JOURNAL OF CHEMICAL EDUCATION

Article

pubs.acs.org/jchemeduc

Differential Use of Study Approaches by Students of Different Achievement Levels

Diane M. Bunce,*,† Regis Komperda,‡ Maria J. Schroeder,§ Debra K. Dillner,§ Shirley Lin,§ Melonie A. Teichert,§ and JudithAnn R. Hartman§

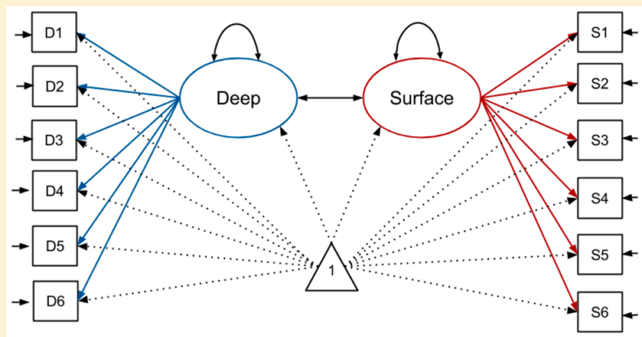
†The Catholic University of America, Washington, D.C. 20064, United States

‡Portland State University, Portland, Oregon 97207, United States

§United States Naval Academy, Annapolis, Maryland 21402, United States

Supporting Information

ABSTRACT: This study examined similarities and differences in study approaches reported by general chemistry students performing at different achievement levels. The study population consisted of freshmen enrolled in a required year-long general chemistry course at the U.S. Naval Academy. Students in the first and second semesters of the course were surveyed using a modified version of the published Approaches and Study Skills Inventory for Students (ASSIST) referred to as the M-ASSIST (Modified Approaches and Study Skills Inventory). Responses to items associated with using deep or surface approaches to studying were examined for students of three achievement levels (A/B, C, and D/F course grades) using both ANOVA and Structured Means Modeling to look for differences in study approaches between achievement levels. Results show that, with only 12 items, the M-ASSIST can be used to measure differences in reported use of deep and surface approaches by students in different achievement groups; that Structured Means Modeling can uncover significant differences that are not apparent with an ANOVA analysis of the same data; and that A/B and D/F students can be classified as reporting using either using primarily deep (A/B students) or primarily surface (D/F) study approaches. C students reported study approaches characteristic of both the A/B and D/F groups, leading to the interpretation that C students may be in an intermediate and possibly transitional state between the higher- and lower-grade groups. These results suggest a new understanding of C students as those who may not fully implement deep approaches to studying but, in general, demonstrate less reliance on surface approaches than lower-achieving students.



KEYWORDS: *First Year Undergraduate/General, Chemical Education Research, Testing/Assessment, Student-Centered Learning*

FEATURE: Chemical Education Research

INTRODUCTION AND THEORETICAL FRAMEWORK

Understanding the different ways students learn can be traced back to Ausubel's[1] work on meaningful and rote learning. Meaningful learning is defined by Ausubel as relating a concept to other concepts in an individual's cognitive structure while rote learning is a condition where concepts are learned in isolation.[2] In the 1970s, Svensson[3,4] suggested that there were implications for Ausubel's two levels of learning in terms of achievement. Chan and Bauer's[5] research related affective and metacognitive characteristics with higher exam performance. On this basis, approaches which can be classified as deep and surface can now be explained both in terms of the cognitive meaningful and rote learning, and affective and metacognitive characteristics of learners such as students evaluating their understanding throughout the learning process.

The distinction between deep and surface approaches to learning has been investigated both qualitatively through interviews and quantitatively through the development and testing of survey instruments.[6−10] Several inventories targeting students' learning or study approaches exist and have been published and refined over time, including the Approaches and Study Skills Inventory (ASSIST),[11] the Study Process Questionnaire (SPQ),[12] and the Motivated Strategies for Learning Questionnaire (MSLQ).[13] A common feature among all these instruments is a conceptual framework that includes some connection to deep and surface approaches to learning.[10,14]

A deep approach describes learners who demonstrate intrinsic motivation, attempt to understand the underlying meaning of a problem, and generate new connections between the ideas presented in the problem and what is already understood from previous tasks. A surface approach is

Received: March 17, 2017
Revised: July 26, 2017
Published: August 15, 2017

ACS Publications © 2017 American Chemical Society and Division of Chemical Education, Inc.

