Chemistry Education Research and Practice

Measuring Student Motivation in Foundation-Level Inorganic Chemistry Courses: A Multi-Institution Study

Journal:	*Chemistry Education Research and Practice*
Manuscript ID	RP-ART-07-2022-000199.R2
Article Type:	Paper
Date Submitted by the Author:	11-Sep-2022
Complete List of Authors:	Pratt, Justin; University of Rhode Island, Chemistry Stewart, Joanne; Hope College, Chemistry Reisner, Barbara; James Madison University, Chemistry and Biochemistry Bentley, Anne; Lewis & Clark College, Department of Chemistry Lin, Shirley; The United States Naval Academy, Chemistry Department Smith, Sheila; University of Michigan Dearborn Raker, Jeffrey; University of South Florida, Department of Chemistry

Measuring Student Motivation in Foundation-Level Inorganic Chemistry Courses: A Multi-Institution Study

Justin M. Pratt,*[a] Joanne L. Stewart,[b] Barbara A. Reisner,[c] Anne K. Bentley,[d] Shirley Lin,[e] Sheila R. Smith,[f] Jeffrey R. Raker[g]

[a] Department of Chemistry, University of Rhode Island, Kingston, Rhode Island 02881, United States
[b] Department of Chemistry, Hope College, Holland, Michigan 49423, United States
[c] Department of Chemistry & Biochemistry, James Madison University, Harrisonburg, Virginia 22807, United States
[d] Department of Chemistry, Lewis & Clark College, Portland, Oregon 97219, United States
[e] Department of Chemistry, The United States Naval Academy, Annapolis, Maryland 21402, United States
[f] Department of Natural Sciences, University of Michigan—Dearborn, Dearborn, Michigan 48128, United States
[g] Department of Chemistry, University of South Florida, Tampa, Florida 33620, United States
* Corresponding Author E-mail: justin.pratt@uri.edu

Abstract
The association between student motivation and learning, and changes in motivation across a course, were evaluated for students enrolled in one-semester foundation-level inorganic chemistry courses at multiple postsecondary institutions across the United States. The Academic Motivation Scale for Chemistry (AMS-Chemistry) and the Foundations of Inorganic Chemistry American Chemical Society Exam (i.e., a content knowledge measure) were used in this study. Evidence of validity, reliability, and longitudinal measurement invariance for data obtained from the AMS-Chemistry instrument with this population were found using methodologies appropriate for ordinal, non-parametric data. Positive and significant associations between intrinsic motivation measures and academic performance corroborate theoretical and empirical investigations; however, a lack of pre/post changes in motivation suggest that motivation may be less malleable in courses primarily populated by chemistry majors. Implications for inorganic chemistry instructors include paths for incorporating engaging pedagogies known to promote intrinsic motivation and methods for incorporating affect measures into assessment practices. Implications for researchers include a need for more work that disaggregates chemistry majors when evaluating relationships between affect and learning, and when making pre/post comparisons. Additionally, this work provides an example of how to implement more appropriate methods for treating data in studies using Likert-type response and nested data.

Introduction
Research on student experiences in upper-level chemistry courses (i.e., post-general chemistry and organic chemistry courses) is rare in the STEM and chemistry education research literature (National Research Council, 2012); Bodner and

Weaver (2008) argued in this *Journal* that such research, in particular, is necessary in upper-level chemistry courses due to the unique pedagogical challenges of such courses that arise from reliance on prerequisite coursework as a starting place for new learning (Bodner and Weaver, 2008). This lack of scholarship is likely due in part to small populations (*N* typically less than 20) in these courses that limit possible experimental designs, data collection techniques, and statistical power for analyses associated with studies of affect. These limited-sized populations, though, have been more suited to cognitive-focused work in upper-division courses that has in turn resulted in associated literature reviews (e.g., Bain *et al.*, 2014; Bain and Towns, 2016; Rodriguez and Towns, 2020). While insights into the affective experiences of students in general chemistry and organic chemistry courses exists (e.g., Villafañe *et al.*, 2016; Liu *et al.*, 2017, 2018; Gibbons *et al.*, 2018; Raker *et al.*, 2019), there is limited understanding of how findings from experiences of students in gateway chemistry courses translate into more homogeneously populated courses (e.g., majors-focused courses which are often upper-division courses). While students in introductory courses are making decisions about persistence in the major, students in upper-division courses are making decisions about career, graduate school, and professional school (Seymour and Hunter, 2019). Upper-division courses are where students are exposed to the broader array of chemistry subdisciplines and begin to develop specialized interests in chemistry (e.g., laser spectroscopy, computational chemistry, air-sensitive syntheses); Bodner and Weaver (2008) asserted that upper-division courses are where students are exposed to “real” chemistry. Thus, understanding affect in upper-division courses will help us to begin to understand how student experiences impact chemistry learning across the entire post-secondary chemistry curriculum. In this study, we investigated a measure of motivation for students enrolled in foundation-level inorganic chemistry courses at multiple study sites in the United States. Our results suggest that motivation is less malleable for students in non-gateway courses (i.e., is relatively unchanged between pre/post measures), and provides further evidence that the relationship between motivation and performance is persistent and significant in chemistry courses across the undergraduate curriculum.

Self-Determination Theory

Motivation is operationalized in our study through the theoretical lens of Self-Determination Theory (SDT) (Deci and Ryan, 2000; Ryan and Deci, 2000). SDT assumes that people (i.e., students taking an inorganic chemistry course in our study) have fundamental psychological needs that must be satisfied to thrive and learn. These needs include *competence, relatedness,* and *autonomy*. *Competence* is an individual’s need to feel effective in their interactions, express their understanding and abilities, and seek out challenges that align with their cognitive level. *Relatedness* is an individual's need for connection, to be cared for, and for a sense of belonging and community. *Autonomy* is an individual's need to feel in control of their environment, their actions, and their behavior. These needs factor into three types of motivation on a continuum based on autonomy (Howard *et. al.* 2017).

The two ends of the continuum are amotivation and intrinsic motivation. *Amotivation* is characterized by a lack of interest and feeling forced to do something or not being in control of the learning experience. *Intrinsic motivation* is characterized by a person's inner interests, perceived autonomy or control over a task, and competence. In the middle of the continuum lies *extrinsic motivation,* which represents various stages of increasing autonomy, competence, and relatedness. These stages are related to external factors including reward systems, deadlines, the desire to avoid guilt or shame, and finding value in the task itself. Overall, SDT operationalizes three types of motivation in terms of autonomy (i.e., level of self-determination or control), the source of the motivation, and how the type of motivation is regulated (i.e., prompted, responded to, put into action).

When SDT is used as a lens to interpret student learning, intrinsic motivation supports learning by fulfilling basic psychological needs, while amotivation hinders learning as these needs are not met. Extrinsic motivation provides a pathway from amotivation to intrinsic motivation through increasing levels of autonomy, relatedness, and competence with decreasing reliance upon external motivators. Empirical evidence from chemistry course settings and other STEM course settings supports the assumptions of SDT regarding meeting psychological needs and emphasizing intrinsic motivation to support learning (e.g., Black and Deci, 2000; Vaino *et al.*, 2012; Hagger *et al.*, 2015; Kiemer *et al.*, 2015). Previous studies have also shown that intrinsic motivation is linked to increased achievement (Lepper *et al.*, 2005; Tseng and Tsai, 2010) and persistence in STEM (Vallerand and Blssonnette, 1992; Vallerand, 1997; Allen, 1999; French *et al.*, 2005; Lavigne *et al.*, 2007; Maltese and Tai, 2011; Morrow and Ackermann, 2012).

To increase intrinsic motivation and the likelihood of success in a class, focus should be placed on supporting student autonomy, competence, and relatedness. Learning environment structures are known to strongly influence student motivation and actions in the classroom (Potvin and Hasni, 2014). With more teacher-centered instructional practices, where the teacher is perceived as the locus of authority, students have less autonomy and are more likely to feel amotivated due to lacking opportunities to demonstrate competence or feel connected with peers or the instructor (Soenens *et al.*, 2012). With more student-centered instruction that offers students more choices and control, students are more likely to develop autonomy, make connections with peers, and have opportunities to demonstrate their understanding, all of which lead to increased intrinsic motivation (e.g., American Psychological Association Presidential Task Force on Psychology in Education, 1993; Lepper and Henderlong, 2000; Chirkov and Ryan, 2001; Niemiec and Ryan, 2009; Reeve, 2012; Vansteenkiste *et al.*, 2012; León *et al.*, 2015). Situations where students have choices, instructors are encouraging, and students' inner curiosities and interests are primed, all support intrinsic motivation and the likelihood of student learning (Zimmerman, 2000; Deci and Ryan, 2008; Niemiec and Ryan, 2009).

Prior affect work in chemistry education has considered students' affective experiences in gateway chemistry courses (i.e., general chemistry and organic chemistry courses); however, the impact of academic major has been mostly overlooked (e.g., Bauer, 2005; Grove and Bretz, 2007; Liu *et al.*, 2017, 2018). This is likely because gateway courses typically have low numbers of chemistry majors, even in the context of large enrollment courses, making single-institution quantitative studies not feasible. Research outside of chemistry has shown significant relationships between motivation, and other affective measures, with academic major, as well as the alignment between major and course content (e.g., motivation for majors in majors-focused courses is different than the motivation of majors in non-majors-focused courses) (e.g., Chase and Keene, 1981; Debacker and Nelson, 2000; Zhi-ling, 2003; Cole *et al.*, 2006; Allen and Robbins, 2010; Glynn *et al.*, 2011; Kirn and Benson, 2013; Shell and Soh, 2013; Wang and Degol, 2013; Komarraju *et al.*, 2014). Additionally, studies have indicated that year in school is related to motivation (e.g., first-year and third-year students differ) (Mercincavage and Brooks, 1990; Coppola *et al.*, 1997; Lynch, 2008; Cao, 2012; Ackerman *et al.*, 2013; Planchard *et al.*, 2015). As such, there is precedent that we need to investigate motivation across chemistry courses given the concentration of majors, particularly chemistry majors, in courses varies. In addition, as students progress through the chemistry curriculum, a larger fraction of students are chemistry majors; thus, the population becomes more distinct, homogeneous, and warrants investigating how motivation relates to chemistry major and year in school.

Inorganic Chemistry Education

Inorganic chemistry courses provide an opportunity not only to shed light on an understudied context but also to study a sample of students that primarily consists of chemistry majors. Foundation-level inorganic chemistry courses are varied in both curriculum and placement in a chemistry degree program (ranging from first year courses to junior and senior-level courses); however, there is a core set of topics that unite these courses: atoms and electronic structure, covalent bonding and molecular orbital theory, transition metal complexes and coordination chemistry, acids and bases, symmetry and group theory, solids and solid-state chemistry (Raker *et al.*, 2015a). An additional commonality is that inorganic chemistry courses are predominately taken by chemistry majors (Raker *et al.*, 2015a, 2015b). It is through upper-level inorganic chemistry courses (and other upper-level courses) that students develop a deeper understanding of the scope of chemical research and potential chemistry careers. As Reisner and colleagues noted, the postsecondary inorganic chemistry course is a key opportunity to introduce primary literature and promote research skill development (Reisner *et al.*, 2015). Recent work has analyzed inorganic chemistry students' achievement emotions (i.e., anxiety and enjoyment) in relation to content knowledge, and calls for further work focused on student experiences in inorganic chemistry (Pratt and Raker, 2020). As such, based on previous research, we expect to observe differences in how motivation changes and how motivation is related to content knowledge acquisition in our foundation-level inorganic chemistry course contexts, as compared to gateway-course contexts.

