores are calculated based on the average response to each item for both achievement emotions due to the strength of structural validity evidence (Little et al., 2002; Rushton et al., 1983).

A longitudinal measurement invariance analysis is conducted (see Table 4.6). Despite the statistically significant values for change in $\chi^2$, the good fit of the model at each increasing level of constraint enables us to establish that the same constructs of anxiety and enjoyment are similarly measured at each time period and thus can be considered to be longitudinally invariant (Widaman et al., 2010; Widaman & Thompson, 2003).

**Table 4.5.** Confirmatory factor analysis fit information for measurement models.

| | $\chi^2$ | *df* | CFI | SRMR | RMSEA |
|---|---|---|---|---|---|
| Time 1 | 793.51$^*$ | 204 | .924$^\pm$ | .047$^\pm$ | .059$^\pm$ |
| Time 2 | 1950.64$^*$ | 520 | .898 | .064$^\pm$ | .058$^\pm$ |
| Time 3 | 2026.68$^*$ | 520 | .898 | .058$^\pm$ | .061$^\pm$ |

*Note*. $\chi^2$ = conventional chi-square fit statistic (under maximum-likelihood estimation); cut-off values for CFI > .90; SRMR < .09; RMSEA < .05 good, .05-.08 reasonable, .08-.10 mediocre.
"$^*$" indicates a statistically significant value at p < .0001
"$^\pm$" indicates a value within the boundaries for good fit

**Table 4.6**. Longitudinal measurement invariance fit information and model comparisons for measurement models.

| Model | $\chi^2$ | *df* | CFI | RMSEA | SRMR | Δ*df* | SBΔ$\chi^2$ |
|---|---|---|---|---|---|---|---|
| Configural | 8009.405* | 3806 | .908$^\pm$ | .056$^\pm$ | .035$^\pm$ | | |
| Metric | 8103.151* | 3852 | .907$^\pm$ | .057$^\pm$ | .035$^\pm$ | 46 | 93.19* |
| Scalar | 8424.939* | 3896 | .900$^\pm$ | .058$^\pm$ | .036$^\pm$ | 44 | 343.12* |
| Strict | 8568.805* | 3940 | .898 | .036$^\pm$ | .059$^\pm$ | 44 | 134.97* |

*Note*. $\chi^2$ = conventional chi-square fit statistic (under maximum-likelihood estimation); cut-off values for CFI > .90; SRMR < .09; RMSEA < .05 good, .05-.08 reasonable, .08-.10 mediocre.
"$^*$" indicates a statistically significant value at p < .0001
"$^\pm$" indicates a value within the boundaries for good fit

### Structural equation modelling

A series of path models are evaluated to test proposed relationships between the study measures. Table 4.7 includes fit statistics for each model, tested with the Satorra-Bentler-adjusted $\chi^2$ value for comparison (Satorra, 2000; Satorra & Bentler, 2001). Per Pekrun et al. (2014), two of these indexes indicating good fit are acceptable to determine that the model is appropriate to continue with analysis. Models 1, 2 and 5 have two indices indicating good fit; Models 3 and 4 are discarded due to poor fit. Upon comparing Model 1 to 2 and 5, the significant change in $\chi^2$ indicates that Model 1 has significantly better fit, supported by CFI and SRMR values (i.e. higher and lower, respectively) than Models 2 and 5. The model indicating best fit is the reciprocal model with the snowball effect, Model 1. Standardised path coefficients for Model 1 are reported in Figure 4.4.

**Table 4.7.** Fit statistics for the reciprocal causation (1) and alternative models (2-5).

| Model | $\chi^2$ | *df* | CFI | SRMR | $\Delta df$ | SB$\Delta\chi^2$ |
|---|---|---|---|---|---|---|
| Reciprocal (1) | 174.70* | 18 | .943$^{\pm}$ | .067$^{\pm}$ | | |
| Examination Effects (2) | 175.26* | 19 | .941$^{\pm}$ | .070$^{\pm}$ | 1 | 14.050* |
| Achievement Emotions Effects (3) | 300.87* | 22 | .989$^{\pm}$ | .117 | | |
| Autoregressive (4) | 303.69* | 23 | .893 | .126 | | |
| Reciprocal without Snowball (5) | 247.49* | 19 | .917$^{\pm}$ | .071$^{\pm}$ | 1 | 79.12* |

*Note.* $\chi^2$ = conventional chi-square fit statistic (under maximum-likelihood estimation), SB = Sartorra-Bentler adjusted difference in $\chi^2$ ; Cut-off values for CFI > .90; SRMR .09, "*" indicates a statistically significant value at $p$ < .0001 , "±" indicates a value within the boundaries for good fit.

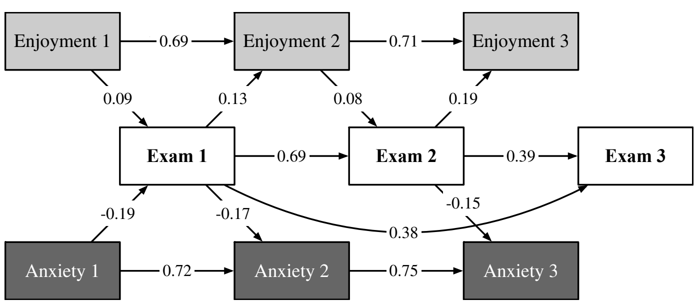


**Figure 4.4.** Standardized coefficients for reciprocal causation structural model of the relationship between anxiety, enjoyment and examination performance. All pathway coefficients are significant at $p$ < .05.

## Discussion

The results of this study provide empirical support for the reciprocal relationship between affect and performance. The AEQ subscales regarding class-related, study-related and exam-related anxiety and enjoyment were utilised and shown to produce reliable and structurally valid data in the population of interest (Pekrun et al., 2002; 2014). This successful method for analysing the interaction of anxiety and enjoyment along with achievement overcomes previous research challenges in measuring affect in a static research design. This study demonstrated longitudinal invariance among measurements of affect. Ultimately, the information presented in this report helps us to understand the emotional landscape of the organic chemistry course.

In response to Research Question 1: What is the relationship between anxiety, enjoyment and examination performance? The best fitting relationship between anxiety, enjoyment and performance in a postsecondary organic chemistry course is reciprocal. Model 1, as represented in Figure 4.2, best fits the relationships of interest.

In response to Research Question 2: What are the strength and valence of these relationships? The relationships between anxiety and exam performance in both directions are significant, small and negative. The relationships between enjoyment and exam performance in both directions are significant, small and positive. In the case of anxiety, the interrelated impacts generally decrease in size over time and are consistently stronger from exam performance to subsequent anxiety measures. For enjoyment, however, the impact of exam performance on subsequent enjoyment scores are on average equivalent to the reciprocal impact, but directional relationships from enjoyment to exam performance decrease in size over time while those from exam performance to enjoyment increase over time.

The results of this study demonstrate that affect and achievement are reciprocally related in organic chemistry classes, as predicted by the control-value theory of achievement emotions. While many of the paths in the best-supported model (Model 1, Figure 4.2) are statistically significant, the extent of the relationships must be understood through the lens of the size of the effects found in this population. The standardised coefficients, i.e. effect size of the relationships found in this study, can be considered moderate in the educational context. The effect sizes of impacts from anxiety measures, enjoyment measures and achievement measures (i.e. examination performance) to subsequent measures of the same are the largest. The impacts from affect measures to achievement are largest at the first examination, while those from examination results to future affect measures grow over time regarding enjoyment and shrink regarding anxiety. These effect size values demonstrate that the results presented here should not be over interpreted; there are many factors which impact academic performance that were not measured in this study.

### Implications for Instruction

The data presented here can enable faculty to recognise the importance of affect on predicting performance. This study demonstrates that a relationship between the first exam and subsequent emotional experiences is the largest (see Figure 4.4). This result could inform pedagogical decision-making. The incorporation of active learning strategies has demonstrated learning gains through emphasising perceived autonomy (i.e., control) in learning in organic chemistry (Black & Deci, 2000). Others have found that a brief introduction followed by more in-depth instruction on each key subject in the organic chemistry course was successful in reducing anxiety (Grove et al., 2008). A qualitative study of a nursing chemistry course utilising peer-led team learning pedagogy found decreased anxiety with the reformed curriculum (White, Rowland, & Paesis-Katz, 2012; Wilson & Varma-Nelson, 2016). In one studio-based course and another Process-Oriented Guided Inquiry Learning classroom, chemistry students showed no difference in anxiety from a traditional lecture, indicating that instructional style is not the only influencing factor on affect responses (Chase, Pakhira, & Stains, 2013; Oliver-Hoyo & Allen, 2005). One intervention for reducing anxiety in an introductory chemistry course was the use of a counsellor to provide guidance during scheduled laboratory time; the study found reduced anxiety in the group visited by the counsellor but no significant association with performance (Abendroth & Friedman, 1983). Efforts to decrease anxiety, like these examples, can be incorporated with efforts to increase control over the content to improve overall performance.

Control-value theory posits that the nine achievement emotions (Table 4.1) are moderated by control and value placed on the achievement activities (Pekrun, 2000, 2006). The unique environment of organic chemistry among the chemistry disciplines and among science courses at the postsecondary level is highlighted by its presence on applications for graduate and professional school (Grove et al., 2008;

Muller, 2013). The observed increase in anxiety over time aligns with previous findings (Ahmed et al., 2013). The non-directional longitudinal changes in enjoyment may be related to interest, as has been found in other contexts (Ainley & Ainley, 2011; Nieswandt, 2007; Singh et al., 2002); as the content shifts throughout the term, students may place particularly high value on content, and value appraisals may change. Measuring these emotions throughout the semester can empower faculty members and teachers to maximise the effectiveness of instruction.

**Implications for Research**

Future researchers interested in analysing the effect of affect on achievement can use the methodology outlined here, while improving in areas that were not assessed in this study. Other studies utilising the achievement emotions questionnaire (AEQ) and exploring relationships with affect through the lens of the control-value theory have found mediation relationships with motivation and self-regulated learning, both of which could be measured simultaneously with the achievement emotions in the interest of measuring their differential impact (Aydin et al., 2011; Goetz et al., 2007; Pekrun et al., 2002). One limitation of this study is that other predictive factors for anxiety and enjoyment were not measured, including instructor enjoyment- which might result in transmission (i.e., Frenzel, Goetz, Ludke, Pekrun, & Sutton, 2009)- and the other achievement emotions: hope, pride, anger, shame, relief, boredom, hopelessness (Pekrun, 2000). Because mathematics is known to be related to chemistry achievement, measuring attitudes and efficacy towards mathematics might provide insight into achievement in future studies. Future work should also incorporate measures of reasonable mediators as has been done in prior research; psychological correlates such as self-efficacy and motivation are potential mediators of the impact of affect on achievement (Aydin et al., 2011; Kurbanoglu & Akim, 2010; Lynch & Trujillo, 2011; Teo et al., 2014; Villafañe et al., 2016).

One key contribution of this work is the designation of the subscales for anxiety and enjoyment of the AEQ as providing valid and reliable data in the population of undergraduate organic chemistry students.

Determining which tools are best suited, as the AEQ subscales for anxiety and enjoyment are supported in this study, are important for future research (Pekrun et al., 2002). The importance of integrated measures of affect, motivation and performance is vital to the growth of the field (Linnenbrink, 2006). This study has demonstrated a successful method for the integration of such measures, and such methodology should be utilised in future research in order to gain a richer context of affect and student learning.

## Conclusion

Achievement emotions such as enjoyment and anxiety are ubiquitous in the undergraduate environment (Pekrun, 2000); these emotions contribute to academic success in an influential way (Daniels et al., 2009; Huang, 2011; Pekrun et al., 2011; Pekrun et al., 2002; Pekrun et al., 2009; Putwain et al., 2013). To increase classroom performance, affect must be engaged to empower students to learn new content. The results of this study demonstrate that anxiety and enjoyment predict performance. Instructors could act to reduce anxiety and increase enjoyment by utilising active learning strategies (Abendroth & Friedman, 1983; Aydin et al., 2011; Kurbanoglu & Akim, 2010). Researchers could utilise measures of achievement emotions when evaluating the effects of classroom reforms with the methodology described above.

## References


Abendroth, W., & Friedman, F. (1983). Anxiety reduction for beginning chemistry students. *Journal of Chemical Education, 60*, 25-26. doi:10.1021/ed060p25

Ahmed, W., van der Werf, G., Kuyper, H., & Minnaert, A. (2013). Emotions, self regulated learning, and achievement emotions in mathematics: A growth curve analysis. *Journal of Educational Psychology, 105*, 150-161. doi:10.1037/a0030160

Ainley, M., & Ainley, J. (2011). Students engagement with science in early adolescence: The contribution of enjoyment to students' continuing interest in learning about science. *Contemporary Educational Psychology, 36*, 4-12. doi:10.1016/j.cedpsych.2010.08.001

Anderson, T. L., & Bodner, G. M. (2008). What can we do about 'Parker'? A case study of a good student who didn't 'get' organic chemistry. *Chemistry Education Research and Practice, 9*, 93-101. doi:10.1039/B806223B

Anderson, T. L., & Evans, F. B. (1974). Causal models in education research: Recursive models. *American Educational Research Journal, 11*, 29-39. doi:10.3102/00028312011001029

Artino, A. R., & Jones, K. D. (2012). Exploring the complex relations between achievement emotions and self-regulated learning behaviors in online learning. *The Internet and Higher Education, 15*, 170-175. doi:10.1016/j.iheduc.2012.01.006

Artino, A. R., La Rochelle, J. S., & Durning, S. J. (2010). Second-year medical students' motivational beliefs, emotions, and achievement. *Medical Education, 44*, 1203-1212. doi:10.1111/j.1365-2923.2010.03712.x

Aydin, Y. C., Uzuntiryaki, E., & Demirdogen, B. (2011). Interplay of motivational and cognitive strategies in predicting self-efficacy and anxiety. *Educational Psychology, 31*, 55-66. doi:10.1080/01443410.2010.518561

Bauer, C. F. (2008). Attitude toward chemistry: A semantic differential instrument for assessing curriculum impacts. *Journal of Chemical Education, 85*, 1440-1445. doi:10.1021/ed085p1440

Black, A. E., & Deci, E. L. (2000). The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: A self determination theory perspective. *Science Education, 84*, 740-756. doi:10.1002/1098-237X(200011)84:6<740::AID-SCE4>3.0.CO;2-3

Bowen, C. W. (1999). Development and score validation of a chemistry laboratory anxiety instrument (CLAI) for college chemistry students. *Educational and Psychological Measurement, 59*, 171-185. doi:10.1177/0013164499591012

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research* (2nd ed.). New York: Guilford Press.