DOI: 10.1021/acs.jchemed.7b00202
J. Chem. Educ. 2017, 94, 1415−1424

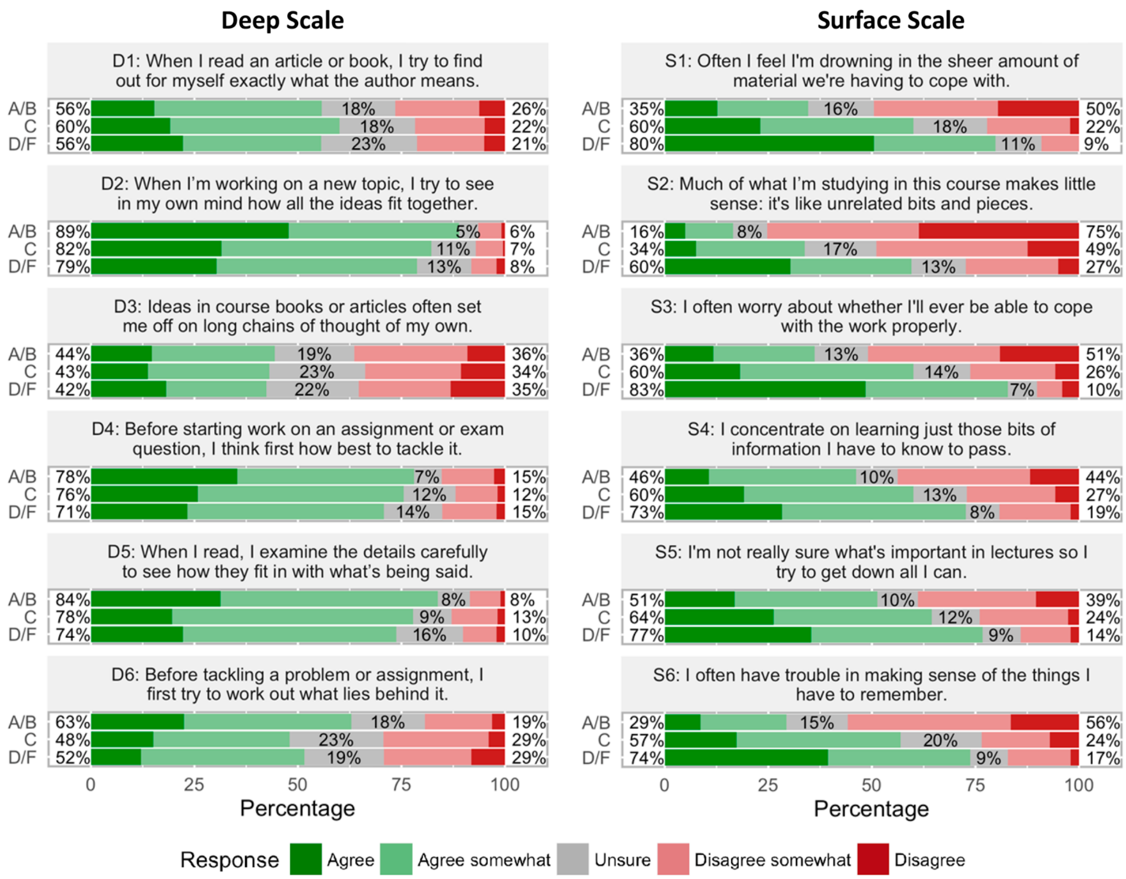


Figure 1. Distribution of responses to deep (D) and surface (S) scale items of the M-ASSIST in 1st semester general chemistry.

associated with extrinsic motivation and is typically concerned with the use of memorized facts and previously seen algorithms to solve a problem. Surface learners attempt to match the attributes of a new problem to those previously seen without analyzing what might be unique to the new problem. This division between deep and surface approaches to learning may become more pronounced at the college level where grades in large classes are much more dependent on exam scores than they typically are in high school.[15]

Providing data on how to identify the approach used by students through a simple method such as a survey and its relation to course achievement is a worthwhile preliminary step before attempting to influence students' learning and studying approaches. Marton and Säljö[7,16] assert that the uses of deep and surface study approaches are fluid, with deep learners sometimes reverting to surface study approaches based on the expectations of the course. This leads to the question of whether students adopting a surface approach can gain knowledge of and begin to practice deep approaches after both specific exposure to deep approaches in the course and the level of learning expected in the course. Sinapuelas and Stacy[6] support this position through their research showing that students can move toward deep approaches when course expectations, conveyed by the type of exam, supported such a move. Christian and Talenquer[17] further demonstrate that trying to develop a deep approach in students is not effective if questions used in class or on achievement measures can be answered through surface approaches. In this study, we limited our work to whether students could be identified as reporting using either a deep or surface approach through a self-report instrument and if the choice of study approach was related to course grade.

In order to more fully understand the difference between deep and surface approaches, we used the levels of understanding developed by Sinapuelas and Stacy[6] and used by Ye et al.[18] to guide our interpretation of the data. An advantage of using these levels of understanding to differentiate between deep and surface approaches is that they can also be used to characterize the levels of metacognition demonstrated at each level. Metacognition is commonly understood in the literature[19] as the knowledge and regulation of one's cognitive system. We interpret this definition along with Sinapuelas and Stacy[6] as students' understanding of and actions taken in their study approach. These four levels were related to two categories of study approaches (deep and surface). Typically, level 1 students are overly concerned with "getting the answer". These students equate time they spend studying with how they expect to perform on an exam. Level 2 students memorize the procedures or algorithms for solving problems from external sources such as notes or books. Level 3 students work out some problems on their own and collaborate with peers to try out ideas and strengthen their own understanding while level 4 students often act as "teachers" of their peers and monitor/evaluate their understanding. This body of work on deep and surface learning helps frame the questions of this study regarding the relationship between study approach and course achievement, as measured by course grade.