Research Questions

This study is an initial investigation of the learning experience of students within an inorganic chemistry course. The purpose is to characterize student experiences in foundation-level inorganic chemistry courses, including affective experiences using the Academic Motivation Scale - Chemistry (AMS-Chemistry) (Liu *et al.*, 2017) and cognitive experiences using the 2016 ACS Foundations of Inorganic Chemistry Exam (ACS Exams, 2016). Results provide insight into how student motivation relates to inorganic chemistry content knowledge. To this end, we address three research questions (RQs):

1. What validity and reliability evidence supports using the AMS-Chemistry with students taking a foundation-level inorganic chemistry course?
2. What differences in student motivation towards inorganic chemistry are found when comparing motivation at the beginning and end of the course?
3. How is motivation towards inorganic chemistry associated with student content knowledge assessed by a summative examination?

Methods

This study had three sequential stages: (1) gathering evidence for validity and reliability of data generated from the AMS-Chemistry, (2) gathering evidence for longitudinal measurement invariance between AMS-Chemistry administrations, and (3) determining changes in AMS-Chemistry scores and associations with student content knowledge.

Participants

During Fall 2018 and Spring 2019 semesters, inorganic faculty members from eighteen universities within the United States participated in a multi-institution project focused on curricular change in foundation-level inorganic chemistry courses. While each studied foundation-level inorganic chemistry course is unique, all of the studied courses had general chemistry as a prerequisite; other perquisite requirements varied with some including organic chemistry while others required physical chemistry. Institution Review Board applications were submitted and approved at each study site.

Faculty members administered achievement/performance and affective measures to students in their foundation-level inorganic chemistry courses. At the beginning of the semester, students completed a consent form agreeing to have their data de-identified and shared with the research team. To minimize students' perceptions that participation would influence their course grades, consent forms were returned in a sealed envelope that was opened only after final grades were submitted for the term. Regardless of consent status, all students completed the cognitive and affect measures as part of their normal classroom experiences. After consent status was known, faculty members collated data for consenting students (N = 449), removed identifiers, and provided the de-identified matched data to the research team.

The focus of the study herein is the affect measures of motivation towards chemistry courses using the AMS-Chemistry (Liu *et al.*, 2017). The AMS-Chemistry was administered to students at two timepoints: *pre* (approximately the second week of the term) and *post* (approximately the last week of the term). The 2016 ACS Foundations of Inorganic Chemistry Exam (called the *ACS exam* herein) was given at all sites at the conclusion of the semester as a summative assessment of student content knowledge (ACS Exams, 2016). The ACS exam is a 60-question assessment designed to assess an array of topics covered in foundation-level inorganic chemistry courses. Only a subset of students participated in all three data collections (i.e., completed pre AMS-Chemistry, post AMS-Chemistry, and the ACS exam); N = 396 students participated in at least one of the three data collections (82% of total sample). For each analysis presented below, listwise deletion was used to subset the data corpus appropriately.

Academic Motivation Scale – Chemistry (AMS-C)

The AMS-Chemistry (Liu *et al.*, 2017) is a 28-question survey designed to measure three types of course-specific motivation aligned with Self-Determination Theory: amotivation, extrinsic motivation, and intrinsic motivation (Deci and Ryan, 2000; Ryan and Deci, 2000). The survey was adapted for use in postsecondary chemistry contexts from the Academic Motivation Scale (Vallerand *et al.*, 1992). Work with general chemistry (Liu *et al.*, 2017) and organic chemistry (Liu *et al.*,

2018) students supports that the AMS-Chemistry has seven subscales (see Figure 1). Each subscale corresponds to a motivation sub-construct with differing degrees of regulation (i.e., the thoughts, actions, or behaviors through which students act to influence their choice, effort, or persistence) which is how SDT was operationalized into the AMS-C. For each subscale there are four questions/items with Likert-type response formats (Carifio and Perla, 2007). Students indicate "*To what extent each of the following statements corresponds to one of the reasons why you are enrolled in this chemistry course*" with five response options: *Not at all* (A), *A Little* (B), *Moderately* (C), *A Lot* (D), and *Exactly* (E). To make the AMS-Chemistry more course-specific, the stem was modified to specify "*…why you are enrolled in this inorganic chemistry course*" (see Appendix 1). Modifying the stem in this way is supported by previous work measuring achievement emotions in another chemistry-specific disciplinary context (Raker *et al.*, 2019; Pratt and Raker, 2020).

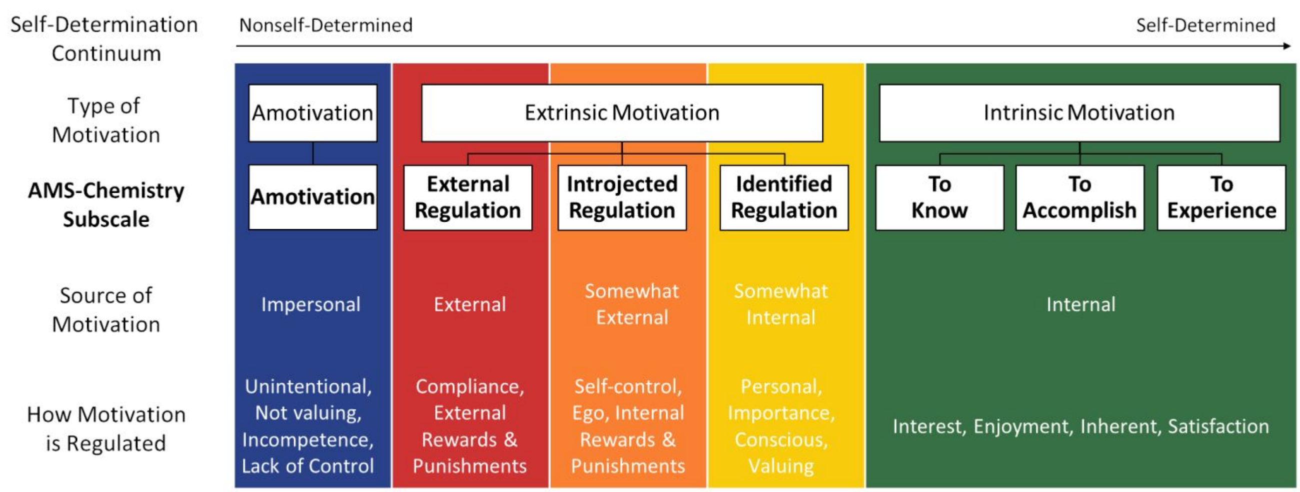

Fig. 1: Diagram showing the subscales of the AMS-Chemistry (Liu *et al.*, 2017) and the alignment with Self-Determination Theory (Deci and Ryan, 2000; Ryan and Deci, 2000; Reiche, 2013).

Gathering Evidence of Validity and Reliability for AMS-Chemistry Data

While the AMS-Chemistry has been shown to function well with general chemistry and organic chemistry student populations, the AMS-Chemistry has not previously been used with an inorganic chemistry student population. Therefore, it is important to provide evidence for validity and reliability of generated data, in line with the *Standards for Educational and Psychological Testing* (called the *Standards* herein) (Arjoon *et al.*, 2013; American Education Research Association *et al.*, 2014). One key evidence for validity in this study is internal structure validity to determine if the AMS-Chemistry functioned similarly with this new sample of students (i.e., measured the intended seven subscales/factors of motivation with the new sample). To investigate this, categorical responses (A-E) were numerically transformed (1-5, respectively) to represent the ordinal nature of the data, and confirmatory factor analyses (CFA) were conducted. Previous studies of the internal structure validity of data from the AMS-Chemistry have treated the data as continuous when conducting CFA analyses (Liu *et al.*, 2017, 2018). However, the data generated are ordinal and should be treated as such by using either diagonal weighted least squares (DWLS) or weighted least squares—mean and variance adjusted (WLSMV) estimators. DWLS is the default estimator for ordinal data in the LISREL program (Jöreskog, 2019) and the *R* (R Core Team, 2017) package Lavaan (Rosseel, 2012), while WLSMV is the default ordinal estimator in Mplus (Muthén and Muthén, 2017). When sample sizes are small and data are non-normally distributed, DWLS can produce problematic results and WLSMV is recommended (DiStefano and Morgan, 2014). As such, we conducted a seven factor CFA using the WLSMV estimator in Mplus 8.3. A schematic of the seven-factor model is provided in Figure 2. Note: all latent constructs (subscales) are modeled as correlated based on previous analyses of the AMS-Chemistry (Liu *et al.*, 2017, 2018) and based on the theoretical conjecture that all seven constructs/subscales are interrelated with motivation. Additionally, this CFA is a congeneric model where the only constraints are on which items load onto which latent constructs, and which latent constructs are correlated; no correlated error variances were modelled. This model aligns with previous investigations of the internal structure of the AMS-Chemistry. While the congeneric CFA model evaluates the internal structure of the AMS-Chemistry for the overall inorganic chemistry student sample, a multilevel CFA was also conducted to evaluate the internal structure of the AMS-Chemistry when accounting for the nested nature of the data; since students are nested within multiple classrooms/institutions, a multilevel CFA evaluates the impact of between-institution differences on the internal structure validity of the data (Dyer *et al.*, 2005).

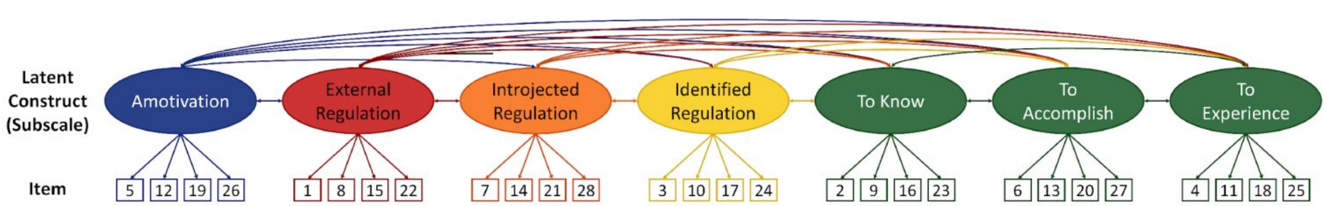

Fig. 2: Diagram showing the seven factor CFA model testing for internal structure validity of AMS-Chemistry data with students taking a foundation-level inorganic chemistry course. To simplify interpretation, errors are not displayed in the representation.