Chan, J. Y. K., & Bauer, C. F. (2016). Learning and studying strategies used by general chemistry students with different affective characteristics. *Chemistry Education Research and Practice, 17*, 675-684. doi:10.1039/C5RP00205B

Chase, A., Pakhira, D., & Stains, M. (2013). Implementing Process-Oriented, Guided Inquiry Learning for the first time: Adaptations and short-term impacts on students' attitude and performance. *Journal of Chemical Education, 90*, 409-416. doi:10.1021/ed300181t

Cortina, J. M. (1999). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*, 98-104. doi:10.1037/0021-9010.78.1.98

Daniels, L. M., Stupinsky, R. H., Pekrun, R., Haynes, T. L., Perry, R. P., & Newall, N. E. (2009). A longitudinal analysis of achievement goals: From affective antecedents to emotional effects and achievement outcomes. *Journal of Educational Psychology, 101*. doi:10.1037/a0016096

Duis, J. M. (2011). Organic chemistry students' perspectives on fundamental concepts and misconceptions: An exploratory study. *Journal of Chemical Education, 88*, 346-350. doi:10.1021/ed1007266

Dwyer, A. O., & Childs, P. (2014). Organic chemistry in action! Developing an intervention program for introductory organic chemistry to improve learners' understanding, interest, and attitudes. *Journal of Chemical Education, 91*, 987-993. doi:10.1021/ed400538p

Eddy, R. M. (2000). Chemophobia in the college classroom: Extent, courses, and student characteristics *Journal of Chemical Education, 77*, 514-517. doi:10.1021/ed077p514

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology. *American Psychologist, 56*, 218-226. doi:10.1037/0003-066X.56.3.218

Frenzel, A. C., Goetz, T., Ludke, O., Pekrun, R., & Sutton, R. E. (2009). Emotional transmission in the classroom: Exploring the relationship between teacher and student enjoyment. *Journal of Educational Psychology, 101*, 705-716. doi:10.1037/a0014695

Goetz, T., Frenzel, A. C., Hall, N. C., & Pekrun, R. (2007). Antecedents of academic emotions: Testing the internal/external frame of reference model of academic enjoyment. *Contemporary Educational Psychology, 33*, 9-33. doi:10.1016/j.cedpsych.2006.12.002

Goetz, T., Pekrun, R., Hall, N. C., & Haag, L. (2006). Academic emotions from a social cognitive perspective: Antecedents and domain specificity of students' affect in the context of Latin instruction. *British Journal of Educational Psychology, 76*(289-308). doi:10.1348/00709905X42860

Grove, N. P., & Bretz, S. L. (2010). Perry's scheme of intellectual and epistemological development as a framework for describing difficulties in learning organic chemistry. *Chemistry Education Research and Practice, 11*, 207-211. doi:10.1039/C005469K

Grove, N. P., Cooper, M. M., & Cox, E. L. (2012). Does mechanistic thinking improve student success in organic chemistry? *Journal of Chemical Education, 89*(850-853). doi:10.1021/ed200394d

Grove, N. P., Hershberger, J. W., & Bretz, S. L. (2008). Impact of a spiral organic curriculum on student attrition and learning. *Chemistry Education Research and Practice, 9*, 157-162. doi:10.1039/B806232N

Henderlierer, J., Smart, R., Anderson, J., & Elian, O. (2001). How do organic chemistry students understand and apply hydrogen bonding? *Journal of Chemical Education, 78*, 1126-1130. doi:10.1021/ed078p1126

Hoyle, R. H., & Panter, A. T. (1995). Writing about structural equation models. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 158-176). Newbury Park, CA: Sage.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modelling: Multidisciplinary Journal, 6*, 1-55. doi:10.1080/10705519909540118

Huang, C. (2011). Achievement goals and achievement emotions: A meta-analysis. *Educational Psychology Review, 23*, 359-388. doi:10.1007/s10648-011-9155-x

Kurbanoglu, N. I., & Akim, A. (2010). The relationships between university students' chemistry laboratory anxiety, attitudes, and self-efficacy beliefs. *The Australian Journal of Teacher Education, 35*(8), 48-59. doi:10.14221/ajte.2010v35n8.4

Linnenbrink, E. A. (2006). Emotion research in education: Theoretical and methodological perspectives on the integration of affect, motivation, and cognition. *Educational Psychology Review, 18*, 307-314. doi:10.1007/s10648-006-9028-x

Linnenbrink-Garcia, L., & Pekrun, R. (2011). Students' emotions and academic engagement: Introduction to the special issue. *Contemporary Educational Psychology, 36*, 1-3. doi:10.1016/j.cedpsych.2010.11.004

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling, 9*, 151-173. doi:10.1207/S15328007SEM0902_1

Lynch, D. J., & Trujillo, H. (2011). Motivational beliefs and learning strategies in organic chemistry. *International Journal of Science and Mathematics Education, 9*, 1351-1365. doi:10.1007/s10763-101-9264-x

Mandler, G. (1990). A constructivist theory of emotion. In N. L. Stein, B. Leventhal, & T. R. Trabasso (Eds.), *Psychological and biological approaches to emotion* (pp. 21-43). Hillsdale, NJ: Lawrence Erlbaum Associates.

McCarthy, W. C., & Widanski, B. B. (2009). Assessment of chemistry anxiety in a two year college. *Journal of Chemical Education, 86*, 1447-1449. doi:10.1021/ed086p1447

Muller, D. (2013). Reforming premedical education- out with the old, in with the new. *New England Journal of Medicine, 368*, 1567-1569. doi:10.1056/NEJMp1302259

Muthén, L. K., & Muthén, B. O. (1998 - 2017). *Mplus User's Guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.

Nieswandt, M. (2007). Student affect and conceptual understanding in learning chemistry. *Journal of Research in Science Teaching, 44*, 908-937. doi:10.1002/tea.20169

Oliver-Hoyo, M., & Allen, D. (2005). Attitudinal effects of a student-centered active learning environment. *Journal of Chemical Education, 82*, 944-949. doi:10.1021/ed082p944

Pekrun, R. (2000). A social-cognitive, control-value theory of achievement emotions. In J. Heckhausen (Ed.), *Motivational psychology of human development* (pp. 143-163). Oxford: Elsevier.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*, 315-341. doi:10.1007/s10648-006-9029-9

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Academic Emotions Questionnaire (AEQ). *Contemporary Educational Psychology, 36*, 36-48. doi:10.1016/j.cedpsych.2010.10.002

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research *Educational Psychologist, 37*, 91-105. doi:10.1207/S15326985EP3792_4

Pekrun, R., Hall, N. C., Goetz, T., & Perry, R. P. (2014). Boredom and academic achievement: Testing a model of reciprocal causation. *Journal of Educational Psychology, 106*, 696-710. doi:10.1037/a0036006

Pekrun, R., J., E. A., & Maier, M. A. (2009). Achievement goals and achievement emotions: testing a model of their joint relations with academic performance. *Journal of Educational Psychology, 101*, 115-135. doi:10.1037/a0013383

Putwain, D. W., Larkin, D., & Sander, P. (2013). A reciprocal model of achievement goals and learning related emotions in the first year of undergraduate study. *Contemporary Educational Psychology, 38*, 361-374. doi:10.1016/j.cedpsych.2013.07.003

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance A systematic review and meta analysis. *Psychological Bulletin, 138*, 353-387. doi:10.1037/a0026838

Rushton, J. P., Brainerd, C. J., & Pressley, M. (1983). Behavior development and construct validity: The principle of aggregation. *Psychological Bulletin, 94*, 18-38.

Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In R. D. H. Heijmans, D. S. G. Pollock, & A. Satorra (Eds.), *Innovation in multivariate statistical analysis* (pp. 233-247). London: Kluwer Academic Publishers.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychomerika, 66*, 5507-5514. doi:10.1007/BF02296192

Singh, K., Granville, M., & Dika, S. (2002). Mathematics and science achievement: Effects of motivation, interest, and academic engagement. *Journal of Educational Research, 95*, 323-332.

Spielberger, C. D. (1972). *Anxiety: Current trendds in theory and research*. New York, NY: Academic Press.

StataCorp. (2015). *Stata Statistical Software: Release 14*. College Station, TX: StataCorp LLC.

Steiner, R., & Sullivan, J. (1984). Variables correlating with student success in organic chemistry. *Journal of Chemical Education, 61*, 1072-1074. doi:10.1021/ed061p1072

Teo, T. W., Goh, M. T., & Yeo, L. W. (2014). Chemistry education research trends: 2004-2013. *Chemistry Education Research and Practice, 15*(470-487). doi:10.1039/C4RP00104D

Tien, L. T., Roth, V., & Kampmeier, J. A. (2022). Implementation of a peer-led team learning instructional approach in an undergraduate organic chemistry course. *Journal of Research in Science Teaching, 39*, 606-632. doi:10.1002/tea.10038

Villafañe, S. M., Xu, X., & Raker, J. R. (2016). Self-efficacy and academic performance in first-semester organic chemistry: Testing a model of reciprocal causation. *Chemistry Education Research and Practice, 17*(973-984). doi:10.1039/C6RP00119J

Villavicencio, F. T., & Bernardo, A. B. I. (2013). Positive academic emotions moderate the relationship between self-regulation and academic achievement. *British Journal of Educational Psychology, 83*, 329-340. doi:10.1111/j.2044-8279.2012.02064.x

White, P., Rowland, A. B., & Paesis-Katz, I. (2012). Peer-led team learning model in a graduate-level nursing course. *Journal of Nursing Education, 51*, 471-475. doi:10.3928/01484834-20120706-03

Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives, 4*, 10-18. doi:10.1111/j.1750-8606.2009.00110.x

Widaman, K. F., & Thompson, J. S. (2003). On specifying the null model for incremental fit indices in structural equation modelling. *Psychological Methods, 8*, 16-37. doi:10.1037/1082-989X.8.1.16

Wilson, S. B., & Varma-Nelson, P. (2016). Small groups, significant impact: A review of peer-led team learning research with implications for STEM education. *Journal of Chemical Education, 93*, 1686-1702. doi:10.1021/acs.jchemed.5b00862

Xu, X., & Lewis, J. E. (2011). Refinement of a chemistry attitude measure for college students. *Journal of Chemical Education, 88*(5), 561-568. doi:10.1021/ed900071q

Zoller, U., & Pushkin, D. (2007). Matching higher-order cognitive skills (HOCS) promotion goals with problem-based laboratory practice in a freshman organic chemistry course. *Chemistry Education Research and Practice, 8*, 153-171. doi:10.1039/B6RP90028C

# CHAPTER FIVE:

# "I'M SO BORED!" TESTING A MODEL OF RECIPROCAL EFFECTS BETWEEN BOREDOM AND ACADEMIC PERFORMANCE IN FIRST-SEMESTER ORGANIC CHEMISTRY

## Introduction

Academic boredom is a negative emotion linked to limited engagement with and attention given to a task such as solving homework problems, taking notes in a lecture, or completing examinations (Eastwood, Frischen, Fenske, & Smilek, 2012; Tze, Daniels, & Klassen, 2016). Boredom has been studied in the context of repetitive work environments and found to be an antecedent of destructive activities such as truancy (Watt & Hargis, 2010). In the field of science education research, there is little mention of boredom; discussions of boredom are limited to reports about the instructional laboratory in the chemistry context (e.g., Galloway & Bretz, 2016; Supalo, Humphrey, Mallouk, Wohlers, & Carlsen, 2016; Trehan, Brar, Arora, & Kad, 1997). Efforts to incorporate contextualized examples into the postsecondary chemistry curriculum, such as climate change in general chemistry (King et al., 2015) or drug leads and design in organic chemistry (Forbes, 2004), are targeted at reducing boredom and increasing interest.

Boredom has been shown to vary between instructional environments, subject areas, and testing situations (Acee et al., 2010; Bench and Lench, 2013; Goetz, Frenzel, Pekrun, & Hall, 2006; Goetz, Pekrun, Hall, & Haag, 2006). In the context of learning, boredom has been defined as a negative, deactivating achievement emotion resulting from a lack of control over and low level of value assigned to a given achievement task (Pekrun, Goetz, Daniels, Stupnisky, & Perry, 2010). Achievement emotions such as anxiety have been shown to have negative impacts on academic achievement (Aydin, Uzuntiryaki, & Demirdogen, 2011; Black & Deci, 2000; Chan & Bauer, 2016; Eddy, 2000; Lynch & Trujillo, 2011);

however, emotions such as boredom have received little attention in the literature (Frenzel, Pekrun, & Goetz, 2007; Pekrun et al., 2010).

**Theoretical Framework**

The complex nature of emotions experienced by students poses a challenge to researchers who seek to understand the affective factors that influence achievement and for instructors hoping to increase learning (Frenzel, Thrash, Pekrun, & Goetz, 2007). Pekrun (1992, 2006) developed the control-value theory of achievement emotions (CVT) as a framework for understanding the reciprocal relationship between emotions and academic performance. CVT posits that there are nine achievement emotions that occur in the academic environment characterized by their valence (i.e., positive or negative) and activation (see Table 5.1). Boredom is characterized as a negative, deactivating emotion (Pekrun, 2006; Pekrun et al., 2010; Pekrun & Perry, 2014).

**Table 5.1**. Valence and activation of achievement emotions within control-value theory.

| | Positive Valence | Negative Valence |
|---|---|---|
| Activating | Enjoyment, Hope, Pride | Anxiety, Anger, Shame |
| Deactivating | Relief | Boredom, Hopelessness |

Pekrun (2006) posited that boredom is the result of a student's perceived knowledge level, referred to as control, and the value ascribed to learning and assessment activities being performed. Prior to CVT, Csikszentmihalyi’s (1975) flow theory prevailed, which suggests that boredom is only experienced in under-challenging situations and that anxiety was prevalent when students were over-challenged. More recent research, however, asserts that due to a misalignment of control, under-challenging situations will appear repetitive and dull to students while over-challenging situations will result in frustration and ennui, both inducing boredom (e.g., Acee et al., 2010; Pekrun et al., 2010). Empirical studies of antecedents to boredom have found a negative relationship between boredom and control (e.g., Daniels et al., 2009; Goetz et al., 2006; Niculescu, Tempelaar, Dailey-Herbert, Segers, & Gijselaers, 2015; Perry, Hladkyj, Pekrun, &

Pelletier, 2001). Studies have used the Precursors to Boredom Scale to measure boredom due to over- and under-challenge (e.g., Daschmann, Goetz, & Stupnisky, 2011; Tze, Daniels, & Klassen, 2013).

The experience of boredom can be influenced by control and value assigned to a task (Pekrun, 2006). A student may not experience boredom despite challenge level if the value assigned to an achievement task is high; high value increases attention (Eastwood et al., 2012). Boredom, then, can be considered antonymous to the construct *interest*, an affective state which encourages greater engagement and has been positively linked to chemistry achievement (Ferrell & Barbera, 2015). Interest is linked to increased value assigned to given content (Harackieqicz & Hulleman, 2010; Perkins, Adams, Pollock, Finklestein, & Wieman, 2005). Boredom is a state opposite to interest and has not been frequently explored in the context of chemistry education. A focus on alleviating boredom would complement interest and identify additional means for increasing achievement.