■ RESEARCH QUESTIONS

This research focuses on three questions:

DOI: 10.1021/acs.jchemed.7b00202
J. Chem. Educ. 2017, 94, 1415−1424

1. Is a modified version of the Approaches and Study Skills Inventory for students (M-ASSIST) able to categorize the study approaches used by general chemistry students?
2. Are there differential approaches to studying used by successful (A/B) students versus unsuccessful (D/F) students?
3. Which study approaches do C students use, and how does this compare to the study approaches of the other achievement groups?

METHODOLOGY

Population

This research was conducted at the U.S. Naval Academy (USNA), a selective undergraduate institution of approximately 4500 students who will become officers in the U.S. Navy and Marine Corps upon graduation. The average Math SAT score for students in general chemistry during the academic year of this study was 656 (SD = 81). Students at USNA have little time to study due to their daily schedule which includes 18 academic credit hours, military training, and daily structured time for physical activities. A sample schedule is included in the Supporting Information.

All students participating in the research were enrolled in the full year general chemistry sequence in which each semester is a 4 credit-hour course consisting of 3 h of lecture and 2 h of lab weekly. In the first semester (Fall 2014) there were 990 students who completed the course enrolled in 53 sections of the course taught by 29 instructors using a common syllabus. There were two common multiple-choice exams administered at 6 weeks and 12 weeks during the semester, a third instructor-written exam administered at 16 weeks, and a common multiple-choice final exam. In the second semester (Spring 2015) there were 1005 students who completed a similarly structured course. The enrollment is higher in the second semester to accommodate the freshmen who scored high enough on an entrance exam to place out of the first semester of general chemistry.

Instrument

The 67-item ASSIST[11] formed the basis for the shorter inventory used in the current research to identify the study approaches adopted by USNA general chemistry students. Modifications to the ASSIST included changing the wording of some items to be more in line with American English (e.g., replacing "revising for exams" with "studying for exams") and focusing on only the 12 items most relevant to understanding students' adoption of deep and surface approaches (six deep and six surface), based on the results of Confirmatory Factor Analysis (CFA). The original five-point Likert response scale of the ASSIST was retained. This modified instrument is referred to as the M-ASSIST (Modified Approaches and Study Skill Inventory for Students), and the wording of the individual deep and surface items can be found in Figure 1.

Data Collection

Approval for the overall project was obtained from the USNA Institutional Review Board (Approval #USNA.2015.00001-IR-EP7-A). Student completion of the M-ASSIST was included as part of the Chemistry Department's end-of-course evaluation and administered online using Google Forms. A link to this survey was placed in the online homework system used by all instructors and students during the last week of class. The first page of the M-ASSIST described the research study and asked students to provide consent to access their course grade at the end of the semester. From this data set, duplicate responses and those of students who did not grant permission to release final course grades were removed, resulting in 774 (78%) useable responses for the first semester and 717 (71%) for the second semester. Since this study was an analysis of how students describe their approach to studying without an intervention to address reported differences, there was no control used in this study.

ANALYSIS AND RESULTS

Descriptive Statistics

Students were placed into achievement groups based on their final course letter grade in each semester (Table 1). According to Lewis and Lewis,[20] there are advantages and disadvantages to choosing either exam grades or course grades as the achievement measure for a study. Final course grades in this study include grades from the previously described common exams and a common final exam, as well as grades from instructor-written tests and quizzes which included both multiple-choice and free-response questions, lab reports, and online homework. For USNA students, course grades reflect a more consistent level of achievement over an extended period of time and compensate for possible student decisions regarding on which final exam they choose to spend their limited study time at the end of the semester.

Table 1. Achievement Group by Semester

Group	1st Semester (N = 774)	2nd Semester (N = 717)
A/B	450	345
C	225	254
D/F	99	118

Response patterns by achievement group for the six M-ASSIST items on each study approach scale, deep (D) and surface (S), in both semesters were plotted using the Likert package[21] for R (version 3.3.2)[22] to allow for visual examination. Response patterns for first semester are shown in Figure 1. The second semester responses followed similar patterns and are provided in the Supporting Information along with more detailed descriptive statistics including item means and covariances. In Figure 1, the achievement groups (A/B, C, and D/F) are on the left side of the results for each question. The categories of Agree and Agree Somewhat are shown on the left while Disagree Somewhat and Disagree are shown on the right. The Unsure category is in the middle. Percentages for the students agreeing, disagreeing, or unsure are written either to the left, right, or on the bar, respectively.

An examination of Figure 1 shows that, in general, A/B students report most agreement with the deep scale items, the D/F students report the most disagreement with these items, and the C students are in between. For the surface items, A/B students agree the least while D/F students agree the most and C students again report intermediate levels of agreement.