CFAs are evaluated based on model fit using a combination of statistics, including the chi-squared test (χ^2), the Comparative Fit Index (CFI), the Tucker Lewis Index (TLI), the Root Mean Square Error of Approximation (RMSEA), and the Standardized Root Mean Squared Residual (SRMR) (Hooper *et al.*, 2008; Kline, 2016; Parry, n.d.). However, when the

WLSMV estimator is used, only the CFI, TLI, and RMSEA are appropriate fit indices (Yu, 2002; Beauducel and Herzberg, 2006; Bandalos, 2008; Komperda, Hosbein, *et al.*, 2018). Chi-squared values, while reported, are typically only used for comparison purposes and not as indicators of model fit due to sensitivity to sample size and model complexity (Cheung and Rensvold, 2002; Schermelleh-Engel *et al.*, 2003; Chen, 2007; Mueller and Hancock, 2008; Liu *et al.*, 2017; Komperda, Hosbein, *et al.*, 2018). There are a variety of cutoff values in the literature used to evaluate the fit of CFA models; cutoff values can be chosen based on the estimator used and/or previous work within the discipline. Experts in structured equation modeling do not agree on acceptable cutoffs, particularly when using the WLSMV estimator (Beauducel & Herzberg, 2006, Huggins-Manley & Han, 2016, Padgett & Morgan 2021); however, consensus exists that more conservative cutoff values are necessary when using the WLSMV estimator. For context, previous work within CER, as well as on previous psychometric analyses of the AMS-Chemistry, have used cutoff values of CFI > 0.91, TLI > 0.91, and RMSEA < 0.08 (Liu *et al.*, 2017, 2018; Gibbons *et al.*, 2018; Komperda, Hosbein, *et al.*, 2018; Raker *et al.*, 2019). As we employed the WLSMV estimator in this study, more conservative cutoff values were chosen, in line with recent evaluations of the WLSMV estimator: our measures of fit indices were determined to be “acceptable” based on these cutoff metrics: CFI ≥ 0.98, TLI ≥ 0.98, and RMSEA ≤ 0.07 (Padgett & Morgan 2021).

To investigate the reliability of the data (e.g., the internal consistency of each subscale), Cronbach’s alpha (α) coefficients are typically used (Cronbach, 1951). However, Cronbach’s α assumes unidimensionality and that all items are associated with the underlying factor to the same degree (i.e., equal item loadings in a parallel or tau equivalent model) (Komperda, Pentecost, *et al.*, 2018). McDonald’s omega (ω) coefficient is a less restrictive alternative that relaxes the requirement for equal item loadings (i.e., a congeneric model) (McDonald, 1999; Hancock and An, 2020), but is conceptually like Cronbach’s α in interpretation (Zinbarg *et al.*, 2005; Hancock and An, 2018). To determine which coefficient is appropriate for the AMS-Chemistry in this context, we conducted stepwise factor analyses where more constraints were added to the model, following recommendations by Komperda and colleagues (Komperda, Pentecost, *et al.*, 2018):

1. Equal structure but freely estimated item errors and item loadings to the latent factor (i.e., the congeneric model)
2. Equal structure and equal item loadings but freely estimated item errors (i.e., the tau equivalent model)
3. Equal structure, item loadings, and item errors (i.e., the parallel model)

Given the stepwise, additive nature of these analyses, analysis ends when one of the models fails to have acceptable fit statistics, using the cutoff values articulated above. There are no consistently agreed upon criteria for interpreting internal consistency coefficients (i.e., acceptable values for either Cronbach's α or McDonald's ω) (Arjoon *et al.*, 2013; Taber, 2018). However, for either coefficient, a value of 0.7 or higher is considered acceptable and consistent with previous AMS-Chemistry studies (Liu *et al.*, 2017, 2018) and other affect studies in CER (Gibbons *et al.*, 2018; Raker *et al.*, 2019; Pratt and Raker, 2020).

Investigating Longitudinal Measurement Invariance

The AMS-Chemistry was administered at two timepoints in this study, and it is therefore inappropriate to assume the instrument functioned similarly at both administrations; it is necessary to investigate whether there is evidence for measurement invariance between timepoints prior to any pairwise comparisons. Measurement invariance (or measurement equivalence) is evidence to suggest that the same construct(s) are being measured across some grouping variable (e.g. between demographic variables, timepoints, etc.) (Dimitrov, 2010; Millsap, 2011; Bialosiewicz *et al.*, 2013; Bandalos, 2018; Rocabado, *et al*., 2020). While explicitly discussed in the *Standards* (Standard 7.1), many studies assume that instruments function similarly between groups/timepoints and conduct pairwise comparisons without evidence to support the appropriateness of the analyses (American Education Research Association *et al.*, 2014).

For this study, measurement invariance was investigated between AMS-Chemistry administrations (i.e., longitudinal measurement invariance between pre and post administrations) with no correlated error variances. While multilevel confirmatory factor analysis was investigated, the small number of second-level groups (only 11 institutions had data for consented individuals for both pre & post AMS-Chemistry administrations) as well as small sample sizes at individual sites (ranging from 5 to 55) prohibit analogous measurement invariance studies. To this end, longitudinal measurement invariance was only investigated for the overall data corpus. To conduct the analysis, listwise deletion was used to subset the data corpus into a subset of participants with both pre and post AMS-Chemistry responses. A series of increasingly restrictive CFA models were tested on the subset where various constraints were set equal between groups and timepoints (Dimitrov, 2010; Millsap, 2011; Bialosiewicz *et al.*, 2013). The stepwise, additive process for measurement invariance testing includes:

1. Testing for configural invariance (i.e., same factor structure between administrations)
2. Testing for metric invariance (i.e., same factor structure and equal item loadings to latent constructs between administrations)
3. Testing for scalar invariance (i.e., same factor structure, equal item loadings, and equal item intercepts between administrations)

4. Testing for strict invariance (i.e., same factor structure, equal item loadings, equal item intercepts, and equal item residuals between administrations)

Given the stepwise and additive nature of this analysis, analysis ends when one of the tests fails to have acceptable fit statistics, using the same cutoff values as previously described. If evidence for all levels of measurement invariance are found, the evidence then supports that the AMS-Chemistry functioned similarly between the two administrations and pairwise comparisons are appropriate to conduct.

It is worth noting that the most restrictive model, called “strict” or residual invariance (i.e., adding in equal error variances between groups) (Putnick and Bornstein, 2016), is a necessary step for researchers investigating full factorial invariance (Meredith, 1993) and seeking to compare observed scores. Testing for residual invariance is not a requirement for studies seeking to compare latent factor means between groups (i.e., not observed scores) since residuals are not part of the latent factors (Vandenberg and Lance, 2000).

Interpreting Scores from the AMS-Chemistry

Based on the validity and reliability evidence (i.e., the seven-factor solution has acceptable fit and reliability coefficients, see the Results & Discussion section), it was appropriate to analyze student responses to the AMS-Chemistry in line with Research Questions 2 and 3. To interpret AMS-Chemistry scores in relation with a summative measure of content knowledge (i.e., ACS Exam), two simultaneous analyses were conducted. The first analysis used a measured variable framework where each factor was treated as individual subscales as suggested by the validity and reliability evidence. Factor scores/means were calculated for each subscale for the AMS-Chemistry by averaging the numeric responses (1-5) for the four items that comprised each subscale (a higher average score on the subscale is indicative of increasing agreement with those items and therefore evidence that the student has that type of motivation regulation, see Figure 1); factor scores were then correlated with ACS exam raw score (total number of questions correct, max of 60). Because of sample size and non-normal distribution of data, the non-parametric correlation coefficient Spearman’s Rho (ρ) was used (Spearman, 1904; Dodge, 2008). The second analysis used structured equation modeling to retain the original ordinal nature of the response options and is more in line with state-of-the-art quantitative research; in this analysis, AMS-Chemistry responses were grouped by subscale/latent factor and then used to predict ACS exam raw scores.

Additionally, because evidence for longitudinal measurement invariance was found (i.e., the AMS-Chemistry functioned similarly between the two administrations, see the Results & Discussion section), comparisons were conducted between pre and post scores. The non-parametric omnibus Friedman Test (Friedman, 1937, 1939) and the non-parametric paired samples Wilcoxon Signed Ranks Test (Wilcoxon, 1945; Rey and Neuhäuser, 2011) were used to compare subscale scores between timepoints due to the ordinal and non-normally distributed data. Because multiple tests were conducted (i.e., each test for each subscale), the chance of a Type 1 error increases (i.e., increases the chance of concluding that there is a significant result when there is not one in reality). Therefore, to minimize Type 1 error, a Bonferroni correction was employed (Dunn, 1961; Haynes, 2013). The correction is performed by dividing our original significance level ($\alpha = 0.05$) by the number of tests performed (i.e., seven). Therefore, if $p \leq 0.007$ it is considered significant.

Results and Discussion

The results of this study showed positive evidence for the validity and reliability of data gathered using the AMS-Chemistry with foundation-level inorganic chemistry students. Additionally, evidence suggested measurement invariance between timepoints indicating that pre-post comparisons were appropriate. Lastly, associations between intrinsic motivation factors and content knowledge (as measured by an ACS exam) were found. In this section we provide a discussion of how these results support prior work, offer new insights into investigating the learning experiences of a predominantly chemistry major student population within a non-gateway course, provide a foundation for considering motivation in the context of inorganic chemistry courses, and provide a useable exemplar for rigorous measurement analyses.

Evidence of Validity and Reliability for AMS-Chemistry Data

Table 1 includes the response rates for both administrations of the AMS-Chemistry and the ACS exam. Not all students completed every measure; as such, we have delineated subsample sizes for students who completed both pre and post administrations of the AMS-Chemistry (i.e., measurement invariance analysis), as well as both the post AMS-Chemistry and ACS exam (i.e., relationship between motivation and content knowledge analysis).

Table 1. Number of student responses for survey and examination administrations.

Measure(s)	Students (*n*)	Percentage of Student Sample (*N* total = 449)
Pre AMS-Chemistry	243	54.1
Post AMS-Chemistry	153	34.1
Pre & Post AMS-Chemistry	133	29.6
ACS Exam	269	59.9

Post AMS-Chemistry & ACS Exam	134	29.8

To investigate the internal structure validity of AMS-Chemistry data, multiple CFAs were conducted using the WLSMV estimator and allowing the models to freely estimate all factor loadings and errors. CFA fit information for the seven-factor solution of the AMS-Chemistry that was previously proposed in prior investigations are provided in Table 2 for the entire data corpus (collapsing administrations together, n = 396), the pre administration (n = 243), and the post administration (n = 153). Conducting this analysis on individual administrations (pre/post) as well as on the entire data corpus (collapsing pre/post together) allows us to comment on the overall fit of the imposed factor structured on the data. As our sample sizes from individual administrations are considered small, conducting the full data corpus analysis was useful in evaluating the overall factor structure of our dataset. In all instances, fit indices were well within or at the boundaries of good fit. These results support that with this sample of students taking a foundation-level inorganic chemistry course, the AMS-Chemistry functions in accordance with SDT and previous work with general chemistry and organic chemistry students (Liu *et al.*, 2017, 2018). Furthermore, both pre and post AMS-Chemistry models have evidence for internal structure validity, thus analyzing the administrations separately is supported.

Table 2. Confirmatory factor analysis fit information for 7-factor congeneric measurement models using the WLSMV estimator.