Research on boredom outside of academic environments involves two widely-used self-report instruments: the Boredom Proneness Scale (Farmer and Sundberg, 1986) and the Zuckerman Boredom Susceptibility Scale (Mercer-Lynn, Flora, Fahlman, & Eastwood, 2011; Vodanovich & Watt, 2016; Zuckerman, 1979). With these instruments, boredom has been shown to be negatively correlated with performance in repetitive tasks (Fisher, 1993). Higher boredom was also associated with higher cognitive failure on tasks in a military environment, providing evidence for the lack of attention resulting from boredom understood within the control-value theory (Eastwood et al., 2012; Pekrun, 2006; Pekrun et al., 2010; Tze et al., 2016; Wallace, Vodanovich, & Restino, 2003). When bored, a student experiences an aversive, uncomfortable state (Bench & Lench, 2013), an overwhelming feeling that time is moving more slowly (Watt, 1991), and is motivated to escape the situation (Bench & Lench, 2013), accompanied by a low level of physiological excitement (Mikulas & Vodanovich, 1993). The lack of attention resulting from such deactivation leads to a decrease in achievement (Bench & Lench, 2013; Eastwood et al., 2012; Pekrun et al., 2010; Pekrun, Hall, Goetz, & Perry, 2014). Boredom has been found to result in avoidance coping mechanisms (Tanaka & Murayama, 2014), in which case a decrease in academic achievement can also be expected (Mann & Robinson, 2009).

Observational and experimental studies of academic boredom are rare (Goetz et al., 2006; Pekrun et al., 2010). This is particularly true in science education, in which studies referencing boredom are clustered in the context of the laboratory environment (Galloway & Bretz, 2016; Supalo et al., 2016; Trehan, et al., 1997). In classroom (i.e., lecture) settings, boredom is understood to be a state emotion (Fahlman, Mercer-Lynn, Flora, & Eastwood, 2013; Laukenmann et al., 2003). Instruments have been developed to measure academic state boredom such as the Academic Boredom Coping Scale (Nett, Goetz, & Daniels, 2010), the Academic Boredom Scale (Acee et al., 2010), and the Learning-Related Boredom Scale (Pekrun, Goetz, & Perry, 2005). Academic boredom has consistently been shown to have a negative relationship with performance (Singh, Granville, & Dika, 2002) and has been identified as a motivational barrier to achievement (Acee et al., 2010). The growth of boredom over time is associated with a reduction in the use of metacognitive strategies for learning along with academic achievement (Ahmed, van der Werf, Kuyper, & Minnaert, 2013). Teacher ratings of student ability have also been shown to correlate to student reports of boredom (e.g., Robinson, 1975; Wong & Csikszentmihalyi, 1991). Boredom is consistently negatively associated with achievement in the science classroom. Learning more about boredom in the multiple contexts, including chemical education, will enable researchers and instructors alike to evaluate instructional strategies designed to increase outcomes for their impact of the affective domain of learning.

## Reciprocal Effects Models

A reciprocal relationship is integral to CVT (Pekrun, 2006). The reciprocal relationship is the result of low perceived value of the outcome activity and/or low subjective levels of control over the content, resulting in low performance. Students continually appraise value and control throughout the course resulting in fluctuations of achievement emotions (Lüftenegger, Klug, Harrer, Spiel, & Schober, 2016; Pekrun et al., 2014; Putwain, Larkin, & Sander, 2013). At any point before and after an achievement task such as an exam, students reevaluate their level of control over the material in the course. The content may present itself to the student as too easy (i.e., they scored very highly on the exam) or too complex (i.e. the student achieved a low score). In either case, the time after an examination results in a change in emotions. The value attributed to the performance in the course may vary with the interpretation of achievement outcome scores.

We intend to explore the reciprocal relationship between boredom and academic achievement in a first-semester organic chemistry class (see Figure 5.1; Pekrun et al., 2014). Reciprocal effects relationships have been found in an organic chemistry course between performance and self-efficacy (Villafañe, Xu, & Raker, 2016).

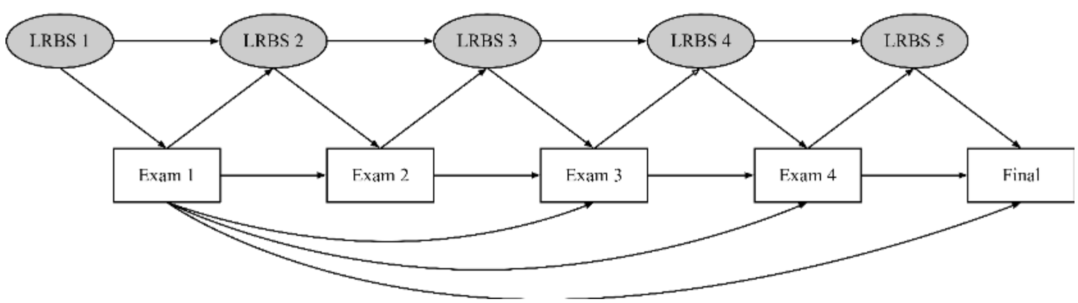


**Figure 5.1**. Model 1 – Hypothesized reciprocal effects model for LRBS and academic performance**.**

**Research Purpose and Goals**

Most research on boredom as an achievement emotion has focused on K-12 students and postsecondary classrooms outside of STEM fields. We, therefore, chose to study boredom in the context of postsecondary organic chemistry, specifically in the first semester of a yearlong course typically taken in the second year of postsecondary studies in the United States. Organic chemistry is a course taken by many students to fulfill entrance requirements for graduate and professional studies including health-related programs; the course is a prerequisite for upper-level courses in chemistry and molecular biology (Muller, 2013). Given the importance of success in the course for academic and career goals in a similar way to many postsecondary science courses, we hypothesize that students in the course have a unique context for perceiving the value of the course that may greatly influence achievement emotions (Weidinger, Steinmayr, & Spinath, 2017).

We have two research goals for this study: To gather evidence to establish that the learning-related boredom scale (LRBS) results in valid, reliable, and longitudinally invariant data with postsecondary organic chemistry students, and to test the reciprocal effects relationship between academic boredom and achievement as compared to direct effects relationships.

## Methods

### Participants and design

Participants were students (*n* = 656) enrolled in an organic chemistry course during Spring 2016 at a large public research-intensive university in the Southeast United States. The population of students in the course were 57% female, 52% white, 18% Hispanic, 13% Asian, 12% black, and 4% international.

Students completed the LRBS questionnaire within 72 hours prior to taking each exam (i.e., five total administrations). The questionnaire was administered online; an initial invitation and two reminder emails were sent to students and the instructors encouraged participation through announcements during lecture periods. The response rate for each administration ranged from 84% ($4^{th}$ administration) to 94% ($5^{th}$ administration). Data were collected in accordance with USF Institutional Review Board application Pro#00017861 approved on 18 June 2014.

### Measures

***Learning-related boredom scale (LRBS).*** The LRBS is a reduced subscale of the Achievement Emotions Questionnaire (AEQ) that has been found to result in valid and reliable data when measuring boredom in postsecondary education environments (Frenzel et al., 2007; Pekrun, 2006; Pekrun, Elliot, & Maier, 2009; Pekrun et al., 2010). In the development of the instrument, students were asked to describe their experiences of boredom in an open-ended questionnaire; their answers were coded to create the items used in the LRBS (Pekrun et al., 2010). The six LRBS items administered in this study are part of a previously used shorter version of the AEQ and have been modified by inserting "organic chemistry" where the original LRBS states "this course" (see Table 5.2). Pekrun et al. (2005) recommend that course-specific language be inserted into the item text. Adaptations of the scale to a specific environment, as done in our study, have shown to produce valid and reliable data (Ahmed et al., 2013; Tze, Klassen, Daniels, Li, & Zhang, 2013).

**Table 5.2**. Learning-related boredom scale.

| Item | not at all true | slightly true | moderately true | very true | completely true |
|---|---|---|---|---|---|
| When studying for organic chemistry, I feel bored. | 1 | 2 | 3 | 4 | 5 |
| The things I have to do for organic chemistry are often boring. | 1 | 2 | 3 | 4 | 5 |
| The content is so boring that I often find myself daydreaming. | 1 | 2 | 3 | 4 | 5 |
| When studying, my thoughts are everywhere else, except on the course material. | 1 | 2 | 3 | 4 | 5 |
| Often I am not motivated to invest effort in this boring course. | 1 | 2 | 3 | 4 | 5 |
| The material in this subject area is so boring that it makes me exhausted even to think about it. | 1 | 2 | 3 | 4 | 5 |

***Academic performance.*** Performance was measured using four term exams, prepared by the two course instructors, and a final exam. Term exams consisted of eleven to fourteen open-ended/free-response items; the instructors and a group of teaching assistants graded the exams following a predefined rubric. Due to slight deviations in the term exams between the two course instructors, z-scores were used for term exam measures. The final exam was the 70 multiple-choice 2014 First Term Organic Chemistry ACS Exam (2014); final exam scores are reported as z-scores for comparative purposes with the four term examinations.

## Data analysis

***Descriptive statistics.*** Descriptive analysis for the Learning-Related Boredom Scale (LRBS) and exam scores were performed in Stata 14 (2015). Means, standard deviations, tests of normality (skewness and kurtosis), and correlations were determined and evaluated for alignment with statistical procedure assumptions.

***LRBS psychometrics.*** Psychometrics, including reliability and internal structure validity analyses, were conducted as recommended by the *Standards for Educational and Psychological Testing* (2014) and in similarity with other studies using the LRBS (Peixoto, Mata, Monteiro, Sanches, & Pekrun, 2015). Reliability was measured using Cronbach's α (1951). When above .70, Cronbach's α indicates low standard error of measurement (Cortina, 1993).

Structural validity evidence based on internal structure was gathered via a confirmatory factor analysis approach (Pekrun, et al., 2010). Confirmatory factor analyses, invariance, and structural analyses were conducted through robust maximum likelihood estimation in MPlus Version 7.1.4 (Muthén & Muthén, 1998 - 2017). The confirmatory model had one continuous latent factor, boredom, indicated by six observed items (see Table 2). Per a recommendation by Pekrun et al. (2014), correlations were modeled between residual error terms for items representing the same domain: items 1 and 2 from the affective domain, items 3 and 4 from the cognitive domain, and items 5 and 6 from the physiological domain.

***Longitudinal measurement invariance.*** We are interested in knowing whether the construct of boredom is measured in the same way across time rather than between groups (e.g., sex or race/ethnicity groups; Dimitrov, 2010; Sass, 2011). Longitudinal measurement invariance is the appropriate measure of instrument functioning when a construct is measured at several points in time for one group (e.g., Widaman, Ferrer, & Conger, 2010; Widaman & Thompson, 2003). Longitudinal invariance is a multi-confirmatory factor analysis approach in which levels of increasing restrictions are placed on factor loadings, intercepts, and residual variances in the model (Brown, 2015). In tests for longitudinal invariance, configural models allow observed variable factor loadings, intercepts, and residual variances to vary at each instance. Strict models, whereby variable loadings are held constant at each time point along with intercepts and unique variances, are then compared with configural models to determine invariance (Widaman et al., 2010). If a strict model has acceptable fit and/or no significant changes in fit from the configural model occur, invariance across time can be considered acceptable.

***Model specification for the reciprocal effects model.*** Reciprocal effects models and the alternative models were evaluated via a structural equation modeling approach (Anderson & Evans, 1974; Levine & Donitsa-Schmidt, 1998; Pekrun et al., 2014). We first tested the hypothesized reciprocal effects model (see Figure 5.1). Then, alternate direct effect models were evaluated: Models 2 and 3 (see Figures 5.2 and 5.3) are limited to the effects of boredom on performance, and performance on boredom, respectively. In Model 4 (see Figure 5.4), there is no interaction between boredom and performance. Model 5 (see Figure 5.5) is identical to Model 1 without the "snowball" effect of performance.

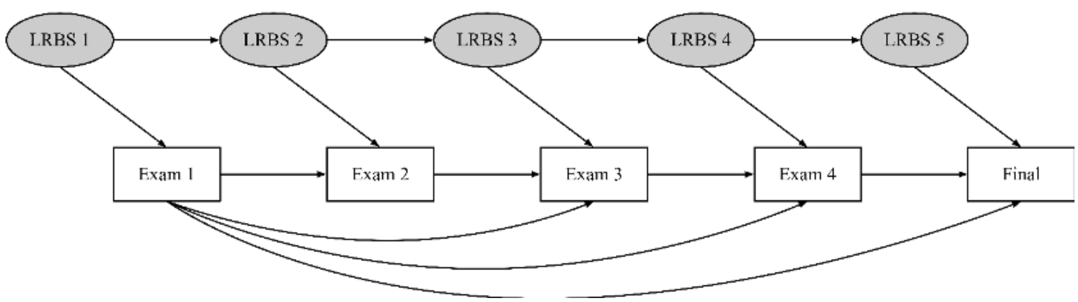


**Figure 5.2**. Model 2 –LRBS boredom effects model.

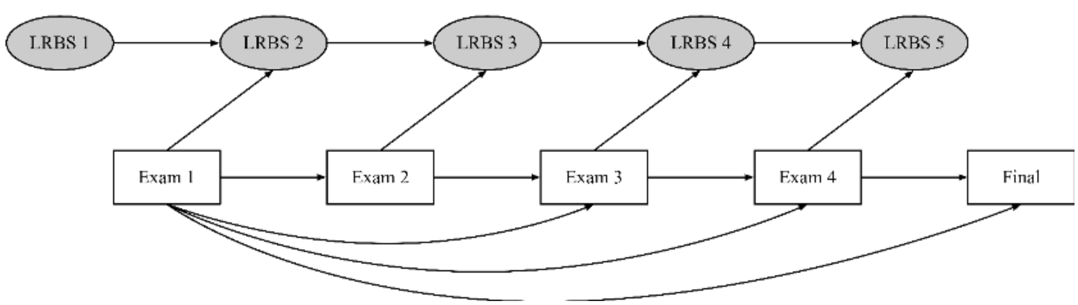


**Figure 5.3**. Model 3 –LRBS performance effects model.

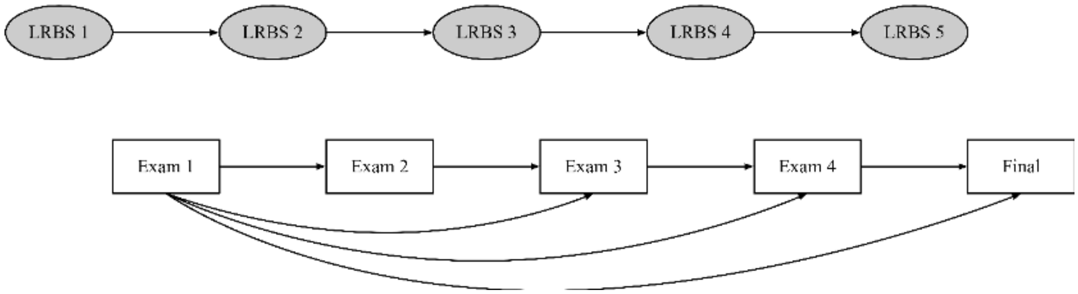


**Figure 5.4**. Model 4 –LRBS autoregressive model.

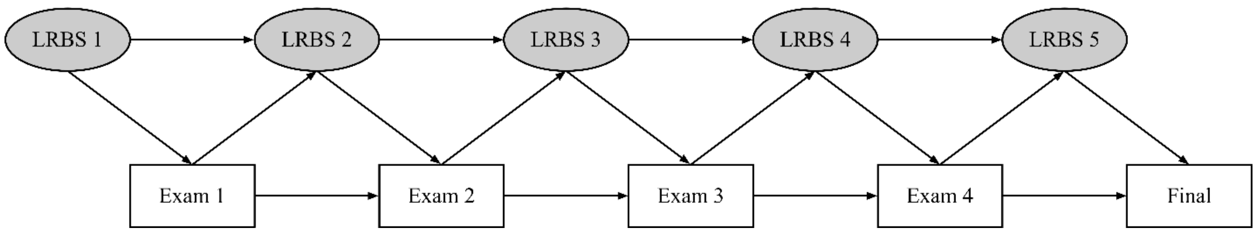


**Figure 5.5**. Model 5 –LRBS reciprocal model without “snowball” effect.

The postsecondary organic chemistry curriculum has been shown in a previous study to be a unique context for exploring affect/achievement relationships due a “snowball” correlation effect between the first exam and all subsequent exams (Villafañe et al., 2016). This correlation is hypothesized due to the cumulative nature of the content which builds in a consistent manner across the targeted course (Villafañe et al., 2016). The content covered in the first weeks of the course include nomenclature (i.e., naming compounds) and three-dimensional structure, topics essential for the understanding of chemical reactions in the later weeks of the term; it is our logical conclusion that an understanding of the information assessed on the first exam is vital for success on future course assessments.