Confirmatory Factor Analysis

Prior to looking for differences across achievement groups on the deep and surface scales, the responses for each semester were examined using confirmatory factor analysis (CFA) to test the fit of the data to a two-factor model of deep and surface approaches to studying (Figure 2). CFA is a standard statistical

DOI: 10.1021/acs.jchemed.7b00202
J. Chem. Educ. 2017, 94, 1415–1424

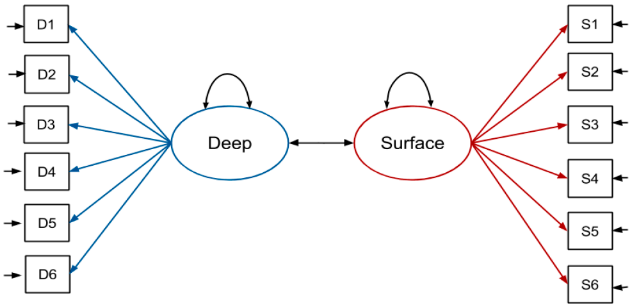


Figure 2. Two-factor model of the M-ASSIST instrument.

procedure used to establish whether items on an instrument fit a previously hypothesized structure of relationships between individual items and the underlying factor the items are designed to measure. The factor is often referred to as a latent variable because it is not measured directly.

The two factors representing approaches to studying (deep and surface) are represented with ovals in Figure 2 and were hypothesized to correlate with each other, represented by the double-headed arrow connecting the two factors. The bent double-headed arrows on each factor represent the variance of each factor. Each of the two factors was also hypothesized to relate to six indicator variables, represented by squares, which are the individual items on the M-ASSIST that were written to relate to either deep or surface study approaches. These items are specified in Figures 1 and 2 as "D1" representing item 1 on the deep scale of the M-ASSIST, "S1" representing item 1 on the surface scale, etc. Within the CFA model, the relationship between each factor and its indicator variables is represented by the single-headed arrows pointing away from each factor toward the indicator item and is referred to as a loading. The other single-headed arrow entering each indicator variable represents the variance of each item that is unexplained by the factor, some of which may be due to measurement error. The numeric values of the loadings and other model parameters for each semester are provided in the Supporting Information.

Overall data-model fit information for the CFA performed with each semester's responses is provided in Table 2 and discussed following the table. All factor analysis was performed in R using the lavaan package,[23] and the code used in the analysis is provided in the Supporting Information. Given the inherently non-normal distribution of the five-point ordinal response scale used for the M-ASSIST, all analyses were executed using robust maximum likelihood estimation with the Satorra–Bentler scaled χ^2 test statistic.[24] All responses contained complete response data; therefore, no cases were removed during analysis.

Table 2. Data-Model Fit for Two-Factor M-ASSIST Model

	Scaled χ^2	df	CFI	RMSEA (90% CI)	SRMR
1st semester (N = 774)	148.52[a]	53	0.95	0.05 (0.04; 0.06)	0.05
2nd semester (N = 717)	126.97[a]	53	0.96	0.04 (0.04; 0.05)	0.04

[a] $p < 0.01$.

Overall, the data from both semesters (Table 2) showed good fit with the two-factor model (deep and surface) of the M-ASSIST. This goodness of fit is supported by literature recommending that the comparative fit index (CFI) values be greater than or equal to 0.95 while root-mean-square error of approximation (RMSEA) values and their associated 90% confidence intervals are below 0.06 and standardized root-mean-square residual (SRMR) values are below 0.08.[25] While the scaled χ^2 values were highly significant for both semesters, χ^2 values are sensitive to sample size so that large samples can lead to larger and therefore more significant χ^2 without necessarily indicating a poor fitting model.[26] The relationship between each of the 12 items of the M-ASSIST and the two factors representing study approaches is discussed in greater detail later in the context of fitting the model to data from students in each grade group. Obtaining good data-model fit values in both semesters of M-ASSIST administration provided evidence for the validity of the data obtained from the instrument[27,28] and therefore supported conducting additional analyses of the data obtained with the M-ASSIST.

Group Comparisons

Comparisons of student study approaches across achievement groups were conducted using two different statistical approaches. First, ANOVA was used to compare average scale scores across achievement groups. With this methodology, each M-ASSIST item contributes equally to its respective scale score (deep or surface). While conducting an ANOVA with scale scores is a common approach,[29−31] it does not take into account the factor structure of the instrument in which each item may not be associated with its respective factor to the same degree due to differences in the values of the loadings. In addition to assuming each item is an equivalent measure of the factor, calculating scale scores assumes each item contains an equal amount of variance due to measurement error. This assumption is equivalent to a level of strict measurement invariance that is unlikely to be true in practice, and this measurement error then becomes incorporated into the composite scale score. To address this limitation, structured means modeling (SMM) was used to provide a latent variable approach to an ANOVA in which group comparisons are made across theoretically error-free latent variable means rather than composite scale means.[32,33] In this study, the results from these two analysis methodologies were compared to look for similarities in conclusions and to identify potential benefits of using SMM with data collected using an instrument with a defined factor structure.