Model	χ^2	*df*	*P*	CFI	TLI	RMSEA	RMSEA 90% Confidence Interval
Entire Data Corpus (n = 396)	961.2	329	< 0.001	0.975	0.971	0.070	0.065 – 0.075
Pre AMS-Chemistry (n = 243)	687.6	329	< 0.001	0.977	0.973	0.067	0.060 – 0.074
Post AMS-Chemistry (n = 153)	549.8	329	< 0.001	0.979	0.975	0.066	0.056 – 0.076
Multilevel – Entire Data Corpus	742.4	658	0.012	0.980	0.977	0.018	

Note: cut-off values for indication of good model fit for this study are CFI ≥ 0.98, TLI ≥ 0.98, and RMSEA ≤ 0.07.

To evaluate the impact of the nested nature of the data (i.e., multiple institutions/sites) on the internal structure validity, intraclass coefficients (ICC) were first calculated; ICC values provide information about the impact of between-institution differences on the model. Previous studies have suggested that ICC values ≥ 0.1 are evidence that the nested nature of the data do impact the model and thus a multilevel CFA is needed (Vajargah and Nikbakht, 2015). ICC values for the 28 AMS-Chemistry items are provided in Appendix 2; the values range from 0.036 to 0.193 with only 11 items having values ≥ 0.1. As such, the ICC values provide inconclusive evidence for whether the between-institution differences impact the overall model. To error on the conservative side, we chose to conduct a multilevel CFAs on the entire data corpus, the pre administration, and the post administration to evaluate the impacts of the nested nature of the data. This multilevel approach adds an institution/context factor (level two) to the model above the seven subscales/latent factors shown in Figure 2 (i.e., a hierarchical model). Unfortunately, we were unable to evaluate multilevel CFAs for the pre and post administrations individually due to the small number of institutions/nests, as well as minimal variability between institutions. As such, only the collapsed full data corpus was evaluated in this manner. Model fit information for the multilevel CFA for the entire data corpus are provided in Table 2; the fit indices are within the boundaries of good fit and show improved fit over the un-nested models. This provides further evidence for internal structure validity of the AMS-Chemistry when administered with foundation-level inorganic chemistry students while accounting for between-institution differences.

To investigate the reliability of the AMS-Chemistry, we determined which reliability coefficient was appropriate (i.e., Cronbach's α or McDonald's ω). Assumptions underlying the two coefficients were tested using additional CFA models. The first assumption for both measures is whether a congeneric model (i.e., freely estimated errors and item loadings) fits the data. Given that the fit indices for the congeneric models for both the pre and post AMS-Chemistry were within the cut-off values for good fit (see Table 2), the next step was to test whether items within each subscale load equally to the latent construct (i.e., tau equivalent models, an assumption for Cronbach's α). The results of the CFA where constraints were added to force items within factors to load equally are given in Table 3.

Table 3. Confirmatory factor analysis fit information for 7-factor tau equivalent measurement models using the WLSMV estimator.

Model	χ^2	*df*	*p*	CFI	TLI	RMSEA
Pre AMS-Chemistry	1366.1	350	< 0.001	0.934	0.928	0.109
Post AMS-Chemistry	829.9	350	< 0.001	0.953	0.950	0.095

Note: cut-off values for indication of good model fit for this study are CFI ≥ 0.98, TLI ≥ 0.98, and RMSEA ≤ 0.07.

Both pre and post AMS-Chemistry tau equivalent models have fit indices outside the limits of good fit. This suggests that both models lack strong overall fit evidence when constraints of equal item loadings are applied. Therefore, *Cronbach's α is not an appropriate reliability coefficient for these data*, and McDonald's ω is more appropriate as it does not require equal item loadings within factors (McDonald, 1999; Komperda, Pentecost, *et al.*, 2018; Hancock and An, 2020). However, an assumption of McDonald's ω is that scales are unidimensional, meaning that each AMS-C subscale needs to be evaluated individually to ensure proper model fit prior to evaluating reliability. Previously, all models evaluated the full 7-factor model without evaluating each factor individually. Reported in Appendix 4 are the results of evaluating each subscale individually. The results of these single factor congeneric models are varied. Both pre and post administrations of the *identified regulation* subscale fail to have fit indices within the boundaries of good fit, indicating that this subscale does not meet the assumptions of reliability. All other subscales and timepoints have CFI and TLI indices approaching or within the boundaries of good fit. However, RMSEA values are varied with some approaching adequate fit and others well outside the bounds of adequate fit. Given the small sample sizes at both timepoints/administrations, these findings are less concerning as RMSEA values are sensitive to sample size resulting in evaluated values when sample sizes are small (Taasoobshirazi and Wang, 2016). As such, we focus on the CFI and TLI values primarily in this analysis which all meet or exceed the cutoff values. McDonald's ω coefficient values for the subscales with adequate fit, for both administrations of the AMS-Chemistry, are reported in Table 4. Across both administrations, all scales with evidence to support unidimensionality/the assumptions of reliability have $\omega > 0.80$. These findings support the internal consistency/reliability of the data generated from the AMS-Chemistry with students taking a foundation-level inorganic chemistry course.

Table 4. McDonald's ω reliability coefficients for the seven subscales of the AMS-Chemistry at both timepoints

	McDonald's ω	
Subscale	Pre (n = 243)	Post (n = 153)
Amotivation	0.90	0.86
External Regulation	0.89	0.88
Introjected Regulation	0.90	0.90
Identified Regulation	--	--
To Know	0.87	0.89
To Accomplish	0.90	0.90
To Experience	0.84	0.87

Note: Subscales that do not meet the assumption of unidimensionality do not have a reported ω

Longitudinal Measurement Invariance

Tests of longitudinal measurement invariance were conducted to determine whether the instrument functioned similarly between the two administrations (n = 133 students completed both pre & post administrations). To conduct this analysis, every item on the instrument must have at least one data point for all response options (1-5). Question 2 on the post administration lacked any respondents that selected option *Not at all*. As recommended by the MPlus user guide, data were manipulated where a randomly selected respondent who originally selected *A Little* on Question 2 was recoded to *Not at all* (Muthén and Muthén, 2017). This results in modifying a single data entry of more than 3,000 data points (< 0.03% of the data) and is consistent with recommendations for conducting CFAs with ordinal data (Muthén and Muthén, 2017). Results from this analysis are shown in Table 5.

Table 5. Longitudinal measurement invariance fit information and model comparisons for 7-factor measurement models comparing pre and post administrations using the WLSMV estimator.

Model	χ^2	*df*	*p*	$\Delta\chi^2$	Δdf	*p* for $\Delta\chi^{2*}$	CFI	ΔCFI	TLI	RMSEA	ΔRMSEA
Configural (equal structure)	1236.4	658	< 0.001	--	--	--	0.977	--	0.974	0.067	--
Metric (add in equal item loadings)	1245.5	679	< 0.001	12.9	21	0.913	0.978	0.001	0.975	0.065	-0.002
Scalar (add in equal item intercepts)	1314.8	753	< 0.001	86.9	74	0.145	0.978	0.000	0.978	0.061	-0.004
Strict (add in equal item residuals)	1371.8	796	< 0.001	-43	43	0.292	0.981	0.003	0.982	0.055	-0.006

Note: cut-off values for indication of good model fit fir this study are CFI ≥ 0.98, TLI ≥ 0.98, and RMSEA ≤ 0.07.

*$\Delta\chi^2$ difference test conducted at α = 0.05 using the *difftest* approach in Mplus (Asparouhov and Muthén, 2019; Koziol, 2010; Muthén and Muthén, 2017; Suh and Cho, 2014)

When conducting measurement invariance analyses, it is necessary to see how the fit indices change as increasing constraints are added. These additional criteria provide evidence for measurement invariance (Chen, 2007; Koziol, 2010; Rutkowski and Svetina, 2014; Suh and Cho, 2014; Putnick and Bornstein, 2016; Asparouhov and Muthén, 2019):

- Non-significant differences in the χ^2 values for consecutive models (i.e., the additional constraints did not significantly change the χ^2 value)
- A ΔCFI of ≤ 0.01 between consecutive models
- A ΔRMSEA of ≤ 0.015 between consecutive models

Shown in Table 5 are fit indices and fit indices changes between consecutive models. Generally, all provide evidence for measurement invariance. It is worth noting that the ΔRMSEA for the strict model is outside the criteria for good fit; all other measures are within cutoff criteria. The complexity of the model (i.e., seven factors and two administrations), coupled with the impact small sample sizes have on RMSEA values, help to alleviate any concerns and supports strict measurement invariance of the data (Taasoobshirazi and Wang, 2016). Therefore, there is evidence to suggest that the AMS-Chemistry functioned similarly at both timepoints and pre/post pairwise comparisons are acceptable using observed scores.

Interpreting Scores from the AMS-Chemistry

Descriptive statistics for students with subscale scores at both timepoints (n = 133) as well as for the ACS Foundations of Inorganic Chemistry Exam (n = 134) are reported in Table 6. While descriptive statistics are provided for all subscales, it is worth reminding the reader that both pre and post *identified regulation* data failed to meet the assumptions of reliability; the results of this subscale/factor should be interpreted cautiously.

Table 6. Descriptive statistics for the seven subscales of the AMS-Chemistry for pre and post administrations and the ACS exam.

Measure	Min	Max	Mean	Median	Mode	Standard Deviation	Skew	Kurtosis
Pre Amotivation	1	4	1.2	1	1	0.43	2.91	10.65
Pre External Regulation	1	5	3.0	3	4	1.02	-0.04	-1.13
Pre Introjected Regulation	1	5	3.1	3	3	1.08	-0.18	-0.67
Pre Identified Regulation	1	5	4.0	4	5	0.81	-0.97	0.70
Pre To Know	2	5	3.8	4	4	0.85	-0.49	-0.26
Pre To Accomplish	1	5	3.5	4	4	0.95	-0.30	-0.66
Pre To Experience	1	5	3.1	3	3	1.00	-0.19	-0.53
Post Amotivation	1	4	1.3	1	1	0.53	2.22	5.21
Post External Regulation	1	5	2.9	3	3	0.98	-0.21	-0.91
Post Introjected Regulation	1	5	3.2	3	4	1.03	-0.16	-0.87
Post Identified Regulation	1	5	3.9	4	4	0.78	-1.02	1.07
Post To Know	1	5	3.6	4	4	0.84	-0.41	-0.38
Post To Accomplish	1	5	3.4	4	4	0.94	-0.29	-0.57
Post To Experience	1	5	3.2	3	4	1.02	-0.28	-0.69
ACS Exam Raw Score (out of 60)	14	52	31.3	29.5	26	8.90	0.43	-0.69

Scores for all measures are non-normally distributed (Shapiro-Wilk test for normality for all measures is $p < 0.05$) and data are ordinal (Shapiro and Wilk, 1965); therefore, non-parametric statistics were used to analyze the data. A Friedman test was first performed to indicate if there were detectable differences between timepoints across the seven subscales of the AMS-Chemistry; results showed that there were detectable differences: χ^2 (13) = 817.2, $p < 0.001$. Follow-up tests to determine the source(s) of differences by comparing pre/post subscale scores using Wilcoxon Signed Ranks tests are shown in Table 7 (Wilcoxon, 1945; Rey and Neuhäuser, 2011). No evidence of significant differences were found after the Bonferroni correction ($p \leq 0.007$) between the pre and post administrations of the AMS-Chemistry with this sample.