To determine the best model, a scaled difference chi-square ($\chi^2$) test was used to compare the fitness of alternate models (Satorra & Bentler, 2001).

## Results

### Descriptive statistics

Descriptive statistics for each administration of the LRBS as well as for each examination are presented in Table 5.3. Mean LRBS scores (i.e., average responses to the items found in Table 5.3) ranged from 1.75 to 2.14 with gradual increases in boredom across the term. Univariate normality was evaluated using skewness and kurtosis; except for LRBS 1, skewness and kurtosis measures fall within acceptable ranges, with all values for skewness within the range of acceptable values calculated by Stata14: below 2 and above 5. The single high kurtosis value is of little concern for this analysis, as the LRBS scale is measured on a 5-point scale and therefore perfect normality is not expected (Leung, 2011).

**Table 5.3**. Descriptive statistics and Pearson correlations for study variables.

| | Learning-Related Boredom Scale | | | | | Exam 1 | Exam 2 | Exam 3 | Exam 4 | Final |
|---|---|---|---|---|---|---|---|---|---|---|
| | t1 | t2 | t3 | t4 | t5 | | | | | |
| *Mean* | 1.75 | 1.85 | 2.03 | 2.07 | 2.14 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| *SD* | 0.83 | 0.88 | 0.96 | 0.94 | 1.01 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Min | 1 | 1 | 1 | 1 | 1 | -3.67 | -3.70 | -2.11 | -2.20 | -2.49 |
| Max | 5 | 5 | 5 | 5 | 5 | 2.20 | 2.10 | 2.61 | 2.57 | 2.67 |
| Skewness | 1.54 | 1.30 | 0.95 | 0.78 | 0.74 | -0.36 | -0.76 | 0.37 | 0.19 | 0.24 |
| Kurtosis | 5.30 | 4.43 | 3.39 | 3.10 | 2.92 | 2.79 | 3.45 | 2.45 | 2.44 | 2.66 |
| t1 | 1.00 | | | | | | | | | |
| t2 | $.64^{***}$ | 1.00 | | | | | | | | |
| t3 | $.63^{***}$ | $.70^{***}$ | 1.00 | | | | | | | |
| t4 | $.57^{***}$ | $.72^{***}$ | $.70^{***}$ | 1.00 | | | | | | |
| t5 | $.52^{***}$ | $.62^{***}$ | $.66^{***}$ | $.73^{***}$ | 1.00 | | | | | |
| Exam 1 | $-.10^{*}$ | $-.17^{***}$ | $-.19^{***}$ | $-.19^{***}$ | $-.11^{**}$ | 1.00 | | | | |
| Exam 2 | -.08 | $-.16^{***}$ | $-.22^{***}$ | $-.15^{***}$ | $-.14^{**}$ | $.63^{***}$ | 1.00 | | | |
| Exam 3 | $-.13^{**}$ | $-.13^{**}$ | $-.20^{***}$ | $-.17^{***}$ | $-.16^{***}$ | $.62^{***}$ | $.60^{***}$ | 1.00 | | |
| Exam 4 | $-.14^{**}$ | $-.20^{**}$ | $-.21^{***}$ | $-.21^{***}$ | $-.18^{***}$ | $.62^{***}$ | $.66^{***}$ | $.74^{***}$ | 1.00 | |
| Final | $-.11^{*}$ | $-.17^{**}$ | $-.21^{***}$ | $-.21^{***}$ | $-.20^{***}$ | $.65^{***}$ | $.64^{***}$ | $.72^{***}$ | $.76^{***}$ | 1.00 |

*Note*. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

Pearson-product moment correlation coefficients were obtained for the study measures (see Table 5.3). All correlations are significant between study measures except for between LRBS 1 and Exam 2.

Given that LRBS1 was taken prior to any performance measure, a weak relationship between these two measures is appropriate. In addition, we note a significant correlation between other LRBS measures and Exam 2 including the LRBS measure directly preceding Exam 2; therefore, the non-significant relationship is inconsequential.

**LRBS psychometrics**

Observed levels of Cronbach's $\alpha$ for each LRBS administration indicate that the data collected by the instrument are reliable (see Table 5.4); Vodanovich and Watt (2016) observed similarly high Cronbach's α in a review of studies using the boredom subscale of the AEQ. Given similar findings with the LRBS instrument, our observed Cronbach's α values are considered acceptable (Ahmed et al., 2013; Pekrun et al., 2014; Tze et al., 2013).

**Table 5.4.** Cronbach's α for LRBS.

| Boredom | *N* | α | Lower (95% c.i.) | Upper (95% c.i.) |
|---|---|---|---|---|
| LRBS 1 | 567 | .94 | .93 | .94 |
| LRBS 2 | 548 | .95 | .94 | .96 |
| LRBS 3 | 533 | .95 | .95 | .96 |
| LRBS 4 | 492 | .95 | .95 | .96 |
| LRBS 5 | 549 | .97 | .96 | .97 |

*Note*. c.i. = confidence interval

Confirmatory factor analyses of the internal structure of the LRBS instrument at each administration had acceptable model fit (see Table 5.5). Measures of model fit include $\chi^2$, in which a statistically significant value indicates poor fit. Other fit measures that are less susceptible to error inflation due to sample size were used (i.e., CFI, TLI, RMSEA, and SRMR). Traditional cut-off values for CFI (comparative fit index) are ≥ .95; TLI (Tucker-Lewis index) are ≥ .95 with .90 as acceptable; RMSEA (root mean square standard error of approximation) are ≤ .05, good, .05-.08 reasonable, and .08-.10 mediocre;

and SRMR (standardized root mean square residual) are ≤ .08 (Hoyle & Panter, 1995; Hu & Bentler, 1999). RMSEA measures in models with small degrees of freedom like ours are not anticipated to have acceptable fit; therefore, emphasis is placed on CFI, TLI, and SRMR measures (Kenny, Kaniskan, & McCoach, 2014). At least two indexes were within acceptable cut offs for each administration of the LRBS (Hooper, Coughlan, & Mullen, 2008; Hu & Bentler, 1999).

**Table 5.5**. Confirmatory factor analyses of the LRBS at each administration.

| | *N* | $\chi^2$ (*df*= 6) | CFI | TLI | RMSEA | SRMR |
|---|---|---|---|---|---|---|
| LRBS 1 | 567 | 17.4 | .993 | .981ᵻ | .058 | .016ᵻ |
| LRBS 2 | 551 | 42.5* | .981ᵻ | .952ᵻ | .105 | .019ᵻ |
| LRBS 3 | 535 | 27.7 | .986ᵻ | .965ᵻ | .082 | .017ᵻ |
| LRBS 4 | 494 | 35.5* | .977ᵻ | .942 | .100 | .018ᵻ |
| LRBS 5 | 551 | 23.6 | .990ᵻ | .976ᵻ | .073 | .012ᵻ |

*Note.* * p < .00001, “ᵻ” indicates a value within the boundaries for good fit.

**Longitudinal measurement invariance**

We conducted a test of longitudinal measurement invariance (results in Table 5.6) to determine if the LRBS was measuring the construct consistently across the five administrations. Goodness-of-fit statistics are reported for both configural and strict longitudinal invariance models; the same standard of goodness-of-fit for CFA models are applied to invariance models. Similar goodness-of-fit statistics are found for the strict model as for the configural model.

**Table 5.6**. Longitudinal measurement invariance analysis of LRBS across five administrations.

| Model | $\chi^2$ | *df* | CFI | SRMR | RMSEA | ΔCFI |
|---|---|---|---|---|---|---|
| Configural | 1333* | 371 | .926ᵻ | .030ᵻ | .063 | .005 |
| Strict | 1457* | 435 | .921ᵻ | .034ᵻ | .060 | |

*Note.* * *p* < .0001, “ᵻ” indicates a value within the boundaries for good fit

The goodness-of-fit statistics indicate that the strict model still has good fit. An acceptable determination of invariant fit is a change in CFI less than .01, which is demonstrated in our analysis (Cheung & Rensvold, 2002). In addition, because of the general indication of acceptable fit (i.e., we would accept the strict model to represent the data as if we were not comparing two models), we will accept the strict model as representing acceptable longitudinal invariance (Brown, 2015; Widaman et al., 2010; Widaman & Thompson, 2003).

**Reciprocal effects model and alternate models**

We conducted structural equation models of our hypothesized reciprocal effects model with a snowball effect (Model 1, Figure 5.1), LRBS effects model (Model 2, Figure 5.2), performance effects model (Model 3, Figure 3), autoregressive model (Model 4, Figure 5.4), and reciprocal effects model without a snowball effect (Model 5, Figure 5.5). We report goodness-of-fit statistics for these models in Table 5.7.

**Table 5.7**. Model fit indices results for reciprocal causation model and alternate models.

| Model | $\chi^2$ | *df* | CFI | SRMR | RMSEA | S-B Factor | $\Delta df$ | $SB\Delta\chi^2$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1409* | 530 | .943 | .078 | .050 | 1.2328 | - | - |
| 2 | 1432* | 534 | .942 | .088 | .051 | 1.2315 | 4 | 25ᵻ |
| 3 | 1423* | 535 | .942 | .092 | .052 | 1.2301 | 5 | 14ᵻ |
| 4 | 1450* | 539 | .940 | .112 | .052 | 1.2288 | 9 | 44ᵻ |
| 5 | 1598* | 533 | .932 | .082 | .055 | 1.2330 | 3 | 184ᵻ |

*Note*. * $p < .001$, ᵻ = Fit significantly worse than the reciprocal model with $p < .05$

A statistically significant value for the change in $\chi^2$ by the Sartorra-Bentler adjustment in model complexity (Satorra, 2000; Satorra & Bentler, 2001) was observed for the difference between Model 1 and Models 2, 3, 4, and 5, respectively. These results suggest that Model 1 best fits the data. In addition, Model 1 has goodness-of-fit statistics that are superior to acceptable cutoffs. Therefore, the reciprocal model is tenable. Standardized estimates of the reciprocal model with a snowball effect are reported in Figure 5.6. In all cases, the relationship between exam performance and boredom at each time point is small, but negative.

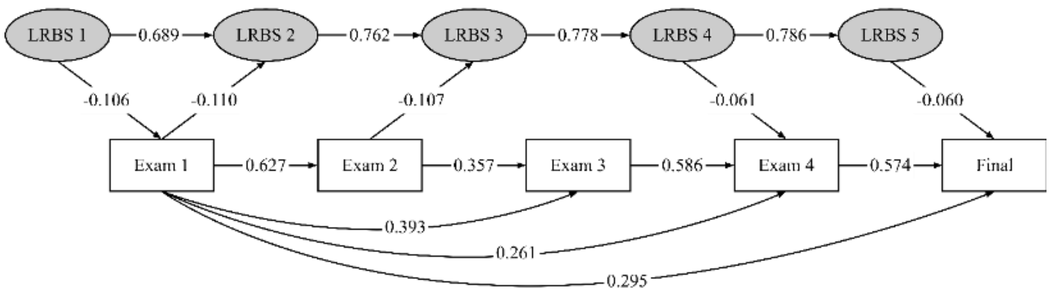


**Figure 5.6**. Standardized coefficients for reciprocal causation structural model of the relationship between LRBS and academic performance. All pathway coefficients are significant at $p < .05$.

**Discussion**

In support of achieving the first research goal set for this study, our analyses suggest that the learning-related boredom scale (LRBS) has acceptable psychometric and longitudinal measurement invariance properties in organic chemistry students. High values of Cronbach's α and strong evidence of internal structure of the LRBS mirror those found in other settings (Pekrun et al., 2014; Preckel, Götz, & Frenzel, 2010; Tze et al., 2013). Likewise, our conclusion of longitudinally invariant measures of boredom via the LRBS match prior results in other settings (Pekrun et al., 2014).

The empirical evidence described in this report support the achievement of the second research goal set for this study: a reciprocal relationship between boredom and achievement. The indication here that students who perform worse on academic performance outcomes will be more bored in subsequent classes supports the theoretical interpretation by Pekrun et al. (2010) and Acee et al. (2010) that boredom occurs as a negative, deactivating emotion for students regardless of prior knowledge levels. The evidence of a reciprocal relationship indicates that any opportunity to reduce boredom at any point in time will have a positive effect on achievement. The results of this study imply that an increase in boredom will reflect in lower scores on future academic performance outcomes, creating a situation in which escape is unlikely (Pekrun et al., 2010; Pekrun et al., 2014). Giving students an opportunity to increase control over the course content or reduce learning-related boredom by increasing value given to the content can help remove a barrier to academic success (Pekrun et al., 2010).

We conclude that through the lens of the control-value theory (Pekrun, 2006), learning-related boredom has a small but negative relationship with academic achievement across the first semester of a yearlong postsecondary organic chemistry course. This relationship is similar to that found in a study between boredom and academic achievement involving an introductory psychology course (Pekrun et al., 2014). A unique finding in this study is that the relationship between boredom and achievement is larger at the beginning of the semester, which demonstrates a key time for the implementation at intervention-type efforts to reduce boredom.

We cannot conclusively claim that boredom is caused by performance or vice versa; testing reciprocal effects models lends credence to the complex relationships between affect and achievement. A possible limitation of this study is the potentially multidimensional aspect of boredom; the LRBS is a six-item measure that is conducive to repeated measures across time, but is limited in the breadth and depth of the dimensional aspects of boredom. Multidimensionality has been detected in research on boredom in the academic environment (Acee et al., 2010; Tze et al., 2013); this dimensionality is inherent in affective constructs and should be considered while realizing the practical nature of the measures applied in this study.