■ ANOVA

Two composite scale means were calculated for each student by computing an average score for the six items on each scale (deep and surface) with Disagree coded as 1 and Agree as 5. Separate ANOVAs were conducted for the deep and surface scales in each semester providing four comparisons in total. All four scales had approximately normal distributions of means. Levene's test indicated that in both semesters the surface scale variances were not equal across groups ($p < 0.05$), so the results of Welch's test are reported for the surface scale to account for the violation of the homogeneity of variance. All posthoc comparisons were completed using the Games−Howell technique to account for the differences in variance and sample size across groups.[34] In all four cases, the ANOVA results indicated significant overall differences in group means for the three achievement (A/B, C, D/F) groups (Table 3).

The effect sizes of the ANOVA results show that the differences in scale means between achievement groups were smaller for the deep scale (η^2 = 0.01 and 0.02) than the surface

DOI: 10.1021/acs.jchemed.7b00202
J. Chem. Educ. 2017, 94, 1415−1424

Table 3. ANOVA Results

Scale	Semester	F	Effect Size (η^2)
Deep	1st	$F_{(2,771)} = 4.11^a$	0.01
	2nd	$F_{(2,714)} = 7.92^a$	0.02
Surface	1st	$F_{(2,274.12)} = 119.74^a$	0.22
	2nd	$F_{(2,348.46)} = 136.84^a$	0.27

[a] $p < 0.05$.

scale (η^2 = 0.22 and 0.27). These results are visualized in Figure 3 with scale means and associated 95% confidence intervals.

Though the omnibus ANOVA indicated significant group differences for the three achievement groups on the deep scale in the first semester, the posthoc comparisons did not detect any significant pairwise differences between achievement groups on the deep scale ($p > 0.05$). This is consistent with the overlapping confidence intervals of the achievement groups in Figure 3 for the first semester deep scale means. Posthoc comparisons for the second semester showed that students earning A/B course grades had significantly higher ($p < 0.05$) mean scores on the deep scale than students earning C or D/F grades. However, the C and D/F grade groups showed no significant difference in their deep scale means ($p > 0.05$). On the surface scale in Figure 3, all three groups had significantly different means in both semesters ($p < 0.05$).

Structured Means Modeling

An alternative approach to analyzing the data using ANOVA is to use SMM. Figure 4 shows the two-factor CFA model of the M-ASSIST from Figure 2 with the addition of a mean structure using dotted lines and a central triangle identified with the number 1 to indicate a constant.[35] The mean structure is always present within the CFA framework, but is not always an explicit part of the analysis. In SMM, the mean structure is explicitly analyzed to provide information about differences in factor means across groups. Mathematically, the mean structure is related to the intercept term of the regression equation connecting each item and its associated factor. In the regression equation, the slope parameter is symbolized by the solid single-headed arrows directed from the factors to the individual items

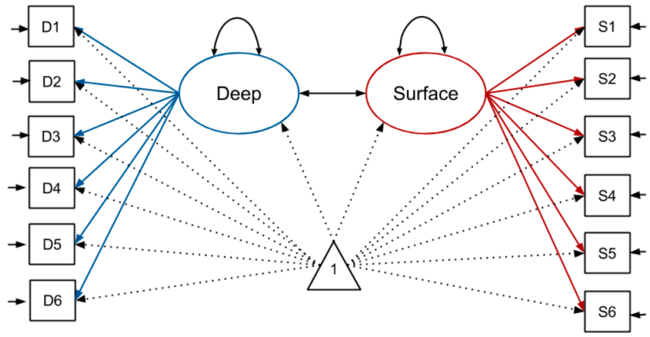


Figure 4. Two-factor model of the M-ASSIST instrument with mean structure shown.

(the item loadings). By setting the mean of the factor to zero, the mean of each item is represented by the intercept parameter, symbolized by the dashed arrow leading from the constant to each item. Analyzing the mean structure for the model in Figure 4 for each achievement group allows for a comparison of factor means across groups.[36] In this way, SMM provides a latent variable approach to ANOVA-type comparisons. The benefit to the latent variable approach is that the factor mean does not contain the measurement error associated with creating composite scale means since the unique variance of each item (the solid arrow entering each item that does not originate from the factor) remains separated from the factors themselves.

The fit information in Table 2 demonstrates that the two-factor M-ASSIST model is a good fit for the data from each semester when all achievement groups are treated as a single data set. To use SMM, the data from each of the three achievement groups (A/B, C, D/F) must also separately show a good fit to the model to establish that students in each group respond to the items in a way that is consistent with the factor structure of the instrument (Table 4). This level of consistency is known as configural invariance and is necessary to demonstrate that the instrument is measuring the same factors (deep and surface) for each group. The assumption of measurement invariance is met by constraining both the