Table 7. Wilcoxon Signed Ranks Test comparing pre and post AMS-Chemistry subscale scores (*n* = 133)

Subscale	*Z*	*p**
Amotivation	1.713	0.087
External Regulation	0.670	0.503
Introjected Regulation	0.635	0.525
Identified Regulation	--	--
To Know	2.626	0.009
To Accomplish	0.252	0.801
To Experience	0.873	0.383

*Note: significance determined at $\alpha \leq 0.007$ with Bonferroni correction. The Identified Regulation subscale was not evaluated due to failing to meet reliability assumptions.

These results differ from previous investigations where significant changes over time were found with *External Regulation* with general chemistry students (Liu *et al.*, 2017) and *Amotivation*, *Introjected Regulation*, and *Identified Regulation* with organic chemistry students (Liu *et al.*, 2018). Given that most students enrolled in foundation-level inorganic chemistry courses are chemistry majors, these results may suggest that chemistry majors are a different population than non-chemistry majors in terms of how motivation towards chemistry courses is related to course performance. Previous findings from the AMS-Chemistry with general chemistry and organic chemistry students, where major was not studied or disaggregated, should be interpreted cautiously when extrapolating to non-gateway courses, and future studies should explicitly investigate the impact of major and/or career goals on the development of motivation related to chemistry coursework.

To investigate the association between motivation and content knowledge, two simultaneous analyses were conducted. The first analysis used a measured variable framework that involved calculating correlations between the post administration of the AMS-Chemistry and raw scores from the ACS Foundations of Inorganic Chemistry Exam (ACS Exams, 2016). Previous analyses of the AMS-Chemistry have used similar, measured variable frameworks so analyzing the data in this manner will allow for comparisons to other studies. Additionally, the measured variable analysis is most similar to how chemistry practitioners may analyze these results (i.e., calculating scores for all seven AMS-Chemistry subscales) making this analysis useful for readers.

In addition to the measured variable framework, a more rigorous analysis was also conducted that considered both the non-normal distribution of the data as well as the ordinal nature of response options; in this analysis, structured equation modeling was used to predict raw scores from the ACS exam using post AMS-Chemistry responses, grouped by latent factor/subscale. While this is a more robust analysis that is aligned with the types of data collected, it is not an analysis that a chemistry instructor would typically use. Thus, providing both the measured framework analysis and the structured equation modeling analysis is warranted to speak to both researchers and practitioners, a specific goal of this *Journal*.

In both analyses, post administrations of the AMS-Chemistry were compared with ACS exam raw scores due in part to the temporal proximity in which each were completed, as well as literature precedence from other studies using the AMS-Chemistry (Liu *et al.*, 2018). Results from the measured variable framework (see Table 8) indicate that two of the seven subscales have significant correlations with the ACS exam: *To Know* and *To Accomplish*. These subscales are measures of intrinsic motivation, and both subscales have positive, but small correlations with ACS exam raw score. These results agree with Self-Determination Theory which suggests that increased intrinsic motivation (i.e., internalized motivation associated with interest and enjoyment in the topic) supports learning (Deci and Ryan, 2000; Ryan and Deci, 2000). While no other subscales have significant correlations at the Bonferroni corrected level, it is worth noting the magnitude and signs of the correlation values for *To Experience* (small and positive) and *Amotivation* (small and negative). These two subscales further support the claims of Self-Determination Theory; *Amotivation* (i.e., the lack of motivation) hinders learning while *To Experience* (an intrinsic motivation component) supports learning. As such, these results provide further empirical evidence of the impact of affect on learning, specifically student motivation and student content knowledge. Previous work in general chemistry and organic chemistry courses have further supported this claim, with similar results showing the impact of amotivation and intrinsic motivation on learning (Liu *et al.*, 2017, 2018). Additionally, work in organic chemistry has also demonstrated a strong connection between student motivation and emotions directly related to achievement (enjoyment, anxiety, etc.) (Raker *et al.*, 2019).

Table 8. Spearman's ρ correlations comparing post AMS-Chemistry subscale

scores with ACS Foundations of Inorganic Chemistry Exam raw scores (n = 134)

Subscale	Spearman's ρ	p*
Amotivation	-0.172	0.047
External Regulation	-0.004	0.963
Introjected Regulation	0.130	0.133
Identified Regulation	--	--
To Know	0.281	0.001*
To Accomplish	0.248	0.004*
To Experience	0.222	0.010

*Note: significant at $\alpha \leq 0.007$ with Bonferroni correction. The Identified Regulation subscale was not evaluated due to failing to meet reliability assumptions.

Results from the structured equation modeling analysis are included in Table 9. From this analysis, three AMS-Chemistry subscales are significant predictors of raw ACS exam scores: *To Know*, *To Accomplish*, and *To Experience*. All three of these subscales are measures of intrinsic motivation, which Self-Determination Theory suggests supports student learning (Deci and Ryan, 2000; Ryan and Deci, 2000). Similar to the measured variable framework, signs for the non-significant predictors provide further support for SDT as a useful framework for conceptualizing student learning in relationship to motivation.

Table 9. Structured equation modeling analysis using the WLSMV estimator to predict ACS Exam raw scores from post AMS-Chemistry responses, grouped by subscale (n = 134)

Subscale	Standardized Coefficient	Standard Error	p*
Amotivation	-0.148	0.076	0.051
External Regulation	-0.026	0.085	0.756
Introjected Regulation	0.136	0.083	0.100
Identified Regulation	--	--	--
To Know	0.282	0.076	< 0.001*
To Accomplish	0.263	0.084	0.002*
To Experience	0.253	0.081	0.002*

*Note: significant at $\alpha \leq 0.007$ with Bonferroni correction. The Identified Regulation subscale was not evaluated due to failing to meet reliability assumptions.

Overall, results from both analyses suggest that instructors of inorganic chemistry courses should consider affect in addition to measures of content knowledge when evaluating the success of students and impacts of instructional practices. Considerations of affect should be incorporated into classroom culture, practices, assessments, etc. and used as additional evidence to support teaching decisions. Results comparing pre/post administrations (considering previous findings with general chemistry and organic chemistry students) also suggest that while motivation is related to learning, motivation may also be difficult to change. As such, using instructional practices that promote intrinsic motivation (e.g., active learning approaches) can help support student success while also impacting affective characteristics that are key components of the learning process (Farrell *et al.*, 1999; Eberlein *et al.*, 2008; National Research Council, 2012; Freeman *et al.*, 2014; Wieman, 2014).

Conclusions and Implications

Our results support using the AMS-Chemistry with students taking foundation-level inorganic chemistry courses. Evidence from our study suggests that the instrument elicits valid and reliable data and provides insights into the student experience (RQ 1). Additionally, the lack of changes in motivation over a semester may indicate unique characteristics of chemistry majors enrolled in chemistry courses (RQ 2). Despite the lack of observed changes over time, associations found between intrinsic motivation and content knowledge add to the growing body of literature emphasizing the impacts of affect on student learning (RQ 3). Associations found with amotivation and extrinsic motivation complement previous research examining motivation with general chemistry and organic chemistry students (Liu *et al.*, 2017, 2018) and provide additional empirical evidence to support SDT as a lens for interpreting student experiences in the classroom (Deci and Ryan, 2000; Ryan and Deci, 2000).

Overall, these findings have clear implications for researchers and practitioners. For researchers, our analyses investigating the validity, reliability, and invariance of AMS-Chemistry data serves as a model for future work where *Likert-type response data are treated as ordinal and not continuous* (Carifio and Perla, 2007; Bishop and Herron, 2015) and the nested (i.e., multi-level) nature of collected data are considered in analyses. Additionally, our efforts to determine the appropriate reliability coefficient for the data answers the call from recent work (Komperda, Pentecost, *et al.*, 2018) to *move away from using Cronbach's α as a default measure and to become more purposeful in analytic choices.* Lastly, results showing the relationships between motivation and student learning provide evidence to support further investigations of affective measures and the student experience, and to develop interventions that can help students move towards intrinsic motivation to foster learning and success.

For practitioners, these results add to our *understanding of students* and *understanding of assessment* (Rodriguez and Towns, 2019). First and foremost, these findings support that affect is important for student learning. As such, instruction should emphasize the affective experience of students in promoting student success. Additionally, more than a decade of research shows the positive impacts of active learning teaching approaches on student learning (National Research Council, 2012; Freeman *et al.*, 2014; Wieman, 2014); while not explicitly investigated in this study, relating active learning pedagogies to student motivation provides one lens for interpreting these findings. Active learning promotes student ownership (autonomy), active involvement in knowledge construction (competence), and feeling part of a classroom community (relatedness). These components impact the affective experiences of students (particularly intrinsic motivation) as shown by SDT (Deci and Ryan, 2000; Ryan and Deci, 2000), other learning theories (Bretz, 2001; Novak, 2010; Sousa, 2011), previous work in chemistry (Liu *et al.*, 2017, 2018; Pratt and Raker, 2020) and other disciplines (Lepper and Henderlong, 2000; Chirkov and Ryan, 2001; Reeve, 2012), and this study, and are positively related to student learning. As such, chemistry assessments and evaluations should focus not only on measures of student content knowledge, but also on measures of student experiences. Affective measures can provide evidence to inform teaching decisions, as well as provide evidence in interpreting student success and the effectiveness of instructional practices. By incorporating considerations of student experiences (e.g., motivation) into classroom culture, practices, assessments, etc., chemistry educators can more effectively support chemistry learning.

Limitations

A key limitation of this study is the lack of additional institution and student demographics, as well as course and institution type analyses. Given eighteen institutions/instructors participated in this project, all responses should be considered nested within their unique institutional contexts. However, given only thirteen sites provided AMS-Chemistry data (see Appendix 3), analyses that considered the multilevel nature of the data were limited due to limited variability and small *n* at the institution level. This resulted in a multilevel analysis focused on the entire data corpus where pre & post administrations were collapsed; this calls into question the assumption of data independence for such an analysis. A more thorough analysis would analyze the pre and post administrations individually using a multilevel approach. However, as previously mentioned, sample sizes at individual institutions (i.e., nests) limited that ability. Instead, choosing to conduct the analysis on the data corpus allows us to provide empirical evidence for the imposed factor structure/SDT when taking into consideration the "noise" associated with multiple institutions/nests. It does not allow us to comment on the pre and post administrations nor use the nested data in subsequent analyses (i.e., comparing institutions/nests). Therefore, our multilevel approach, while novel for this *Journal*, should be interpreted cautiously and conservatively. It is only intended to provide further support for the 7-factor structure/Self-determination Theory constructs when considering the additional "noise" or complexity of multiple institutions/nests.