**Implications for Research**

Future research inspired by this work will include the use of the LRBS to measure student boredom. Explicit efforts were not made in the studied course to decrease boredom; the work presented increases our capacity to evaluate future curricular and pedagogical reforms in postsecondary organic chemistry education. The psychometric evidence presented in support of the LRBS suggests that it is a viable instrument for evaluating organic chemistry curricula targeted at reducing boredom by introducing relevant and meaningful examples (e.g., Clayden, Greeves, & Warren, 2012; Doxsee, 1990; Ferguson, 1980; Harrison, 1989; Kelley & Gaither, 2007; Kolb & Kolb, 1981) or the incorporation of active learning strategies such as peer-led team learning, process-oriented guided inquiry learning, or problem-based learning. Reform implementation in the wild is typically deemed successful based on cognitive assessment performance outcomes rather than other important factors such as implementation strategy and other assessment measures (Holme et al., 2010; Stains & Vickery, 2016). The contribution of affective outcomes such as boredom can serve as a further indication of the effectiveness of new pedagogical methods. Measures such as the LRBS can assist in the evaluation of new instructional strategies, rather than relying on cognitive assessments alone, as noncognitive evaluation can empower researchers to a better understanding of achievement (Rhöneck & Grob, 1991; von Rhöneck, Grob, Schnaitmann, & Völker, 1998).

It is important to consider that our interpretation of boredom is from the context of organic chemistry, a course in which high value is assigned by students to the performance outcome (i.e., entrance to graduate and professional programs) and students approach with highly negative anticipation (Grove & Bretz, 2010; Muller, 2013). When students are presented with an environment in which low value is assigned to the achievement activities, the relationship between boredom and achievement will present itself differently. Therefore, the claims made in this report regarding the relationship between boredom and achievement should be contextualized in similar high-stakes environments and studied uniquely in others. The context also disguises the impact of students who meet the traditional interpretation of boredom, i.e., students who are under-challenged (Csikszentmihalyi, 1975). There are students in the course studied here and in many educational contexts who experience boredom but still achieve high scores on performance outcomes; CVT would posit that this is due to their high level of control over the content. These students, while present, are not the majority in postsecondary science classrooms because of their unique nature

(Grove & Bretz, 2010), and would need to be considered more strongly in an environment in which they are more likely the majority.

**Implications for Instruction**

For instructors, boredom may appear ubiquitous in the classroom. The results of this study indicate that bringing an understanding of student boredom to any classroom can benefit students and overall course grades. As noted by chemical education researchers, learning is an emotional phenomenon, and considering affect in instruction is vital to success (Chan & Bauer, 2016; Ferrell & Barbera, 2015; Nieswandt, 2007). Gauging the level of boredom in the classroom using the LRBS as described in this report can empower instructors to meet their students' needs in a meaningful way.

In addition to the ready-to-use method of incorporating the LRBS in to a classroom, curricular innovation is also supported by the results of this work. Incorporating simple innovations such as regular formative assessment, with feedback throughout the course, may enable students to gauge their control over the content and encourage them to seek resources to improve level of control prior to summative performance outcome assessments, even in large courses such as those typically found in organic chemistry (Black & William, 2010; Broadbent, Panadero, & Boud, 2017; Nicol & Macfarlane-Dick, 2007). Active learning pedagogies such as evidence-based instructional practices build value and control in the classroom through developing student autonomy (Gonzalez & Paoloni, 2015; Tien, Roth, & Kampmeier, 2002).

Our work indicates a direction for further research that can provide more insight into the student experience of postsecondary science. Emotions do not occur in a vacuum; additional studies on other emotions and contexts measured by the Achievement Emotions Questionnaire (Pekrun et al., 2011) would broaden our understanding of the way students experience chemistry courses. Instructional innovations incorporating efforts to reduce negative emotions such as boredom can enhance the aspects of the innovation designed to increase learning in the cognitive domain. Incorporating measures of a variety of these affective states can inform the process of developing and implementing instructional innovations (Osborne, Simon, & Collins, 2003).

**References**


Acee, T. W., Kim, H., Kim, H. J., Kim, J., Chu, H. N., Kim, M., . . . Wicker, F. W. (2010). The Boredom Research Group. Academic boredom in under- and over-challenging situations. *Contemporary Educational Psychology, 35*(1), pp. 17-27. doi:10.1016/j.cedpsych.2009.08.002

ACS Examinations Institute. (2014). *American Chemical Society Division of Chemical Education, First Term Organic Chemistry (Form 2014)* Ames, Iowa: ACS Exams.

Ahmed, W., van der Werf, G., Kuyper, H., & Minnaert, A. (2013). Emotions, self-regulated learning, and achievement in mathematics: a growth curve analysis. *Journal of Educational Psychology, 105*(1), pp. 160-161. doi:10.1037/a0030160

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing* Washington, D.C.: AERA Publication.

Anderson, J. G., & Evans, F. B. (1974). Causal models in education research: Recursive models. *American Educational Research Journal, 11*(1), pp. 29-39. doi:10.3102/00028312011001029

Aydin, Y. C., Uzuntiryaki, E., & Demirdogen, B. (2011). Interplay of motivational and cognitive strategies in predicting self-efficacy and anxiety. *Educational Psychology, 31*, pp. 55-66. doi:10.1080/01443410.2010.518561

Bench, S. W., & Lench, H. C. (2013). On the function of boredom. *Behavioral Sciences, 3*(3), pp. 459-472. doi:10.3390/bs3030459

Black, A. E., & Deci, E. L. (2000). The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: A self determination theory perspective. *Science Education, 84*, pp. 740-756. doi:10.1002/1098-237X(200011)84:6<740::AID-SCE4>3.0.CO;2-3

Black, P., & William, D. (2010). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 92*, pp. 81-90. doi:10.1177/003172171009200119

Broadbent, J., Panadero, E., & Boud, D. (2017). Implementing summative assessment with a formative flavour: A case study in a large class. *Assessment & Evaluation in Higher Education*, pp. 1-16. doi:10.1080/02602938.2017.1343455

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research* (2nd ed.) New York: Guilford Press.

Chan, J. Y. K., & Bauer, C. F. (2016). Learning and studying strategies used by general chemistry students with different affective characteristics. *Chemistry Education Research and Practice, 17*, pp. 675-684. doi:10.1039/C5RP00205B

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, pp. 233-255. doi:10.1207/S15328007SEM0902_5

Clayden, J., Greeves, N., & Warren, S. (2012). *Organic Chemistry (2nd Ed.)*: Oxford.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*(1), pp. 98-104. doi:10.1037/0021-9010.78.1.98

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), pp. 297-334. doi:10.1007/BF02310555

Csikszentmihalyi, M. (1975). *Beyond Boredom and Anxiety* San Francisco, CA: Jossey-Bass.

Daniels, L. M., Stupnisky, R. H., Pekrun, R., Haynes, T. L., Perry, R. P., & Newall, N. E. (2009). A longitudinal analysis of achievement goals: From affective antecedents to emotional effects and achievement outcomes. *Journal of Educational Psychology, 101*(4), pp. 948-963. doi:10.1037/a0016096

Daschmann, E. C., Goetz, T., & Stupnisky, R. H. (2011). Testing the predictors of boredom at schools: Development and validation of the precursors to boredom scales. *British Journal of Educational Psychology, 81*, pp. 421-440. doi:10.1348/000709910X526038

Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development, 43*(2), pp. 121-149. doi:10.1177/0748175610373459

Doxsee, K. M. (1990). Development of an advanced synthesis laboratory course in Nobel prize winning chemistry. *Journal of Chemical Education, 67*(12), pp. 1057-1060. doi:10.1021/ed067p1057

Eastwood, J. D., Frischen, A., Fenske, M. J., & Smilek, D. (2012). The unengaged mind: Defining boredom in terms of attention. *Perspectives on Psychological Science, 7*(5), pp. 482-495. doi:10.1177/1745691612456044

Eddy, R. M. (2000). Chemophobia in the college classroom: Extent, courses, and student characteristics *Journal of Chemical Education, 77*, pp. 514-517. doi:10.1021/ed077p514

Fahlman, S. S., Mercer-Lynn, K. B., Flora, D. B., & Eastwood, J. D. (2013). Development and validation of the multidimensional state boredom scale. *Assessment, 20*, pp. 68-85. doi:10.1177/1073191111421303

Farmer, R., & Sundberg, N. D. (1986). Boredom proneness- the development and correlates of a new scale. *Journal of Personality Assessment, 50*(1), pp. 4-17. doi:10.1207/s15327752jpa5001_2

Ferguson, L. N. (1980). Content and structure of the chemistry course: New approaches. *Journal of Chemical Education, 57*(1), pp. 46-48. doi:10.1021/ed057p46

Ferrell, B., & Barbera, J. (2015). Analysis of students' self-efficacy, interest, and effort beliefs in general chemistry. *Chemistry Education Research and Practice, 16*(2), pp. 318-337. doi:10.1039/C4RP00152D

Fisher, C. D. (1993). Boredom at work: A neglected concept. *Human Relations, 46*(3), pp. 395-417. doi:10.1177/001872679304600305

Forbes, D. C. (2004). Incorporation of medicinal chemistry into the organic chemistry curriculum. *Journal of Chemical Education, 81*(7), pp. 975-976. doi:10.1021/ed081p975

Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Perceived learning environment and students' emotional experiences: A multilevel analysis of mathematics classrooms. *Learning and Instruction, 17*(5), pp. 478-493. doi:10.1016/j.learninstruc.2007.09.001

Frenzel, A. C., Thrash, T. M., Pekrun, R., & Goetz, T. (2007). Achievement emotions in Germany and China: A cross-cultural validation of the Academic Emotions Questionnaire—Mathematics. *Journal of Cross-Cultural Psychology, 38*(3), pp. 302-309. doi:10.1177/0022022107300276

Galloway, K. R., & Bretz, S. L. (2016). Video episodies and action cameras in the undergraduate chemistry laboraotory: eliciting student perceptions of meaningful learning. *Chemistry Education Research and Practice, 17*(1), pp. 139-155. doi:10.1039/C5RP00196J

Goetz, T., Frenzel, A. C., Pekrun, R., & Hall, N. C. (2006). The domain specificity of academic emotional experiences. *Journal of Experimental Education, 75*(1), pp. 5-29. doi:10.3200/JEXE.75.1.5-29

Goetz, T., Pekrun, R., Hall, N., & Haag, L. (2006). Academic emotions from a social- cognitive perspective: Antecedents and domain specificity of students' affect in the context of Latin instruction. *British Journal of Educational Psychology, 76*(2), pp. 289-308. doi:10.1348/000709905X42860

Gonzalez, A., & Paoloni, P. (2015). Perceived autonomy-support, expectancy, value, metacognitive strategies and performance in chemistry: A structural equation model in undergraduates *Chemical Education Research and Practice, 16*, pp. 640-653. doi:10.1039/C5RP00058K

Grove, N. P., & Bretz, S. L. (2010). Perry's scheme of intellectual and epistemological development as a framework for describing difficulties in learning organic chemistry. *Chemistry Education Research and Practice, 11*, pp. 207-211. doi:10.1039/C005469K

Harackieqicz, J. M., & Hulleman, C. S. (2010). The importance of interest: The role of achievement goals and task values in promoting the development of interest. *Soc. & Personal. Psych. Compass. , 4*(1), pp. 42-52. doi:10.1111/j.1751-9004.2009.00207.x

Harrison, A. M. (1989). Medicinal chemistry/pharmacology in sophomore organic chemistry. *Journal of Chemical Education, 1989*(66), p 10. doi:10.1021/ed066p825

Holme, T., Bretz, S. L., Cooper, M., Lewis, J. E., Paek, P., Pienta, N., . . . Towns, M. (2010). Enhancing the role of assessment in curriculum reform in chemistry. *Chemical Education Research and Practice, 11*(2), pp. 92-97. doi:10.1039/C005352J

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Eletronic Journal of Business Research Methods, 6*(1), pp. 53-60.

Hoyle, R. H., & Panter, A. T. (1995). Writing about structural equation models. In R. H. Hoyle (Ed.), *Structural Equation Modeling: Concepts, Issues, and Applications*. Newbury Park, C.A.: Sage

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: Multidisciplinary Journal, 6*(1), pp. 1-55. doi:10.1080/10705519909540118

Kelley, C., & Gaither, K. K. (2007). Integrating pharmacology into the organic chemistry course: Understanding the synergy of biology and chemistry. *Journal of College Science Teaching, 30*(7), pp. 450-453.

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2014). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research, 44*(3), pp. 1-22. doi:10.1177/0049124114543236

King, D. E., Lewis, J. E., Anderson, K., Latch, D., Moog, R., Sutheimer, S., & Wester, G. (2015). Choosing appropriate models- Incorporating climate change into general chemistry. In K. C. Lanigan, E. S. Roberts-Kirchoff, K. R. Evans, M. A. Benvenuto & A. Rihana-Abdallah (Eds.), *Chemistry and the Environment: Pedagogical Methods and Tools* (Vol. 1214, pp. 1-15): American Chemical Society.

Kolb, K. E., & Kolb, D. K. (1981). Treatment of industrial organic chemistry in current textbooks. *Journal of Chemical Education, 58*(7), pp. 557-558. doi:10.1021/ed058p557

Laukenmann, M., Bleicher, M., Fuß, S., Gläser-Zikuda, M., Mayring, P., & von Rhöneck, C. (2003). An investigation of the influence of emotional factors on learning in physics instruction. *International Journal of Science Education, 25*(4), pp. 489-507. doi:10.1080/09500690210163233 Retrieved from http://dx.doi.org/10.1080/09500690210163233

Leung, S. (2011). A comparison of psyhometric properties and normality in 4-, 5-, 6-, and 11-point Likert scales *Journal of Social Service Research, 37*(4), pp. 412-421. doi:10.1080/01488376.2011.580697

Levine, T., & Donitsa-Schmidt, S. (1998). Computer use, confidence, attitudes, and knowledge: A causal analysis. *Computers in Human Behavior, 14*(1), pp. 125-146. doi:10.1016/S0747-5632(97)00036-8

Lüftenegger, M., Klug, J., Harrer, K., Spiel, C., & Schober, B. (2016). Students' achievement goals, learning-related emotions and academic achievement. *Frontiers in Psychology, 7*, pp. 1-10. doi:10.3389/fpsyg.2016.00603

Lynch, D. J., & Trujillo, H. (2011). Motivational beliefs and learning strategies in organic chemistry. *International Journal of Science and Mathematics Education, 9*, pp. 1351-1365. doi:10.1007/s10763-101-9264-x

Mann, S., & Robinson, A. (2009). Boredom in the lecture theatre: An investigation into the contributors, moderators and outcomes of boredom amongst university students. *British Journal of Educational Psychology, 35*(2), pp. 243-258. doi:10.1080/01411920802042911

Mercer-Lynn, K. B., Flora, D. B., Fahlman, S. A., & Eastwood, J. D. (2011). The measurement of boredom: Differences between existing self-report scales. *Assessment, 20*(5), pp. 585-596. doi:10.1177/1073191111408229

Mikulas, W. L., & Vodanovich, S. J. (1993). The essence of boredom. *The Psychologial Record, 43*(1), pp. 3-12.