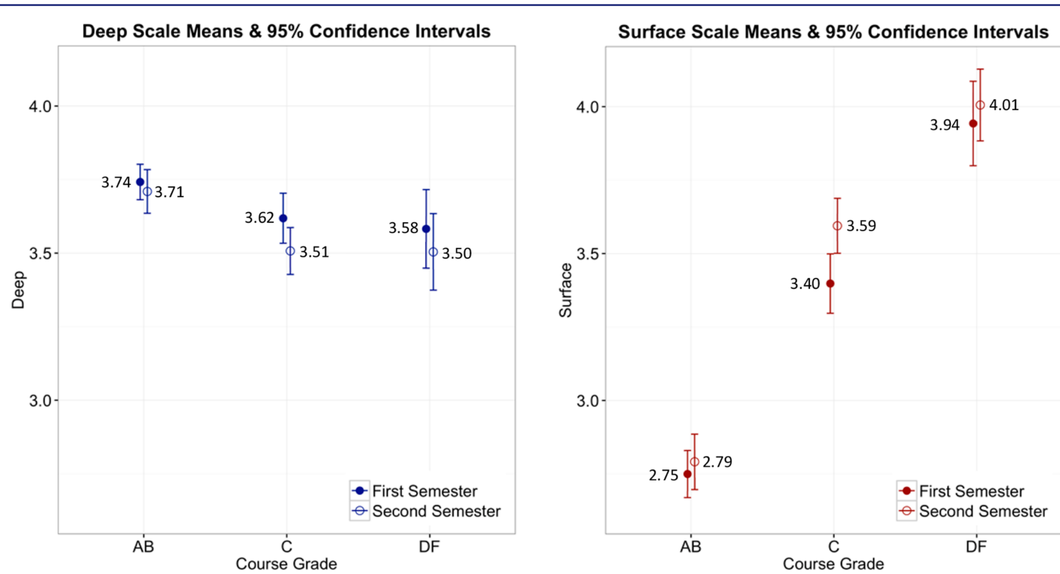


Figure 3. Composite scale means plot for deep and surface scales in 1st and 2nd semester. Larger mean values indicate more agreement with items on that scale.

DOI: 10.1021/acs.jchemed.7b00202
J. Chem. Educ. 2017, 94, 1415−1424

Table 4. Data-Model Fit for Invariance Testing of Two-Factor M-ASSIST

Semester	Model	Scaled χ^2	df	CFI	RMSEA (90% CI)	SRMR
1st	Configural invariance	270.33[a]	159	0.93	0.05 (0.04; 0.06)	0.06
	Measurement invariance	335.01[a]	199	0.91	0.05 (0.04; 0.06)	0.07
2nd	Configural invariance	217.54[a]	159	0.96	0.04 (0.03; 0.05)	0.05
	Measurement invariance	281.48[a]	199	0.95	0.04 (0.03; 0.05)	0.06

[a] $p < 0.01$.

loadings and intercepts of each item to be equal for each of the three achievement groups. Demonstrating this strong level of measurement invariance is necessary to show that the instrument is measuring the factors in the same way for each achievement group, and therefore, comparisons of factor means across achievement groups are meaningful; thus, any differences in factor means are a result of true differences across groups, not differences in instrument functioning (Table 4). The R code for this analysis is provided in the Supporting Information.

As before, the data-model fit was borderline acceptable for the first semester data with a somewhat low CFI but acceptable RMSEA and SRMR values. Due to the borderline acceptance of the first semester data, the analysis of the second semester M-ASSIST data provided a check on the validity of the results for first semester. All three fit indices proved acceptable for the second semester data (Table 4), which supported the use of M-ASSIST data in SMM. The very small change in fit indices under the strict measurement invariance conditions in the second semester also lends support for the interpretation that the instrument functions similarly across the three achievement groups, and therefore, factor means (deep and surface) can be compared across achievement groups.[37]

Table 5 provides values for the item loadings and intercepts (means), which were constrained to be the same value for all the achievement groups, for the six deep items and six surface items on the M-ASSIST in each semester. Comparing loadings and intercepts across semesters shows that the values are similar, supporting the acceptability of using the M-ASSIST to measure differences in study approaches for general chemistry students at USNA due to consistency in the functioning of the M-ASSIST. The covariances between the deep and surface factors, symbolized by the two-headed arrow connecting the two factors in Figure 4, were not constrained to be equal for all the achievement groups and are also reported in Table 5. The only significant covariance between the factors representing deep and surface study approaches was for the A/B students in second semester. This value of −0.11 indicates that, for these students, there was an inverse relationship between reported use of deep and surface study approaches. Since these students were more likely to report using deep approaches to studying, they were also less likely to report using surface study approaches. None of the other groups showed a significant relationship between reported use of the two study approaches demonstrating that an increase in one approach did not result in an increase or decrease in the other approach.

Table 5. Unstandardized Loadings and Intercepts of Deep (D) and Surface (S) Items and Factor Covariances[a]

	1st Semester		2nd Semester	
Item	Loading	Intercept	Loading	Intercept
D1	1	3.51[b]	1	3.54[b]
D2	1.03[b]	4.26[b]	1.15[b]	4.14[b]
D3	0.90[b]	3.19[b]	0.91[b]	3.10[b]
D4	1.09[b]	3.98[b]	1.17[b]	4.05[b]
D5	1.13[b]	4.04[b]	1.28[b]	3.97[b]
D6	1.44[b]	3.59[b]	1.37[b]	3.55[b]
S1	1	2.78[b]	1	2.78[b]
S2	0.83[b]	2.12[b]	0.86[b]	2.44[b]
S3	0.99[b]	2.73[b]	1.00[b]	2.60[b]
S4	0.56[b]	3.01[b]	0.49[b]	3.21[b]
S5	0.58[b]	3.17[b]	0.51[b]	3.10[b]
S6	0.85[b]	2.69[b]	0.87[b]	2.67[b]

Deep−Surface Covariance	1st Semester	2nd Semester
A/B Group	−0.05	−0.11[b]
C Group	−0.05	0.00
D/F Group	0.02	0.03

[a] Scale values: 1 = Disagree to 5 = Agree. [b] $p < 0.01$.