Additionally, previous work has shown that sex differences may exist in relation to motivation (Zhi-ling, 2003; Grouzet *et al.*, 2006; Ackerman *et al.*, 2013; Liu *et al.*, 2017). However, student demographic information was not collected as part of IRB procedures for sharing student data outside of individual institutions; therefore, analyses based on demographics were impractical. Lastly, previous work suggests that foundation-level inorganic chemistry courses are highly varied in terms of content taught (Raker *et al.*, 2015a). As such, analyses should consider the specific types of foundation-level inorganic chemistry courses, adding further complexity and nesting to the analyses. However, given the lack of variability and thirteen sites at the nesting level, there is insufficient power and variability to add additional nesting to the analyses. Despite these limitations, the multiple CFAs conducted have fit indices well within boundaries of good fit despite not accounting for instructional contexts and student demographics. Therefore, these analyses provide support that the data collected in this study fit the seven-factor AMS-chemistry model as conceptualized in SDT for students in foundation-level inorganic chemistry courses, and that the constructs are well-defined and rise above any noise associated with factors such as institution, course, or student demographics not accounted for in the models. While we cannot extrapolate these specific results to all populations of inorganic chemistry students, the findings provide empirical support for the theoretical underpinnings of the work (i.e., SDT) as well as previous findings from other disciplines (e.g., Mercincavage and Brooks, 1990; Zhi-ling, 2003; Cole *et al.*, 2006; Allen and Robbins, 2010; Kirn and Benson, 2013; Shell and Soh, 2013; Wang and Degol, 2013; Komarraju *et al.*, 2014; León *et al.*, 2015). Therefore, the implications of this work are applicable to courses in inorganic chemistry, and to contexts in which student samples primarily consist of chemistry majors in chemistry courses. Future work should consider institution type, course type, and student demographics in analyses to expand the evidence regarding the complex relationships between motivation and student learning. However, it is worth noting that sufficient sample size is a limiting factor for these types of analyses; we suggest conducting more multi-institution studies and/or studies conducted over multiple years at a single institution as ways to mitigate sample size limitations.

Conflicts of Interest

There are no conflicts to declare.

Appendices

Appendix 1: Copy of adapted AMS-Chemistry used in this study

Your instructor, in collaboration with the Interactive Online Network of Inorganic Chemists, is interested in your experience taking inorganic chemistry this term. We ask that you complete the survey below. There are no right or wrong answers. Please be candid and honest in responding to this survey. The information will be used to evaluate the course.

WHY ARE YOU ENROLLED IN THIS CHEMISTRY COURSE?

Using the scale below, indicate to what extent each of the following statements corresponds to one of the reasons why you are enrolled in this inorganic chemistry course.

A = NOT AT ALL   B = A LITTLE   C = MODERATELY   D = A LOT   E = EXACTLY

	WHY ARE YOU ENROLLED IN THIS CHEMISTRY COURSE?	A	B	C	D	E
1	Because without having taken chemistry, I would not find a high-paying job later on	A	B	C	D	E
2	Because I experience pleasure and satisfaction while learning new things	A	B	C	D	E
3	Because I think that chemistry courses will help me better prepare for the career I have chosen	A	B	C	D	E
4	For the feelings I experience when I am communicating chemistry ideas to others	A	B	C	D	E
5	Honestly, I don't know; I really feel that I am wasting my time taking chemistry courses	A	B	C	D	E
6	For the satisfaction I experience while improving my understanding of chemistry	A	B	C	D	E
7	To prove to myself that I am capable of succeeding in chemistry	A	B	C	D	E
8	In order to obtain a better job later on	A	B	C	D	E
9	For the pleasure I experience when I learn new things about chemistry	A	B	C	D	E
10	Because taking chemistry will enable me to enter the job market in a field that I like	A	B	C	D	E
11	For the pleasure that I experience when I perform chemistry experiments	A	B	C	D	E
12	I once had good reasons for taking chemistry courses; however, now I wonder whether I should continue	A	B	C	D	E
13	For the satisfaction I experience while succeeding in chemistry	A	B	C	D	E
14	Because when I succeed in chemistry I feel smart	A	B	C	D	E
15	Because I want to have a well-paying careers	A	B	C	D	E

16	For the pleasure that I experience in broadening my knowledge about chemistry	A	B	C	D	E
17	Because taking chemistry courses will help me make more informed choices about my career options	A	B	C	D	E
18	For the enjoyment I experience when I think about the world in terms of atoms and molecules	A	B	C	D	E
19	I don't know why I take chemistry courses, I couldn't care less about them	A	B	C	D	E
20	For the satisfaction I feel as I work toward an understanding of chemistry	A	B	C	D	E
21	To show myself that I am an intelligent person	A	B	C	D	E
22	In order to have a better salary later on	A	B	C	D	E
23	Because studying chemistry allows me to continue to learn about things that interest me	A	B	C	D	E
24	Because I believe that chemistry courses will improve my skills in my chosen career	A	B	C	D	E
25	For the satisfaction I experience while learning about various chemistry topics	A	B	C	D	E
26	I don't know; I can't understand what I am doing taking chemistry courses	A	B	C	D	E
27	Because chemistry courses allow me to experience satisfaction in my quest for knowledge	A	B	C	D	E
28	Because I want to show myself that I can succeed in studying chemistry	A	B	C	D	E

Appendix 2: Intraclass Coefficients (ICC)

Intraclass coefficients (ICC) for the 28 items of the AMS-Chemistry

Item Number	ICC value
1	0.037
2	0.096
3	0.054
4	0.087
5	0.036
6	0.125
7	0.167
8	0.115
9	0.142
10	0.041
11	0.127
12	0.052
13	0.144
14	0.082
15	0.051
16	0.167
17	0.051
18	0.103
19	0.046
20	0.158
21	0.089
22	0.039
23	0.104
24	0.053
25	0.193

26	0.039
27	0.113
28	0.099

Appendix 3: Sample sizes delineated by institution/data collection site

Sample sizes (*n*) for each data collection delineated by institution/data collection site

Institution	Pre AMS-Chemistry	Post AMS-Chemistry	ACS Exam
1	43	--	39
2	--	--	5
3	20	15	26
4	--	--	--
5	7	6	7
6	14	13	14
7	--	--	--
8	--	--	--
9	--	--	--
10	15	15	15
11	7	7	7
12	12	12	12
13	13	8	13
14	5	5	5
15	7	7	7
16	55	52	58
17	32	--	48
18	13	13	13

Note: "--" indicates data were either not provided or IRB-approval was not obtained

Appendix 4: Results of congeneric factor models for individual subscales/factors of AMS-C

Confirmatory factor analysis fit information for individual, single factor congeneric measurement models using the WLSMV estimator (pre models *n* = 243, post models *n* = 153).

Model	χ^2	*df*	*P*	CFI	TLI	RMSEA	RMSEA 90% Confidence Interval
Pre Amotivation factor	7.140	2	0.028	0.995	0.984	0.103	0.029 – 0.189
Pre Amotivation factor	2.569	2	0.277	0.999	0.998	0.043	0.000 – 0.172
Pre External Regulation	4.811	2	0.090	0.999	0.998	0.076	0.000 – 0.166
Post External Regulation	2.187	2	0.335	1.000	1.000	0.025	0.000 – 0.164
Pre Introjected Regulation	9.241	2	0.010	0.997	0.990	0.122	0.051 – 0.206
Post Introjected Regulation	0.637	2	0.727	1.000	1.002	0.000	0.000 – 0.114
Pre Identified Regulation	16.495	2	< 0.001	0.985	0.956	0.173	0.102 – 0.254
Post Identified Regulation	12.262	2	0.002	0.979	0.936	0.183	0.094 – 0.287
Pre To Know	4.626	2	0.099	0.999	0.997	0.074	0.000 – 0.164
Post To Know	8.142	2	0.017	0.996	0.989	0.142	0.051 – 0.249
Pre To Accomplish	10.137	2	0.006	0.998	0.993	0.129	0.059 – 0.213

Post To Accomplish	4.581	2	0.101	0.999	0.996	0.092	0.000 – 0.206
Pre To Experience	1.630	2	0.443	1.000	1.001	0.000	0.000 – 0.120
Post To Experience	9.174	2	0.010	0.994	0.981	0.153	0.063 – 0.259

Note: cut-off values for indication of good model fit for this study are CFI ≥ 0.98, TLI ≥ 0.98, and RMSEA ≤ 0.07.

Acknowledgments

We would like to thank the students who participated in this study. We would also like to thank the inorganic chemistry faculty members who participated in this project and gave us access to students in their foundation-level inorganic chemistry courses. In addition, we would like to thank the members of the Leadership Council of the Interactive Online Network of Inorganic Chemists (IONiC) for their support of this project. Lastly, we thank J. Ferron (University of South Florida) for feedback and support with the multilevel analyses.

This material is based upon work supported by the National Science Foundation under IUSE DUE-1726162, 1726133, & 1725822. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Ackerman P. L., Kanfer R., and Beier M. E., (2013), Trait Complex, Cognitive Ability, and Domain Knowledge Predictors of Baccalaureate Success, STEM Persistence, and Gender Differences. *J. Educ. Psychol.*, **105**(3), 911–927.

ACS Exams, (2016), *Foundations of Inorganic Chemistry*, American Chemical Society Division of Chemical Education.

Allen D., (1999), Desire to Finish College: An Empirical Link Between Motivation and Persistence. *Res. High. Educ.*, **40**(4), 461–485.

Allen J. and Robbins S., (2010), Effects of Interest–Major Congruence, Motivation, and Academic Performance on Timely Degree Attainment. *J. Couns. Psychol.*, **57**(1), 23–35.

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

American Psychological Association Presidential Task Force on Psychology in Education, (1993), *Learner-Centered Psychological Principles: Guidelines for School Redesign and Reform*, American Psychological Association and Mid-Continent Regional Educational Laboratory.

Arjoon J. A., Xu X., and Lewis J. E., (2013), Understanding the State of the Art for Measurement in Chemistry Education Research: Examining the Psychometric Evidence. *J. Chem. Educ.*, **90**(5), 536–545.

Asparouhov T. and Muthén B., (2019), Nesting and Equivalence Testing for Structural Equation Models. *Struct. Equ. Model. Multidiscip. J.*, **26**(2), 302–309.

Bain K., Moon A., Mack M. R., and Towns M. H., (2014), A review of research on the teaching and learning of thermodynamics at the university level. *Chem Educ Res Pr.*, **15**(3), 320–335.

Bain K. and Towns M. H., (2016), A review of research on the teaching and learning of chemical kinetics. *Chem. Educ. Res. Pract.*, **17**(2), 246–262.

Bandalos D. L., (2008), Is Parceling Really Necessary? A Comparison of Results From Item Parceling and Categorical Variable Methodology. *Struct. Equ. Model. Multidiscip. J.*, **15**(2), 211–240.