Muller, D. (2013). Reforming premedical education- out with the old, in with the new. *New England Journal of Medicine, 368*, pp. 1567-1569. doi:10.1056/NEJMp1302259

Muthén, L. K. & Muthén, B. D. (1998-2017). *Mplus users guide. Seventh edition.* Los Angeles, CA: Muthén & Muthén.

Nett, U. E., Goetz, T., & Daniels, L. M. (2010). What to do when feeling bored: Student strategies for coping with boredom. *Learning and Individual Differences, 20*(6), pp. 626-638. doi:10.1016/j.lindif.2010.09.004

Nicol, D. J., & Macfarlane-Dick, D. (2007). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), pp. 199-218. doi:10.1080/03075070600572090

Niculescu, A. C., Tempelaar, D., Dailey-Herbert, A., Segers, M., & Gijselaers, W. (2015). Exploring the antecedents of learning-related emotions and their relations with achievement outcomes. *Frontline Learning Research, 3*(1), pp. 1-17. doi:10.14786/flr.v%25vi%25i.136

Nieswandt, M. (2007). Student affect and conceptual understanding in learning chemistry. *Journal of Research in Science Teaching, 44*, pp. 908-937. doi:10.1002/tea.20169

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education, 25*(9), pp. 1049-1079. doi:10.1080/0950069032000032199 Retrieved from http://dx.doi.org/10.1080/0950069032000032199

Peixoto, F., Mata, L., Monteiro, V., Sanches, C., & Pekrun, R. (2015). The Achievement Emotions Questionnaire: Validation for pre-adolescent students. *European Journal of Developmental Psychology, 12*(4), pp. 472-481. doi:10.1080/17405629.2015.1040757

Pekrun, R. (1992). The impact of emotions on learning and achievement: Towards a theory of cognitive/motivational mediators. *Applied Psychology, 41*(4), pp. 359-376. doi:10.1111/j.1464-0597.1992.tb00712.x

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), pp. 315-341. doi:10.1007/s10648-006-9029-9

Pekrun, R., Elliot, A. J., & Maier, M. A. (2009). Achievement goals and achievement emotions: Testing a model of their joint relations with academic performance. *Journal of Educational Psychology, 101*(1), pp. 115-135. doi:10.1037/a0013383

Pekrun, R., Goetz, T., Daniels, L. M., Stupnisky, R. H., & Perry, R. P. (2010). Boredom in achievement settings: Exploring control–value antecedents and performance outcomes of a neglected emotion. *Journal of Educational Psychology, 102*(3), pp. 531-549. doi:10.1037/a0019243

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology, 36*(1), pp. 36-48. doi:10.1016/j.cedpsych.2010.10.002

Pekrun, R., Goetz, T., & Perry, R. P. (2005). *Achievement Emotions Questionnaire (AEQ) - User's Manual*.

Pekrun, R., Hall, N. C., Goetz, T., & Perry, R. P. (2014). Boredom and academic achievement: Testing a model of reciprocal causation. *Journal of Educational Psychology, 106*(3), pp. 696-710. doi:10.1037/a0036006

Pekrun, R., & Perry, R. P. (2014). Control-value theory of achievement emotions. In R. Pekrun & L. Linnenrink-Garcia (Eds.), *International Handbook of Emotions in Education* (pp. 120-141). London: Routledge.

Perkins, K. K., Adams, W. K., Pollock, S. J., Finklestein, N. D., & Wieman, C. E. (2005). Correlating student beliefs with student learning using the Colorado Learning Attitudes about Science Survey. *AIP Conference Proceedings, 61*, p CP790. doi:dx.doi.org/10.1063/1.2084701

Perry, R. P., Hladkyj, S., Pekrun, R., & Pelletier, S. T. (2001). Academic control and action control in the achievement of college students: A longitudinal fild study. *Journal of Educational Psychology, 93*(4), pp. 776-789. doi:10.1037/0022-0663.93.4.776

Preckel, F., Götz, T., & Frenzel, A. (2010). Ability grouping on gifted students: Effects on academic self-concept and boredom. *British Journal of Educational Psychology, 80*(3), pp. 451-472. doi:10.1348/000709909X480716

Putwain, D. W., Larkin, D., & Sander, P. (2013). A reciprocal model of achievement goals and learning related emotions in the first year of undergraduate study. *Contemporary Educational Psychology, 38*(4), pp. 361-374. doi:10.1016/j.cedpsych.2013.07.003

Rhöneck‡, C. v., & Grob, K. (1991). Psychological aspects of learning about basic electricity in rural and urban classes. *International Journal of Science Education, 13*(1), pp. 87-95. doi:10.1080/0950069910130108 Retrieved from http://dx.doi.org/10.1080/0950069910130108

Robinson, W. P. (1975). Boredom at school. *British Journal of Educational Psychology, 45*(2), pp. 141-152. doi:10.1111/j.2044-8279.1975.tb03239.x

Sass, D. A. (2011). Testing measurement invariance and comparing latent factor means within a confirmatory factor analysis framework. *Journal of Psychoeducational Assessment, 29*(4), pp. 347-363. doi:10.1177/0734282911406661

Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In R. D. H. Heijmans, D. S. G. Pollock & A. Satorra (Eds.), *Innovations in Multivariate Statistical Analysis*. London: Kluwer Academic Publishers.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*(4), pp. 507-514. doi:10.1007/BF02296192

Singh, K., Granville, M., & Dika, S. (2002). Mathematics and science achievement: effecs of motivation, interest, and academic engagement. *Journal of Educational Research, 95*(6), pp. 323-332. doi:10.1080/00220670209596607

Stains, M., & Vickery, T. (2016). Fidelity of implementation: An overlooked yet critial construct to establish effectiveness of evidence-based instructional practices. *CBE-Life sciences education, 16*(1), pp. 1-11. doi:10.1187/cbe.16-03-0113

StataCorp. (2015). *Stata Statistical Software: Release 14* College Station, TX: StataCorp LLC.

Supalo, C. A., Humphrey, J. R., Mallouk, T. E., Wohlers, H. D., & Carlsen, W. S. (2016). Examining the use of adaptive technologies to increase the ahnds-on participation of students with blindness or low vision in secondary-school chemistry and physics. *Chemistry Education Research and Practice, 17*(4), pp. 1174-1189. doi:10.1039/C6RP00141F

Tanaka, A., & Murayama, K. (2014). Within-person analyses of situational interest and boredom: Interactions between task-specific perceptions and achievement goals. *Journal of Educational Psychology, 106*(4), pp. 1122-1134. doi:10.1037/a0036659

Tien, L. T., Roth, V., & Kampmeier, J. A. (2002). Implementation of a peer-led team learning instructional approach in an undergraduate organic chemistry course. *Journal of Research in Science Teaching, 39*, pp. 606-632. doi:10.1002/tea.10038

Trehan, I. R., Brar, J. S., Arora, A. K., & Kad, G. L. (1997). Fries rearrangement accelerated by microwave radiation in the undergraduate organic laboratory. *Journal of Chemical Education, 74*(3), p 324. doi:10.1021/ed074p324

Tze, V. M., Daniels, L. M., & Klassen, R. M. (2013). Examining the factor structure and validity of the English Precursors to Boredom Scales. *Learning and Individual Differences, 32*, pp. 254-260. doi:10.1016/j.lindif.2014.03.018

Tze, V. M., Daniels, L. M., & Klassen, R. M. (2016). Evaluating the relationship between boredom and academic outcomes: A meta-analysis. *Educational Psychology Review, 28*(1), pp. 119-144. doi:10.1007/s10648-015-9301-y

Tze, V. M., Klassen, R. M., Daniels, L. M., Li, J. C., & Zhang, X. (2013). A cross-cultural validation of the Learning-related Boredom Scale (LRBS) with Canadian and Chinese college students. *Journal of Psychoeducational Assessment, 31*(1), pp. 29-40. doi:10.1177/0734282912443670

Villafañe, S. M., Xu, X., & Raker, J. R. (2016). Self-efficacy and academic performance in first-semester organic chemistry: Testing a model of reciprocal causation. *Chemistry Education Research and Practice, 17*(4), pp. 973-984. doi:10.1039/C6RP00119J

Vodanovich, S. J., & Watt, J. D. (2016). Self-report measures of boredom: An updated review of the literature. *Journal of Psychology: Interdisciplinary and Applied, 150*(2), pp. 196-228. doi:10.1080/00223980.2015.1074531

von Rhöneck, C., Grob, K., Schnaitmann, G. W., & Völker, B. (1998). Learning in basic electricity: how do motivation, cognitive and classroom climate factors influence achievement in physics? *International Journal of Science Education, 20*(5), pp. 551-565. doi:10.1080/0950069980200504 Retrieved from http://dx.doi.org/10.1080/0950069980200504

Wallace, J. C., Vodanovich, S. J., & Restino, B. M. (2003). Predicting cognitive failures from boredom proneness and daytime sleepiness scores: An investigation within military and undergraduate samples. *Personality and Individual Differences, 34*(4), pp. 635-644. doi:0.1016/S0191-8869(02)00050-8

Watt, J. D. (1991). Effect of boredom proneness on time perception. *Psychological Reports, 69*, pp. 323-327. doi:10.2466/pr0.1991.69.1.323

Watt, J. D., & Hargis, M. B. (2010). Boredom proneness: Its relationship with subjective underemployment, perceived organizational support, and job performance. *Journal of Business and Psychology, 25*(1), pp. 163-174. doi:10.1007/s10869-009-9138-9

Weidinger, A. F., Steinmayr, R., & Spinath, B. (2017). Math grades and intrinsic motivation in elementary school: A longitudinal investigation of their association. *British Journal of Educational Psychology, 87*, pp. 187-204. doi:10.1111/bjep.12143

Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives, 4*(1), pp. 10-18. doi:10.1111/j.1750-8606.2009.00110.x

Widaman, K. F., & Thompson, J. S. (2003). On specifying the null model for incremental fit indices in structural equation modeling. *Psychological Methods, 8*(1), pp. 16-37. doi:10.1037/1082-989X.8.1.16

Wong, M. M., & Csikszentmihalyi, M. (1991). Motivation and academic achievement: The effects of personality traits and the quality of experience. *Journal of Personality, 59*(3), pp. 539-574. doi:10.1111/j.1467-6494.1991.tb00259.x

Zuckerman, M. (1979). *Sensation seeking: Beyond the optimal level of arousal* Hillsdale, NJ: Erlbaum.

# CHAPTER SIX:

# CONCLUSION

The research included in Chapters Two and Three of this dissertation demonstrates a step towards understanding of instructional practices across the United States and the research included in Chapters Four and Five demonstrates a nuanced understating of affect across a postsecondary organic chemistry course. Improvements in methodology in survey research on a national scale and research on affect within the classroom demonstrate efforts at achieving the goals of CER by characterizing enacted practices for understanding dissemination of EBIPs and embracing the nuanced development of affect. All four studies use quantitative methodology, seeking to measure the amount and distribution of certain characteristics present in the population. These studies maintain a high level of both precision and accuracy by utilizing self-report data collection mechanisms and incorporating statistical procedures which enable measurement of instrument validity. The central tenet of the studies included in this dissertation is improvement on previous research methodology – by utilizing a stratified sampling procedure and weighted data analysis in the case of the national survey of postsecondary chemistry faculty in Chapters Two and Three, and a longitudinal panel study design with longitudinal measurement invariance testing in the case of studies of affect in postsecondary organic chemistry in Chapters Four and Five.

Both levels of generalizability in the reports included in this dissertation seek to achieve a similar goal of understanding chemical education within their respective contexts. The studies using the national survey of postsecondary chemistry faculty as their data collection mechanism provide a snapshot of the state of use of a variety of instructional strategies and provide empirical support for a theoretical model in Chapter Two and evaluate classroom response systems in detail in Chapter Three. The studies of the effect of affect in the classroom provide a more nuanced snapshot of what the student experience might look like and similarly provide further empirical support for a theoretical model through exploring anxiety and enjoyment in Chapter Four and boredom in Chapter Five.

### Summary of Survey Research Results

Using a national survey of postsecondary chemistry faculty members, we learn more about the nature of chemical education as it exists in the "real world" in Chapters Two and Three of this dissertation. This perspective is sometimes overlooked, primarily because research products such as EBIPs are developed and tested within the CER community before being more widely distributed. Using the TCSR model of educational change (Woodbury & Gess-Newsome, 2002) as a framework, the study included in Chapter Two provided empirical evidence for the link between faculty members' beliefs about teaching and learning and self-efficacy and their instructional activities in the classroom. This corroborates claims made in the TCSR model and serves as a benchmark upon which efforts to change faculty member belief systems can be based. Despite an association between beliefs that students lean best through student-centered activities and the use of student-centered instructional strategies, groups of faculty members who utilize both student- and teacher- centered instructional strategies have similar levels of belief that students learn best through teacher-centered instructional strategies. This finding indicates that there is more nuance to the selection of instructional strategies than simply faculty members' beliefs.

The study included in Chapter Three analyzed the data from the national survey through the lens of one educational tool rather than the general state of instruction. In the case of CRSs, also known as clickers, the context of the course itself demonstrates the strongest association with classroom use. This study found that CRSs are used in large classrooms (≥ 55 students) and in courses which are taught at the introductory and foundational level (as defined by the Committee on Professional Training, 2015). This finding demonstrates a specific example of what is indicated by the results of the study on the association between beliefs and practice; even if a faculty member has a belief structure which is not aligned with their used instructional strategies, they might be influenced by the context of the course. The TALC framework (Rogers, 1995) allows us to explore the use of tools such as CRSs across the population and identify a niche environment in which the use of such tools flourish. This framework can lend itself to the study of particular EBIPs in the future.

**Summary of Affective Survey Instrument Results**

The studies included Chapters Four and Five are framed by CVT (Pekrun, 2000). The strength of CVT comes from its establishment in the theoretical literature and the support it has gained from empirical studies (e.g., Linnenbrink-Garcia & Pekrun, 2011; Pekrun, 2006; Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011; Pekrun, Goetz, Titz, & Perry, 2002). The studies in Chapters Four and Five increase the amount of empirical evidence supporting the propositions of the CVT: the relationships between achievement emotions and academic achievement are causal and reciprocal. In postsecondary organic chemistry, these links are supported as the fit of the SEM procedure demonstrates that reciprocal relationships are better fitting than the alternative models tested. These works explore the operation of the relationship between affect and achievement within one institution, with the goal of supporting theoretical distinctions with empirical data and seeking a deeper understanding of the way that students experience their postsecondary chemistry classrooms.