Unlike the composite means used in ANOVA comparisons, SMM compares latent variables which are not measured directly and therefore have no reference scale. Therefore, it is not possible to report an isolated value for each achievement group's factor mean. Instead, the means are reported with respect to a reference level, set at zero. These factor mean comparisons, which are different from the covariances reported in Table 5, are given in Table 6. For each comparison in Table 6, the second achievement group listed is the reference group. As an example, in the first semester students in the D/F group had a latent mean for the deep factor that was lower than students in the A/B group by 0.18. Since the D/F students were lower than the reference group (A/B), this mean difference is reported as −0.18. The statistical significance of the mean difference (p value) is calculated by considering the difference from 0 as a z score based on the variance of the factor. Table 6 shows that there is a significant difference between the deep factor between A/B and D/F students and again between A/B and C students. There is no significant difference in the deep factor between C and D/F students. On the surface factor, there are significant differences between all three achievement groups. These patterns are consistent for both semesters.

Table 6. Comparison of Structured Means Modeling Factor Mean Differences

		1st Semester		2nd Semester	
Factor	Comparison	Mean Difference	Effect Size	Mean Difference	Effect Size
Deep	D/F to A/B	−0.18[a]	0.39	−0.21[a]	0.41
	D/F to C	−0.03	0.07	−0.02	0.03
	C to A/B	−0.15[a]	0.32	−0.20[b]	0.40
Surface	D/F to A/B	1.47[b]	1.61	1.52[b]	1.59
	D/F to C	0.67[b]	0.83	0.52[b]	0.66
	C to A/B	0.80[b]	0.89	1.00[b]	1.05

[a] $p < 0.01$. [b] $p < 0.001$.

DOI: 10.1021/acs.jchemed.7b00202
J. Chem. Educ. 2017, 94, 1415−1424

The effect size (Table 6) is calculated as the absolute value of the factor mean difference divided by the pooled variance of the factors. Though this effect size is calculated similarly to a Cohen's *d*, its magnitude cannot be interpreted on the same scale. Since the factor is free from measurement error, it is generally accepted that the effect size for factor mean differences should be larger than corresponding effect sizes for measured variables. Therefore, the effect size of 0.39 found between D/F and A/B students on the deep factor in first semester would likely be interpreted as a small effect size[38] in the context of SMM.

The SMM results in Table 6 are similar to the ANOVA results in Table 3, but provide more clarity in identifying differences between student achievement groups with respect to the deep approach. The latent means of both the D/F and the C students on the deep factor were statistically equivalent to each other, yet significantly lower than the mean for the A/B students in both semesters. However, this mean difference for the A/B students on the deep factor compared to both the C and D/F students was relatively small with a small effect size. The use of SMM brings to light significant differences between achievement groups on the deep factor that were hidden in the ANOVA comparison.

In contrast to the deep factor, the results for the surface factor showed significant differences between all three achievement groups with medium to large effect sizes. The latent means for the D/F students on the surface factor were significantly larger than means for the A/B students with a large effect size in both semesters. The C students had significantly larger means on the surface factor than the A/B students, but significantly lower means than the D/F students with medium to large effect sizes.

DISCUSSION

M-ASSIST

One of the goals of this research was to find an effective and relatively simple way to characterize student study approaches and to see if study approaches had a relationship with achievement, as defined by final course grades. Evidence was provided for the validity of the M-ASSIST data based on CFA of responses from administration in two semesters of general chemistry (Fall and Spring) at USNA. The results of SMM and, to a lesser degree, ANOVA demonstrated that student responses to only 12 Likert-type questions on the M-ASSIST, representative of deep and surface approaches to studying, were significantly different between A/B, C, and D/F students at USNA. Though further testing is needed with other student populations, the brevity of the M-ASSIST potentially makes it a convenient and efficient tool for measuring students' self-reported use of deep and surface study approaches in both classroom and research settings, especially compared to the longer ASSIST and other instruments and techniques reported in the literature.

Structured Means Modeling versus ANOVA for Group Comparisons

Consistent with other research, our results indicate subtle but significant differences in the study approaches across achievement groups. The use of SMM was more effective than ANOVA at discerning these differences. Group comparisons with ANOVA and SMM demonstrated similar patterns of self-reported study approach across achievement groups (Table 3, Figure 3, and Table 6). Both analysis methods identified large differences in means of variables representing surface approaches across achievement groups in both semesters as well as the difference between means of variables representing deep study approaches for A/B and D/F students in the second semester. However, while the omnibus ANOVA indicated differences in the deep study approach scale means across first semester achievement groups (Table 3), the posthoc tests were not sensitive enough to detect any pairwise group differences. In contrast, SMM (Table 6) demonstrated that A/B students in the first semester had higher means on the latent variable representing deep study approaches than either the C or D/F students. This same effect was seen in the second semester.