Bandalos D. L., (2018), *Measurement Theory and Applications for the Social Science*, Guilford Press. New York. USA.
Bauer C. F., (2005), Beyond "Student Attitudes": Chemistry Self-Concept Inventory for Assessment of the Affective Component of Student Learning. *J. Chem. Educ.*, **82**(12), 1864.
Beauducel A. and Herzberg P. Y., (2006), On the Performance of Maximum Likelihood Versus Means and Variance Adjusted Weighted Least Squares Estimation in CFA. *Struct. Equ. Model. Multidiscip. J.*, **13**(2), 186–203.
Bialosiewicz S., Murphy K., and Berry T., (2013), An Introduction to Measurement Invariance Testing: Resource Packet for Participants. *Am. Eval. Assoc.,* **27**(5) 1-37.
Bishop P. A. and Herron R. L., (2015), Use and Misuse of the Likert Item Responses and Other Ordinal Measures. *Int. J. Exerc. Sci.*, **8**(3), 297–302.
Black A. E. and Deci E. L., (2000), The Effects of Instructors' Autonomy Support and Students' Autonomous Motivation on Learning Organic Chemistry: A Self-Determination Theory perspective. *Sci. Educ.*, **84**(6), 740–756.
Bodner G. M. and Weaver G., (2008), Introduction: Research and Practice in Chemical Education in Advanced Courses. *Chem Educ Res Pr.*, **9**(2), 81–83.
Bretz S. L., (2001), Novak's Theory of Education: Human Constructivism and Meaningful Learning. *J. Chem. Educ.*, **78**(8), 1107–1116.
Cao L., (2012), Differences in Procrastination and Motivation between Undergraduate and Graduate Students. *J. Scholarsh. Teach. Learn.*, **12**(2), 39–64.
Carifio J. and Perla R. J., (2007), Ten Common Misunderstandings, Misconceptions, Persistent Myths and Urban Legends about Likert Scales and Likert Response Formats and their Antidotes. *J. Soc. Sci.*, **3**(3), 106–116.
Chase C. I. and Keene J. M., (1981), Major Declaration and Academic Motivation. *J. Coll. Stud. Pers.*, **22**(6), 496–502.
Chen F. F., (2007), Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Struct. Equ. Model. Multidiscip. J.*, **14**(3), 464–504.
Cheung G. W. and Rensvold R. B., (2002), Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Struct. Equ. Model. Multidiscip. J.*, **9**(2), 233–255.
Chirkov V. I. and Ryan R. M., (2001), Parent and Teacher Autonomy-Support in Russian and U.S. Adolescents. *J. Cross-Cult. Psychol.*, **32**(5), 618–635.
Cole J. S., Cole S. T., and Ferguson A. P., (2006), Students' Motivation to Learn: A Comparison Between Undergraduate Students Majoring in Parks, Recreation, and Tourism and Those in Other Majors. *J. Teach. Travel Tour.*, **6**(2), 61–70.
Coppola B. P., Ege S. N., and Lawton R. G., (1997), The University of Michigan Undergraduate Chemistry Curriculum 2. Instructional Strategies and Assessment. *J. Chem. Educ.*, **74**(1), 84.
Cronbach L. J., (1951), Coefficient alpha and the internal structure of tests. *Psychometrika*, **16**(3), 297–334.
Debacker T. K. and Nelson R. M., (2000), Motivation to Learn Science: Differences Related to Gender, Class Type, and Ability. *J. Educ. Res.*, **93**(4), 245–254.
Deci E. L. and Ryan R. M., (2008), Self-Determination Theory: A Macrotheory of Human Motivation, Development, and Health. *Can. Psychol.*, **49**(3), 182–185.
Deci E. L. and Ryan R. M., (2000), The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychol. Inq.*, **11**(4), 227–268.
Dimitrov D. M., (2010), Testing for Factorial Invariance in the Context of Construct Validation. *Meas. Eval. Couns. Dev.*, **43**(2), 121–149.
DiStefano C. and Morgan G. B., (2014), A Comparison of Diagonal Weighted Least Squares Robust Estimation Techniques for Ordinal Data. *Struct. Equ. Model. Multidiscip. J.*, **21**(3), 425–438.
Dodge Y., (2008), Spearman Rank Correlation Coefficient, in *The Concise Encyclopedia of Statistics*, Springer New York, pp. 502–505.
Dunn O. J., (1961), Multiple Comparisons among Means. *J. Am. Stat. Assoc.*, **56**(293), 52–64.
Dyer N. G., Hanges P. J., and Hall R. J., (2005), Applying multilevel confirmatory factor analysis techniques to the study of leadership. *Leadersh. Q.*, **16**(1), 149–167.
Eberlein T., Kampmeier J., Minderhout V., Moog R. S., Platt T., Varma-Nelson P., and White H. B., (2008), Pedagogies of Engagement in Science. *Biochem. Mol. Biol. Educ.*, **36**(4), 262–273.
Farrell J. J., Moog R. S., and Spencer J. N., (1999), A Guided-Inquiry General Chemistry Course. *J. Chem. Educ.*, **76**(4), 570.
Freeman S., Eddy S. L., McDonough M., Smith M. K., Okoroafor N., Jordt H., and Wenderoth M. P., (2014), Active learning increases student performance in science, engineering, and mathematics. *Proc. Natl. Acad. Sci.*, **111**(23), 8410–8415.
French B. F., Immekus J. C., and Oakes W. C., (2005), An Examination of Indicators of Engineering Students' Success and

Persistence. *J. Eng. Educ.*, **94**(4), 419–425.
Friedman M., (1939), A Correction: The use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance. *J. Am. Stat. Assoc.*, **34**(205), 109–109.
Friedman M., (1937), The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance. *J. Am. Stat. Assoc.*, **32**(200), 675–701.
Gibbons R. E., Xu X., Villafañe S. M., and Raker J. R., (2018), Testing a Reciprocal Causation Model between Anxiety, Enjoyment and Academic Performance in Postsecondary Organic Chemistry. *Educ. Psychol.*, **38**(6), 838–856.
Glynn S. M., Brickman P., Armstrong N., and Taasoobshirazi G., (2011), Science Motivation Questionnaire II: Validation with Science Majors and Nonscience Majors. *J. Res. Sci. Teach.*, **48**(10), 1159–1176.
Grouzet F. M. E., Otis N., and Pelletier L. G., (2006), Longitudinal Cross-Gender Factorial Invariance of the Academic Motivation Scale. *Struct. Equ. Model. Multidiscip. J.*, **13**(1), 73–98.
Grove N. and Bretz S. L., (2007), CHEMX: An Instrument To Assess Students' Cognitive Expectations for Learning Chemistry. *J. Chem. Educ.*, **84**(9), 1524.
Hagger M. S., Koch S., and Chatzisarantis N. L. D., (2015), The Effect of Causality Orientations and Positive Competence-Enhancing Feedback on Intrinsic Motivation: A Test of Additive and Interactive Effects. *Personal. Individ. Differ.*, **72**, 107–111.
Hancock G. R. and An J., (2020), A Closed-Form Alternative for Estimating ω Reliability under Unidimensionality. *Meas. Interdiscip. Res. Perspect.*, **18**(1), 1–14.
Hancock G. R. and An J., (2018), Digital ITEMS Module 2: Scale Reliability in Structural Equation Modeling. *Educ. Meas. Issues Pract.*, **37**(2), 73–74.
Haynes W., (2013), Bonferroni Correction, in *Encyclopedia of Systems Biology*, Dubitzky W., Wolkenhauer O., Cho K.-H., and Yokota H. (eds.)., Springer New York, pp. 154–154.
Hooper D., Coughlan J., and Mullen M., (2008), Structural Equation Modelling: Guidelines for Determining Model Fit. *Electron. J. Bus. Res. Methods*, **6**(1), 53–60.
Howard J. L., Gagné M., and Bureau K. S. (2017). Testing a Continuum Structure of Self-Determined Motivation: A Meta-Analysis. *Psychological Bulletin*, **143**(12), 1346-1377.
Huggins-Manley A. C. and Han H., (2017), Assessing the Sensitivity of Weighted Least Squares Model Fit Indexes to Local Dependence in Item Response Theory Models. *Struct. Equ. Model. Multidiscip. J.*, **24**(3), 331-340.
Jöreskog K. G., (2019), *LISREL 10*, Scientific Software International, Inc. https://ssicentral.com/index.php/products/lisrel/
Kiemer K., Gröschner A., Pehmer A.-K., and Seidel T., (2015), Effects of a Classroom Discourse Intervention on Teachers' Practice and Students' Motivation to Learn Mathematics and Science. *Learn. Instr.*, **35**, 94–103.
Kirn A. and Benson L., (2013), Quantitative Assessment of Student Motivation to Characterize Differences between Engineering Majors, in *2013 IEEE Frontiers in Education Conference (FIE)*. IEEE, pp. 69-74.
Kline R. B., (2016), *Principles and Practice of Structural Equation Modeling*, Guilford Press. New York, USA.
Komarraju M., Swanson J., and Nadler D., (2014), Increased Career Self-Efficacy Predicts College Students' Motivation, and Course and Major Satisfaction. *J. Career Assess.*, **22**(3), 420–432.
Komperda R., Hosbein K. N., and Barbera J., (2018), Evaluation of the Influence of Wording Changes and Course Type on Motivation Instrument Functioning in Chemistry. *Chem. Educ. Res. Pract.*, **19**(1), 184–198.
Komperda R., Pentecost T. C., and Barbera J., (2018), Moving beyond Alpha: A Primer on Alternative Sources of Single-Administration Reliability Evidence for Quantitative Chemistry Education Research. *J. Chem. Educ.*, **95**(9), 1477–1491.
Koziol N. A., (2010), Evaluating Measurement Invariance with Censored Ordinal Data: A Monte Carlo Comparison of Alternative Model Estimators and Scales of Measurement. M.A. Thesis. University of Nebraska – Lincoln.
Lavigne G. L., Vallerand R. J., and Miquelon P., (2007), A Motivational Model of Persistence in Science Education: A Self-Determination Theory Approach. *Eur. J. Psychol. Educ.*, **22**(3), 351–369.
León J., Núñez J. L., and Liew J., (2015), Self-Determination and STEM Education: Effects of Autonomy, Motivation, and Self-Regulated Learning on High School Math Achievement. *Learn. Individ. Differ.*, **43**, 156–163.
Lepper M. R., Corpus J. H., and Iyengar S. S., (2005), Intrinsic and Extrinsic Motivational Orientations in the Classroom: Age Differences and Academic Correlates. *J. Educ. Psychol.*, **97**(2), 184–196.
Lepper M. R. and Henderlong J., (2000), Turning "play" into "work" and "work" into "play," in *Intrinsic and Extrinsic Motivation: The Search for Optimal Motivation and Performance*, Sansone C. and Harackiewicz J. M. (eds.)., Academic Press, pp. 257–307.
Liu Y., Ferrell B., Barbera J., and Lewis J. E., (2017), Development and Evaluation of a Chemistry-Specific Version of the