In addition to reinforcing previous findings, the studies in Chapters Four and Five also seek to generate new knowledge by interpreting the parameters estimated from the SEM procedure. The strength of relationships between initial anxiety and achievement are higher than those between initial enjoyment and achievement; however, achievement measures influence future enjoyment more than future anxiety. The standardized effects of initial boredom on achievement are lower than those from initial anxiety to achievement, but higher than from initial enjoyment to achievement. The large influence from achievement measures to subsequent measures of enjoyment is important to note, because positive emotions such as enjoyment are frequently ignored in the empirical literature on the relationship between affect and achievement in educational research (Fredrickson, 2001) and the stronger relationship between achievement and enjoyment indicates an area for possible instructional improvement.

In Chapter Five, the study of the relationship between boredom and achievement through the entire semester and in Chapter Four, the study of the relationship between anxiety and enjoyment and achievement on the first three in-term examinations (out of four in-term examinations and a final examination), a common finding is of particular note: the size of relationships between affect and achievement are largest at the beginning of the term, during the first and second in-term examinations. However, the "snowball" effect from the first in-term examination to all subsequent examinations, demonstrated in both studies included here as well as in Villafañe, Xu, and Raker (2016) may account for the loss in strength of relationships from affect to achievement, as their impact on the first in-term examination is residual. The nuance and increased understanding described here demonstrates information which would have been lost if we had utilized a cross-sectional design without longitudinal measurement, and only explored the relationship of affect with final course achievement.

**Implications for Research**

The most important implications of all four studies included in this dissertation are related to future research. As established techniques are critically evaluated and improved upon with the collection of new data and the development of new theoretical interpretations, we can incorporate these techniques in to our own work and increase the quality and quantity of CER.

This dissertation models a structure for future research to emphasize the collection of snapshots of the current state of the distribution of the products of CER, such as EBIPs and educational tools like CRSs. This constitutes basic research in the field of CER. By exploring the use of specific tools or practices, the techniques used for dissemination can be improved. CER innovators who develop EBIPs and other tools should take note of the niche fields demonstrated by the national survey results reported in Chapter Three. By focusing research on development and dissemination to these niche fields, researchers can provide the optimal experience for the most likely consumers of the material. For those interested in expanding the reach of their research projects, an understanding of the niche market for their tools can allow them to explore possible limitations for use in other markets and remove potential barriers to widespread use.

Support provided in this dissertation for the TCSR model of educational change, particularly the link between faculty beliefs about teaching and learning and self-efficacy demonstrated in Chapter Two, demonstrates a rich area for future research. The state of faculty members' beliefs was captured here, but the antecedents and ultimate effects of these beliefs is a new area which should be explored. If it can be determined that changing these beliefs can cause faculty members to effectively implement change initiatives, CER may see a significant growth in the distribution of EBIPs and other educational tools. Such a strand of research will require an exploration in to the interaction of beliefs and practices within the context of chemistry including from the time of graduate student instructional training (i.e., through the process of graduate teaching assistantship) through professorship.

The studies included in Chapters Two, Four, and Five demonstrate that the application of testing for internal structure and relations to other variables validity is a reasonable expectation for well-designed CER studies. The techniques provided in this dissertation should frame similar work in the future. This is important because as we see increased use of affective surveys in the interest of determining the effectiveness of EBIPs in the literature, the proper evaluation of these tools prior to their widespread use is essential. If an instrument is widely used despite a lack of acceptable evidence for the validity of its data, the information found in the literature base will be less reliable for making decisions within a classroom or on a policy level.

Longitudinal studies are becoming more prevalent and are better for identifying causal links across a course or program, as demonstrated in Chapters Four and Five. Longitudinal panel data collection mechanisms allow for greater interpretability of scores and the determination of differences between groups (Villafañe, Garcia, & Lewis, 2014). These studies also demonstrate how to conduct longitudinal measurement invariance testing, novel to the field of CER, that future studies should emulate in the production of high-quality research.

**Implications for Instruction**

The TCSR framework provides a means for thinking about and encouraging reflective practice. Reflective practice is recognized as the activities in which instructions partake after instruction, considering which aspects of the classroom practice were effective and which were not (see Kane, Sandretto, & Heath, 2004 for more on reflective practice). Because the study in Chapter Two demonstrates a link between beliefs and enacted practice, reflecting on how an individual's beliefs influence the way they teach can help foster change in instructional strategy in order to increase alignment between the two. Classroom change is known to be preempted by faculty member dissatisfaction with current instruction (Bauer, Libby, Scharberg, & Reider, 2013); through reflection, instructors can cultivate their own impression of satisfaction or dissatisfaction with the current state of instruction and its alignment with their belief systems.

The other study utilizing the national survey of faculty members focuses on CRSs and their niche market. Instructors are often encouraged to consider the adoption of new tools in the classroom, such as CRSs. Adopting such tools, however, is challenging for many reasons, described in Chapter Three. The contextual information that the study in Chapter Three provides enables faculty members to consider the context in which CRSs are more frequently used. These faculty members can compare this context (i.e., large courses taught at the lower level) their own classroom to provide evidence for the potential for increased or decreased fit of the use of CRSs.

The studies included in Chapters Four and Five, which explore the effect of affect on achievement in postsecondary chemistry classrooms, provide empirical evidence to support theoretical links and demonstrate a trend which can be explored in other contexts and acted upon accordingly. It has been recognized that the classroom climate is set by the instructor and can dictate the experiences that students undergo (e.g., Larsen, 1986). The findings in this dissertation corroborate those from other researchers in other contexts that emotional experiences are directly linked to student achievement (e.g., Daniels et al., 2009; Pekrun, Elliot, & Maier, 2009; Pekrun, Hall, Goetz, & Perry, 2014; Xu, Villafañe, & Lewis, 2013). The findings also support that a singular focus on the removal of negative emotions on the classroom would limit an instructor from reaching the potential for improved achievement through enhancing positive emotional states, such as enjoyment. This is a positive note for instructors who seek to improve their classroom climate; attempts at increasing positive affect in the classroom such as contextualized instruction are supported by this study. These attempts to improve positive emotions in the classroom should be associated with efforts at decreasing deactivating emotions like boredom in order to further encourage increased achievement. Instructors should note that setting the tone and climate of the classroom early in the semester can produce the largest effect regarding change in affect and increased achievement.

CVT posits that the achievement emotions are in direct relation to academic achievement, theoretical links that are supported by the evidence provided by Chapters Four and Five of this dissertation. Per CVT, the control that the student feels over the content as well as their learning environment and the value that they place on the course are the antecedents to the achievement emotions (Pekrun, 2000). Control and value, then, dictate the emotional experience which students undergo, therefore indirectly influencing achievement. This provides an additional venue for instructors to influence achievement by focusing on non-cognitive factors. This is recommended based on the demonstration in Chapters Four and Five that students re-evaluate their situation before and after achievement measures, at which times emotions fluctuate. Value assigned to the topic is challenging to change in contexts like organic chemistry in which value is traditionally high due to the importance of the subject for entrance to pre-professional school (i.e., medical, dental, veterinary) and on admissions examinations. Control, however, is a variable that can be directly influenced by instructional choices. Support of student autonomy provides a lens through which instructors can consider changing student control over the content in order to increase positive emotions and reduce negative emotions (Black & Deci, 2000). Autonomy can be supported through student-centered learning practices, in which the instructor provides critical information but allows student freedom in the use of the information; autonomy support in an active learning environment has been found to be associated with decreased anxiety and higher achievement (Black & Deci, 2000). By incorporating such instructional techniques throughout the semester, especially before and after achievement measurements (i.e., examinations), as demonstrated in Chapters Four and Five, an instructor can enhance the educational experience.

**Implications for Policy**

Resource allocation and funding opportunities for faculty members to learn about and implement EBIPs or technological tools such as CRSs in their classroom is essential for the progress of the field. The study in Chapter Two supports that an instructor's belief system is associated with their enacted practice, and encouragement from funding bodies for the use of EBIPs in association with classroom improvement initiatives should incorporate information designed to change faculty members' beliefs. This can be incorporated through professional development opportunities for faculty members to learn about EBIPs and observe demonstrations of their effectiveness, beginning the process of dissonance required to catalyze change. Funding and professional development opportunities would also be an appropriate venue to communicate results such as those in Chapter Three of this dissertation regarding the context of EBIP and educational tool use.

Policy change regarding resource allocation that can provide faculty members with opportunities to change classroom activities is encouraged by the results of Chapters Two and Three. The data presented in Chapters Four and Five support policy initiatives that relate to student activities. A variety of national bodies interested in postsecondary STEM education are concerned about a leaking pipeline, through which many students, particularly those from underrepresented groups in science, depart from STEM majors or college altogether (American Association for the Advancement of Science, 2012; National Research Council, 2012; President's Council of Advisors on Science and Technology, 2012). Affective states like those studied in Chapters Four and Five are posited to be related to retention in chemistry degree programs. The removal of negative affect has been cited as a possible tool for the remediation of student departure, along with the increased use of EBIPs and other instructional strategies improved over the traditional lecture (see Seymour, 1995; 2002, for more on the use of improved instructional strategies for retention). The work demonstrated in this dissertation indicates that enhancing positive emotions such as enjoyment can improve performance as measured by exam achievement, and increased achievement is a step in a positive direction for increased retention overall. A recommendation on the part of policymakers on the national and institutional level for an emphasis on affect in the classroom can help provide instructors with a framework for increased success.

## Summary

The four studies included in this dissertation explore different aspects of CER, however, they all seek to provide examples of improved methodology from previous studies. The study included in Chapter Two demonstrates the association between beliefs and practice in the population of chemistry faculty members within the United States. The study in Chapter Three demonstrates that context is important to consider when adoption of EBIPs or educational tools. The studies in Chapters Four and Five lend credence to the longitudinal reciprocal relationship between anxiety, enjoyment, and boredom with examination achievement in organic chemistry and the increased effect of affect on achievement at the beginning of the course. The improvements on previous research methodology included in the studies in this dissertation demonstrate a step forward for the field of CER.

## References

American Association for the Advancement of Science. (2012). *Describing & Measuring Undergraduate STEM Teaching Practices* Retrieved from Washington, DC:

Bauer, C. F., Libby, R. D., Scharberg, M., & Reider, D. (2013). Transformative research-based pedagogy workshops for chemistry graduate students and postdocs. *Journal of College Science Teaching, 43*(2), 36-43.

Black, A. E., & Deci, E. L. (2000). The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: A self determination theory perspective. *Science Education, 84*, 740-756. doi:10.1002/1098-237X(200011)84:6<740::AID-SCE4>3.0.CO;2-3

Committee on Professional Training. (2015). *Undergraduate professional education in chemistry: ACS guidelines and evaluaiton procedures for bachelor's degree programs*. Washington, D.C.: American Chemical Society.

Daniels, L. M., Stupinsky, R. H., Pekrun, R., Haynes, T. L., Perry, R. P., & Newall, N. E. (2009). A longitudinal analysis of achievement goals: From affective antecedents to emotional effects and achievement outcomes. *Journal of Educational Psychology, 101*. doi:10.1037/a0016096

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology. *American Psychologist, 56*, 218-226. doi:10.1037/0003-066X.56.3.218

Kane, R., Sandretto, S., & Heath, C. (2004). An investigation into excellent tertiary teaching: Emphasising reflective practice. *Higher Education, 47*(3), 283-310.

Larsen, R. D. (1986). Attitude reconstruction: First-day classroom advocacy. *Journal of Chemical Education, 63*(10), 864. doi:10.1021/ed063p864

Linnenbrink-Garcia, L., & Pekrun, R. (2011). Students' emotions and academic engagement: Introduction to the special issue. *Contemporary Educational Psychology, 36*, 1-3. doi:10.1016/j.cedpsych.2010.11.004

National Research Council. (2012). *Discipline-based education research: Understanding and improving learning in undergraduate science and engineering*. Washington, DC: The National Academies Press.

Pekrun, R. (2000). A social-cognitive, control-value theory of achievement emotions. In J. Heckhausen (Ed.), *Motivational psychology of human development* (pp. 143-163). Oxford: Elsevier.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*, 315-341. doi:10.1007/s10648-006-9029-9

Pekrun, R., Elliot, A. J., & Maier, M. A. (2009). Achievement goals and achievement emotions: testing a model of their joint relations with academic performance. *Journal of Educational Psychology, 101*, 115-135. doi:10.1037/a0013383

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Academic Emotions Questionnaire (AEQ). *Contemporary Educational Psychology, 36*, 36-48. doi:10.1016/j.cedpsych.2010.10.002

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research *Educational Psychologist, 37*, 91-105. doi:10.1207/S15326985EP3792_4

Pekrun, R., Hall, N. C., Goetz, T., & Perry, R. P. (2014). Boredom and academic achievement: Testing a model of reciprocal causation. *Journal of Educational Psychology, 106*, 696-710. doi:10.1037/a0036006

President's Council of Advisors on Science and Technology. (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics* Retrieved from Washington, DC:

Rogers, E. M. (1995). *Diffusion of Innovation* (4th ed.). New York: Free Press.

Seymour, E. (1995). Guest Comment: Why undergraduates leave the sciences *American Journal of Physics, 63*(3), 199-202. doi:https://doi.org/10.1119/1.17954

Seymour, E. (2002). Tracking the processes of change in US undergraduate education in science, mathematics, engineering, and technology. *Science Education, 86*(1), 79-105. doi:10.1002/sce.1044

Villafañe, S. M., Garcia, C. A., & Lewis, J. E. (2014). Exploring diverse students' trends in chemistry self-efficacy throughout a semester of college-level preparatory chemistry. *Chemistry Education Research and Practice, 15*(2), 114 -127. doi: 10.1039/c3rp00141e

Villafañe, S. M., Xu, X., & Raker, J. R. (2016). Self-efficacy and academic performance in first-semester organic chemistry: Testing a model of reciprocal causation. *Chemistry Education Research and Practice, 17*(973-984). doi:10.1039/C6RP00119J

Woodbury, S., & Gess-Newsome, J. (2002). Overcoming the paradox of change without difference: A model of change in the arena of fundamental school reform. *Educational Policy, 16*, 763-782

Xu, X., Villafañe, S. M., & Lewis, J. E. (2013). College students' attitudes toward chemistry, conceptual knoweldge and achievement: structural equation model analysis. *Chemistry Education Research and Practice, 14*(2), 188-200. doi: 10.1039/c3rp20170h

APPENDIX A:

COMMONLY USED ABBREVIATIONS

| | |
|---|---|
| ACS | American Chemical Society |
| CER | chemical education research |
| EBIP(s) | evidence-based instructional practice(s) |
| TALC | technology adoption life cycle |
| CRS | classroom response systems |
| TCSR | teacher-centered systemic reform model |
| SBTL-I | self-efficacy and beliefs about teaching and learning instrument |
| CVT | control-value theory of achievement emotions |
| AEQ | achievement emotions questionnaire |
| LRBS | learning-related boredom scale |
| CFA | confirmatory factor analysis |
| SEM | structural equation model(ing) |
| CFI | comparative fit index |
| TLI | Tucker-Lewis index |
| RMSEA | root-mean-square error of approximation |
| SRMR | standardized root-mean-square residual |

APPENDIX B:

PUBLISHER PERMISSIONS DOCUMENTATION

**Chapter Two**

5/16/2018 RightsLink Printable License

**JOHN WILEY AND SONS**
**LICENSE TERMS AND CONDITIONS**

This Agreement between Rebecca Gibbons ("You") and John Wiley and Sons ("Jo and Sons") consists of your license details and the terms and conditions provided by Wiley and Sons and Copyright Clearance Center.

| | |
|---|---|
| License Number | 4350850284930 |
| License date | May 16, 2018 |
| Licensed Content Publisher | John Wiley and Sons |
| Licensed Content Publication | Journal of Research in Science Teaching |
| Licensed Content Title | Beliefs about learning and enacted instructional practices: investigation in postsecondary chemistry education |
| Licensed Content Author | Rebecca E. Gibbons, Sachel M. Villafañe, Marilyne Stains, Kristen Murphy, Jeffrey R. Raker |
| Licensed Content Date | Feb 7, 2018 |
| Licensed Content Pages | 1 |
| Type of use | Dissertation/Thesis |
| Requestor type | Author of this Wiley article |
| Format | Electronic |
| Portion | Full article |

**Chapter Three**

**Title:** Chasm Crossed? Clicker Use in Postsecondary Chemistry Education
**Author:** Rebecca E. Gibbons, Emily E. Laga, Jessica Leon, et al
**Publication:** Journal of Chemical Education
**Publisher:** American Chemical Society
**Date:** May 1, 2017


Permission for this particular request is granted for print and electronic formats, and translations, at no charge. Figures and tables may be modified. Appropriate credit should be given. Please print this page for your records and provide a copy to your publisher. Requests for up to 4 figures require only this record. Five or more figures will generate a printout of additional terms and conditions.