Academic Motivation Scale (AMS-Chemistry). *Chem. Educ. Res. Pract.*, **18**(1), 191–213.
Liu Y., Raker J. R., and 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. *Chem. Educ. Res. Pract.*, **19**(1), 251–264.
Lynch D. J., (2008), Confronting Challenges: Motivational Beliefs and Learning Strategies in Difficult College Courses. *Coll. Stud. J.*, **42**(2), 416.
Maltese A. V. and Tai R. H., (2011), Pipeline Persistence: Examining the Association of Educational Experiences with Earned Degrees in STEM among U.S. Students. *Sci. Educ.*, **95**(5), 877–907.
McDonald R. P., (1999), *Test Theory: A Unified Treatment*, Lawrence Erlbaum Associates. Mahwah, NJ.
Mercincavage J. E. and Brooks C. I., (1990), Differences in Achievement Motivation of College Business Majors as a Function of Year in College and Classroom Seating Position. *Psychol. Rep.*, **66**(2), 632–634.
Meredith W., (1993), Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, **58**(4), 525–543.
Millsap R. E., (2011), *Statistical Approaches to Measurement Invariance*, Routledge/Taylor & Francis Group. New York, USA.
Morrow J. and Ackermann M., (2012), Intention to Persist and Retention of First-Year Students: The Importance of Motivation and Sense of Belonging. *Coll. Stud. J.*, **46**(3), 483–491.
Mueller R. O. and Hancock G. R., (2008), Best Practices in Structural Equation Modeling, in *Best Practices in Quantitative Methods*, Osborn J. W. (ed.)., Sage Publications, Inc., Thousand Oaks, CA., pp. 488–508.
Muthén L. K. and Muthén B. O., (2017), *Mplus User's Guide: Eighth Edition*, Muthén & Muthén.
National Research Council, (2012), *Discipline-Based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering*, National Academies Press. Washington, DC.
Niemiec C. P. and Ryan R. M., (2009), Autonomy, Competence, and Relatedness in the Classroom: Applying Self-Determination Theory to Educational Practice. *Theory Res. Educ.*, **7**(2), 133–144.
Novak J. D., (2010), Meaningful Learning for Empowerment, in *Learning, Creating, and Using Knowledge*. Routledge, New York, NY, pp. 23–40.
Padgett R. N. and Morgan G. B., (2021), Multilevel CFA with Ordered Categorical Data: A Simulation Study Comparing Fit Indices Across Robust Estimation Methods. *Struct. Equ. Model. Multidiscip. J.*, **28**(1), 51-68.
Parry S., (n.d.), Fit Statistics Commonly Reported for CFA and SEM. *Cornell Stat. Consult. Unit*. Cornell University. https://www.hrstud.unizg.hr/_download/repository/SEM_fit.pdf
Planchard M., Daniel K. L., Maroo J., Mishra C., and McLean T., (2015), Homework, Motivation, and Academic Achievement in a College Genetics Course. *Bioscene J. Coll. Biol. Teach.*, **41**(2), 11–18.
Potvin P. and Hasni A., (2014), Interest, Motivation and Attitude Towards Science and Technology at K-12 Levels: A Systematic Review of 12 years of Educational Research. *Stud. Sci. Educ.*, **50**(1), 85–129.
Pratt J. M. and Raker J. R., (2020), Exploring Student Affective Experiences in Inorganic Chemistry Courses: Understanding Student Anxiety and Enjoyment, in *Advances in Teaching Inorganic Chemistry Volume 1: Classroom Innovations and Faculty Development*, ACS Symposium Series. American Chemical Society, pp. 117–129.
Putnick D. L. and Bornstein M. H., (2016), Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Dev. Rev.*, **41**, 71–90.
R Core Team, (2017), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing.
Raker J. R., Gibbons R. E., and Cruz-Ramírez de Arellano D., (2019), Development and Evaluation of the Organic Chemistry-Specific Achievement Emotions Questionnaire (AEQ-OCHEM). *J. Res. Sci. Teach.*, **56**(2), 163–183.
Raker J. R., Reisner B. A., Smith S. R., Stewart J. L., Crane J. L., Pesterfield L., and Sobel S. G., (2015a), Foundation Coursework in Undergraduate Inorganic Chemistry: Results from a National Survey of Inorganic Chemistry Faculty. *J. Chem. Educ.*, **92**(6), 973–979.
Raker J. R., Reisner B. A., Smith S. R., Stewart J. L., Crane J. L., Pesterfield L., and Sobel S. G., (2015b), In-Depth Coursework in Undergraduate Inorganic Chemistry: Results from a National Survey of Inorganic Chemistry Faculty. *J. Chem. Educ.*, **92**(6), 980–985.
Reeve J., (2012), A Self-determination Theory Perspective on Student Engagement, in *Handbook of Research on Student Engagement*, Christenson S., Reschly A., and Wylie C. (eds.)., Springer, New York, NY, pp. 149–172.
Reiche S., (2013), The Psychology of Expat Motivation: From Theory to Practice. *Expatr. Blog Netw.* https://blog.iese.edu/expatriatus/2013/09/20/the-psychology-of-expat-motivation-from-theory-to-practice/
Reisner B. A., Smith S. R., Stewart J. L., Raker J. R., Crane J. L., Sobel S. G., and Pesterfield L. L., (2015), Great Expectations: Using an Analysis of Current Practices To Propose a Framework for the Undergraduate Inorganic Curriculum. *Inorg. Chem.*, **54**(18), 8859–8868.
Rey D. and Neuhäuser M., (2011), Wilcoxon-Signed-Rank Test, in *International Encyclopedia of Statistical Science*, Springer

Berlin Heidelberg, pp. 1658–1659.
Rocabado G. A., Komperda R., Lewis J. E., and Barbera J., (2020), Addressing diversity and inclusion through group comparisons: a primer on measurement invariance testing. *Chem. Educ. Res. Pract.*, **21**(3), 969–988.
Rodriguez J.-M. G. and Towns M. H., (2019), Alternative Use for the Refined Consensus Model of Pedagogical Content Knowledge: Suggestions for Contextualizing Chemistry Education Research. *J. Chem. Educ.*, **96**(9), 1797–1803.
Rodriguez J.-M. G. and Towns M. H., (2020), Research on Students' Understanding of Michaelis-Menten Kinetics and Enzyme Inhibition: Implications for Instruction and Learning. *The Biophysicist*, **1**(2), 3.
Rosseel Y., (2012), Lavaan: An R Package for Structural Equation Modeling. *J. Stat. Softw.*, **48**(2), 1–36.
Rutkowski L. and Svetina D., (2014), Assessing the Hypothesis of Measurement Invariance in the Context of Large-Scale International Surveys. *Educ. Psychol. Meas.*, **74**(1), 31–57.
Ryan R. M. and Deci E. L., (2000), Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *Am. Psychol.*, **55**(1), 68–78.
Schermelleh-Engel K., Moosbrugger H., and Müller H., (2003), Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. *Methods Psychol. Res.*, **8**(2), 23–74.
Seymour E. and Hunter A.-B., (2019), *Talking about leaving revisited: persistence, relocation, and loss in undergraduate STEM education*, Springer Nature, Switzerland.
Shapiro S. S. and Wilk M. B., (1965), An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, **52**(3–4), 591–611.
Shell D. F. and Soh L.-K., (2013), Profiles of Motivated Self-Regulation in College Computer Science Courses: Differences in Major versus Required Non-Major Courses. *J. Sci. Educ. Technol.*, **22**(6), 899–913.
Soenens B., Sierens E., Vansteenkiste M., Dochy F., and Goossens L., (2012), Psychologically Controlling Teaching: Examining Outcomes, Antecedents, and Mediators. *J. Educ. Psychol.*, **104**(1), 108–120.
Sousa D. A., (2011), How the Brain Processes Information, in *How the Brain Learns*, Sage Publications, Thousand Oaks, CA., pp. 41–60.
Spearman C., (1904), The Proof and Measurement of Association between Two Things. *Am. J. Psychol.*, **15**(1), 72.
Suh Y. and Cho S.-J., (2014), Chi-Square Difference Tests for Detecting Differential Functioning in a Multidimensional IRT Model. *Appl. Psychol. Meas.*, **38**(5), 359–375.
Taasoobshirazi G. and Wang S., (2016), The Performance of the SRMR, RMSEA< CFI, and TLI: An Examination of Sample Size, Path Size, and Degrees of Freedom. *Journal of Applied Quantitative Methods*, **11**(3), 31–39.
Taber K. S., (2018), The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Res. Sci. Educ.*, **48**(6), 1273–1296.
Tseng S.-C. and Tsai C.-C., (2010), Taiwan College Students' Self-Efficacy and Motivation of Learning in Online Peer Assessment Environments. *Internet High. Educ.*, **13**(3), 164–169.
Vaino K., Holbrook J., and Rannikmäe M., (2012), Stimulating Students' Intrinsic Motivation for Learning Chemistry through the Use of Context-Based Learning Modules. *Chem Educ Res Pr.*, **13**(4), 410–419.
Vajargah K. F. and Nikbakht M., (2015), Application REMLModel and determining cut off of ICC by multi-level model based on Markov Chains and simulation in health. *Indian J. Fundam. Appl. Life Sci.*, **5**(S2), 1432–1448.
Vallerand R. J., (1997), Toward A Hierarchical Model of Intrinsic and Extrinsic Motivation. *Advances in Experimental Social Psychology*, **29**, 271–360.
Vallerand R. J. and Blssonnette R., (1992), Intrinsic, Extrinsic, and Amotivational Styles as Predictors of Behavior: A Prospective Study. *J. Pers.*, **60**(3), 599–620.
Vallerand R. J., Pelletier L. G., Blais M. R., Briere N. M., Senecal C., and Vallieres E. F., (1992), The Academic Motivation Scale: A Measure of Intrinsic, Extrinsic, and Amotivation in Education. *Educ. Psychol. Meas.*, **52**(4), 1003–1017.
Vandenberg R. J. and Lance C. E., (2000), A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research. *Organ. Res. Methods*, **3**(1), 4–70.
Vansteenkiste M., Sierens E., Goossens L., Soenens B., Dochy F., Mouratidis A., et al., (2012), Identifying Configurations of Perceived Teacher Autonomy Support and Structure: Associations with Self-Regulated Learning, Motivation and Problem Behavior. *Learn. Instr.*, **22**(6), 431–439.
Villafañe S. M., Xu X., and Raker J. R., (2016), Self-Efficacy and Academic Performance in First-Semester Organic Chemistry: Testing a Model of Reciprocal Causation. *Chem. Educ. Res. Pract.*, **17**(4), 973–984.
Wang M.-T. and Degol J., (2013), Motivational Pathways to STEM Career Choices: Using Expectancy–Value Perspective to Understand Individual and Gender Differences in STEM Fields. *Dev. Rev.*, **33**(4), 304–340.
Wieman C. E., (2014), Large-scale Comparison of Science Teaching Methods Sends Clear Message. *Proc. Natl. Acad. Sci.*,

111(23), 8319–8320.

Wilcoxon F., (1945), Individual Comparisons by Ranking Methods. *Biom. Bull.*, **1**(6), 80–83.

Yu C.-Y., (2002), *Evaluating Cuttoff Criteria of Model Fit Indices for Latent Variable Models with Binary and Continuous Outcomes*, Doctoral Dissertation, University of California, Los Angeles, CA.

Zhi-ling X. I. A. O., (2003), Difference Research for Achievement Motivation of College Students of Different Sex Grades and Majors. *J. Hubei Polytech. Univ.*, **1**.

Zimmerman B. J., (2000), Attaining Self-Regulation, in *Handbook of Self-Regulation*, Boekaerts M., Pintrich P. R., and Zeidner M. (eds.)., Academic Press, San Diego, CA., pp. 13–39.

Zinbarg R. E., Revelle W., Yovel I., and Li W., (2005), Cronbach's α, Revelle's β, and Mcdonald's ω: Their Relations with Each Other and Two Alternative Conceptualizations of Reliability. *Psychometrika*, **70**(1), 123–133.