Appropriate credit should read: "Reprinted with permission from {COMPLETE REFERENCE CITATION}. Copyright {YEAR} American Chemical Society." Insert appropriate information in place of the capitalized words.

I would like reuse in a Thesis/Dissertation

Requestor Author (original work)

**Chapter Four**

Our Ref: KA/CEDP/P18/0698

28 June 2018

Dear Rebecca E Gibbons,

**Material requested:** Rebecca E. Gibbons, Xiaoying Xu, Sachel M. Villafañe & Jeffrey R. Raker (2018): Testing a reciprocal causation model between anxiety, enjoyment and academic performance in postsecondary organic chemistry, Educational Psychology

Thank you for your correspondence requesting permission to reproduce the above mentioned material from our Journal in your printed thesis entitled **"**Survey research in postsecondary chemistry education: Measures of students' affect and faculty members' instructional practices**"** and to be posted in the university's repository – University of South Florida

We will be pleased to grant permission on the sole condition that you acknowledge the original source of publication and insert a reference to the article on the Journals website: http://www.tandfonline.com

This is the authors accepted manuscript of an article published as the version of record in Educational Psychology © 07 Mar 2018 - https://www.tandfonline.com/10.1080/01443410.2018.1447649

This permission does not cover any third party copyrighted work which may appear in the material requested.

Please note that this license does not allow you to post our content on any third party websites or repositories.

Thank you for your interest in our Journal.

Yours sincerely

Kendyl

**Kendyl Anderson** – Permissions Administrator, Journals
Taylor & Francis Group
3 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN, UK.
Tel: +44 (0)20 7017 7617
Fax:+44 (0)20 7017 6336
Web: www.tandfonline.com
e-mail: kendyl.anderson@tandf.co.uk

Taylor & Francis Group
an informa business

APPENDIX C:

ADDITIONAL DATA

**Chapter Two**

**Table A1.** Fit statistics for learning beliefs and self-efficacy subscales of the SBTL-I at the CFA stage (N = 1,026).

| Subscale | $\chi^2$(df) | RMSEA | CFI | SRMR |
|---|---|---|---|---|
| Learning Beliefs | 153(41)* | 0.052ǂ | 0.900 | 0.041ǂ |
| Self-Efficacy | 271(53)* | 0.064ǂ | 0.924 | 0.039ǂ |

*Note.* "*" Indicates $p < 0.001$, "ǂ" indicates a value within the range of good fit according to Hu & Bentler (1999).

**Table A2.** Fit statistics for the learning beliefs and self-efficacy subscales of the SBTL-I for the survey sample (N = 1,282).

| Subscale | $\chi^2$(df) | RMSEA | CFI | SRMR |
|---|---|---|---|---|
| Learning Beliefs | 187(43)* | 0.051 | 0.890 | 0.039ǂ |
| Self-Efficacy | 420(53)* | 0.074 | 0.922 | 0.037ǂ |

*Note.* "*" Indicates $p < 0.001$, "ǂ" indicates a value within the range of good fit according to Hu & Bentler (1999).

**Table A3.** Faculty responses to learning beliefs subscales.

| Scale | Students learn chemistry more effectively… | SA | A | N | D | SD |
|---|---|---|---|---|---|---|
| T | when working individually on problems. | 22.00 | 53.20 | 20.98 | 3.51 | 0.31 |
| T | when taking notes during a lecture. | 15.52 | 47.35 | 28.24 | 7.80 | 1.09 |
| T | by applying a set of rules or steps to algorithmic problems. | 5.69 | 33.46 | 40.09 | 17.32 | 3.43 |
| T | by completing end-of-chapter or homework problems. | 28.39 | 55.49 | 13.49 | 2.26 | 0.47 |
| T | when they have read the textbook before coming to class. | 35.57 | 44.77 | 16.69 | 2.81 | 0.16 |
| S | when working with or constructing physical or theoretical models. | 15.83 | 58.19 | 23.01 | 2.73 | 0.23 |
| S | when they understand the strengths and limitations of models and theories. | 16.46 | 53.59 | 25.82 | 3.67 | 0.47 |
| S | when working in small groups. | 21.53 | 46.80 | 25.66 | 5.46 | 0.55 |
| S | by learning to make connections between chemical concepts and daily life. | 42.43 | 47.43 | 9.13 | 0.78 | 0.23 |
| S | when interacting with computer simulations or animations. | 6.16 | 44.54 | 42.28 | 6.47 | 0.55 |
| S | when they understand the conceptual basis behind an algorithmic problem. | 30.97 | 50.08 | 16.77 | 1.95 | 0.23 |

*Note.* T = Teacher-Centered; S = Student-Centered; SA = Strongly Agree; A = Agree; N = Neutral; D = Disagree; SD = Strongly Disagree.

**Table A4**. Faculty responses to self-efficacy subscales.

| Scale | How confident are you… | CC | VC | MC | SC | NC |
|---|---|---|---|---|---|---|
| P | leading whole-class discussion? | 48.21 | 29.17 | 16.15 | 4.60 | 1.87 |
| P | facilitating small group work? | 33.46 | 34.79 | 23.95 | 6.63 | 1.17 |
| P | showing or conducting demonstrations or experiments? | 35.80 | 36.19 | 20.28 | 6.40 | 1.33 |
| P | lecturing from pre-made slides? | 30.73 | 26.37 | 16.69 | 12.01 | 14.20 |
| P | lecturing using only a whiteboard or chalkboard? | 54.84 | 28.55 | 11.70 | 2.96 | 1.95 |
| P | using technology during instruction? | 41.81 | 41.97 | 13.73 | 2.26 | 0.23 |
| C | in your ability to make connections between chemical concepts and daily life applications? | 47.66 | 38.30 | 11.62 | 2.26 | 0.16 |
| C | explaining a difficult concept in more than one way? | 43.37 | 41.34 | 13.96 | 1.17 | 0.16 |
| C | in your ability to make connections between chemistry concepts and concepts from other chemistry courses? | 40.80 | 41.89 | 15.68 | 1.48 | 0.16 |
| C | in communicating the strengths and limitations of models and theories? | 28.47 | 45.79 | 20.20 | 4.84 | 0.70 |
| C | in your ability to make connections between chemistry concepts and concepts from other non-chemistry science courses? | 32.06 | 36.51 | 24.10 | 5.93 | 1.40 |
| C | in your ability to identify difficult topics and theories? | 36.97 | 44.38 | 16.30 | 2.26 | 0.08 |

*Note.* P = pedagogy; C = content; CC = Completely Confident; VC = Very Confident; MC = Moderately Confident; SC = Somewhat Confident; NC = Not at all Confident.

**Table A5**. Descriptive statistics for enacted instructional practices.

| The following methods can be used when teaching. Please indicate how often you used these methods when you last taught [your course]. | ECM | W | STS | O | N |
|---|---|---|---|---|---|
| Lecturing (presenting content, deriving mathematical results, presenting a problem solution, etc.) | 85.7 | 8.9 | 2.9 | 2.0 | 0.6 |
| Writing on the board, projector, or document camera | 88.9 | 6.4 | 2.7 | 1.2 | 0.9 |
| Posing questions for which you expect a student response | 85.6 | 10.5 | 2.0 | 1.3 | 0.6 |
| Answering questions from students | 88.6 | 9.1 | 2.0 | 0.2 | 0.1 |
| Asking clicker questions | 14.0 | 4.0 | 2.0 | 4.3 | 75.7 |
| Follow-up and provide feedback after a clicker question or other activity | 30.1 | 13.7 | 5.1 | 4.0 | 47.1 |
| Assigning students to work in groups | 15.5 | 26.4 | 21.4 | 16.4 | 20.3 |
| Moving through the class, guiding ongoing student work | 25.1 | 22.2 | 17.9 | 15.8 | 18.9 |
| Extended discussion with small groups or individuals | 15.1 | 21.8 | 19.2 | 20.8 | 23.2 |
| Showing or conducting a demonstration, experiment, simulation, video, or animation | 6.2 | 18.8 | 38.9 | 22.5 | 13.6 |
| Asking students to make a prediction about the outcome of a demonstration or experiment before it is performed | 8.7 | 19.2 | 28.3 | 24.3 | 19.5 |
| Referencing and discussing the primary literature | 6.2 | 12.5 | 35.3 | 30.9 | 15.1 |
| Discussing the process by which a model, theory, or concept was developed | 11.0 | 29.7 | 44.9 | 12.1 | 2.3 |
| Initiating a whole class discussion, including explanation, opinion, or judgment provided by students, often facilitated by instructor | 10.1 | 13.1 | 25.5 | 28.2 | 23.1 |

*Note.* ECM = Every Class Meeting; W = Weekly; STS = Several Times per Semester; O = Once; N = Never.

APPENDIX D:

INSTITUTIONAL REVIEW BOARD APPROVALS

**#Pro00025183**

USF
UNIVERSITY OF SOUTH FLORIDA

RESEARCH INTEGRITY AND COMPLIANCE
Institutional Review Boards, FWA No. 00001669
12901 Bruce B. Downs Blvd., MDC035 • Tampa, FL 33612-4799
(813) 974-5638 • FAX(813)974-7091

January 28, 2016

Jeffrey Raker, PhD
CITRUS - Center for the Improvement of Teaching and Research in Undergraduate STEM Education
4202 East Fowler Avenue
CHE205
Tampa, FL 33620

RE: **Exempt Certification**
IRB#: Pro00025183
Title: National Survey of Postsecondary Chemistry Education (NaSPCe)

Dear Dr. Raker:

On 1/28/2016, the Institutional Review Board (IRB) determined that your research meets criteria for exemption from the federal regulations as outlined by 45CFR46.101(b):

(2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless:
(i) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and (ii) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation.

**Approved Items**:

Protocol Guidelines - JRaker PI.docx

Consent Form - JRaker PI.docx

As the principal investigator for this study, it is your responsibility to ensure that this research is conducted as outlined in your application and consistent with the ethical principles outlined in the Belmont Report and with USF HRPP policies and procedures.

Please note, as per USF HRPP Policy, once the Exempt determination is made, the application is closed in ARC. Any proposed or anticipated changes to the study design that was previously declared submitted to the IRB as a new study prior to initiation of the change. However, administrative changes, including changes in research personnel, do not warrant an amendment or new application.

Given the determination of exemption, this application is being closed in ARC. This does not limit your ability to conduct your research project.

We appreciate your dedication to the ethical conduct of human subject research at the University of South Florida and your continued commitment to human research protections. If you have any questions regarding this matter, please call 813-974-5638.

Sincerely,

Kristen Salomon, Ph.D., Vice Chairperson
USF Institutional Review Board

**#Pro00017861**

UNIVERSITY OF SOUTH FLORIDA

RESEARCH INTEGRITY AND COMPLIANCE
Institutional Review Boards, FWA No. 00001669
12901 Bruce B. Downs Blvd., MDC035 • Tampa, FL 33612-4799
(813) 974-5638 • FAX (813) 974-7091

6/18/2014

Scott Lewis, Ph.D.
USF Department of Chemistry
4202 E. Fowler Ave. CHE205
Tampa, FL 33620

RE: **Expedited Approval for Initial Review**

IRB#: Pro00017861

Title: Improving Large Lecture Gateway Chemistry Courses through Flipped Classes with Peer- Led Team Learning (NSF #1432085)

**Study Approval Period: 6/18/2014 to 6/18/2015**

Dear Dr. Lewis:

On 6/18/2014, the Institutional Review Board (IRB) reviewed and **APPROVED** the above application and all documents outlined below.

**Approved Item(s):**

**Protocol Document(s):**

IRB Research Protocol Gateway Courses.docx

**Consent/Assent Document(s)*:**

IRB Gateway Informed Consent.docx.pdf

*Please use only the official IRB stamped informed consent/assent document(s) found under the "Attachments" tab. Please note, these consent/assent document(s) are only valid during the approval period indicated at the top of the form(s).

It was the determination of the IRB that your study qualified for expedited review which includes activities that (1) present no more than minimal risk to human subjects, and (2) involve only procedures listed in one or more of the categories outlined below. The IRB may review research through the expedited review procedure authorized by 45CFR46.110 and 21 CFR 56.110. The research proposed in this study is categorized under the following expedited review category: Research involving materials (data, documents, records, or specimens) that have been collected, or will be collected solely for nonresearch purposes (such as medical treatment or diagnosis).

(5) Collection of data from voice, video, digital, or image recordings made for research purposes.

(6) Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies.

Your study qualifies for a waiver of the requirements for the informed consent process for records review, as outlined in the federal regulations at 45CFR46.116 (d) which states that an IRB may approve a consent procedure which does not include, or which alters, some or all of the elements of informed consent, or waive the requirements to obtain informed consent provided the IRB finds and documents that (1) the research involves no more than minimal risk to the subjects; (2) the waiver or alteration will not adversely affect the rights and welfare of the subjects; (3) the research could not practicably be carried out without the waiver or alteration; and (4) whenever appropriate, the subjects will be provided with additional pertinent information after participation.

As the principal investigator of this study, it is your responsibility to conduct this study in accordance with IRB policies and procedures and as approved by the IRB. Any changes to the approved research must be submitted to the IRB for review and approval by an amendment.

We appreciate your dedication to the ethical conduct of human subject research at the University of South Florida and your continued commitment to human research protections. If you have any questions regarding this matter, please call 813-974-5638.

Sincerely,

John A Schinka, Ph.D.

John Schinka, Ph.D., Chairperson
USF Institutional Review Board