These results highlight the benefit of SMM to detect small group differences in latent variable means when compared to using an ANOVA to detect small differences in composite scale scores created by averaging (or summing) items on a scale. Simply averaging (or summing) items allows measurement error to become part of the composite score, and this measurement error can obscure group differences. When the items have a good fit to the underlying instrument model, latent variables provide a better measurement of an underlying construct by partitioning out measurement error. Therefore, in situations where an instrument has a clear factor structure, SMM may provide important additional information beyond an ANOVA.

Differential Use of Study Approaches

The results of the model used in this research indicate that A/B students generally report more agreement with items representing deep study approaches than the other achievement groups and as a result have larger latent means for the deep factor. C students report agreement with the deep approach items to a degree more similar to D/F students than to A/B students and also have statistically similar latent means to the D/F group. For the surface approach, all three groups report significantly different levels of agreement with items representing the surface factor, resulting in A/B students having the smallest latent means and D/F students the largest. C students had latent means representing agreement with a surface approach intermediate between that of the A/B and D/F students. Similar patterns were seen in both the first and second semesters of general chemistry with the reported mean of the surface factor for C students increasing slightly during the second semester. This observed increase may be a result of the small (5%) increase in the number of C students in second semester due to the difference in difficulty of topics taught in second semester which typically include more abstract topics which may or may not have been taught in high school.

The results also indicate that a dichotomous grouping of purely deep and surface approaches[6,18] may be too simplistic since it appears that students can agree with items representing adoption of both types of approaches in varying proportions. Our results suggest that higher-performing students are more likely to report agreement with items representing adoption of deep approaches and less likely to agree with items representing adoption of surface approaches while lower-performing students are more likely to agree with items representing adoption of surface approaches and less likely to agree with items representing adoption of deep approaches. However, students in the middle achievement group respond to items representing an adoption of deep approaches similarly to lower-performing students while simultaneously agreeing less with items representing surface approaches than these students. This

DOI: 10.1021/acs.jchemed.7b00202
J. Chem. Educ. 2017, 94, 1415−1424

C group might then be described as intermediate in that there is less agreement with items representing surface approaches but not as much agreement with items representing deep approaches compared to the A/B students. This could represent C students as intermediate or transitional between the other two achievement groups.

Earlier research utilizing a survey of student study resources and interviews with students in general chemistry[39] demonstrated that students earning different final course grades utilized different types of study resources. High-performing students (A/B) were likely to rely on resources that they used individually. These students asked for help from others (usually peers) only when they tried several approaches to solving problems and were still unsuccessful. Similar patterns were identified by Sinapuelas and Stacy[6] during interviews with students in an introductory chemistry course. These successful students can be described as adopting deep approaches to studying. By contrast, both Bunce et al.[39] and Sinapuelas and Stacy found that lower-performing students (D/F) were likely to study by looking for answers from others or memorizing facts and/or algorithmic problem solving strategies. These behaviors are characteristic of surface approaches to studying. Intermediate students (C) were found to utilize study resources similar to both the independent approaches of the successful students (initially trying problems on their own) but also applying intact algorithms to new problems and choosing resources used by lower-performing students (seeking help even before they attempt a problem) and therefore represent an intermediate or transitional stage.[39] A similar transitional point between deep and surface approaches to learning was acknowledged as important by Sinapuelas and Stacy but not identified as its own approach.

Implications for Teaching

With the connection between study approaches (deep and surface) and associated levels of metacognition presented by Sinapuelas and Stacy[6] and others,[40,41] this work provides implications for teaching that support the deep approach to studying over the surface approach by concentrating on the development and practice of higher orders of metacognition in the classroom. On the basis of the difference in study approaches found in this research and choice of study resources as outward evidence of metacognitive level,[39] the results of this research can challenge the way instructors view C students. Advice to C students sometimes is to study harder or longer, but as a result of this research, we propose that studying longer is not the answer if the study approaches used are surface approaches. Results from this study could be used to help C students adopt more of the deep approaches to studying used by A/B students. The data reported here could help instructors view students who are not successful as lacking the necessary study approaches.[42] Nonsuccess could then be viewed in some cases as a mismatch between how we teach and how students approach learning. Resolution of the mismatch will require changes in teaching and learning with the instructor and student working together to achieve a deeper level of understanding.

■ CONCLUSIONS

This study has demonstrated the effectiveness of using a 12-item Likert-scale survey (M-ASSIST) of self-reported deep and surface approaches to differentiate between students of different course achievement levels based on course grades. The study also demonstrates the advantages of using a Structured Means Modeling statistical approach to detect hidden significant differences between student groups of different achievement levels that were not evident from ANOVA. In addition, this work reveals that, for the population examined, high-achieving A/B students reported a significantly higher choice of deep approaches and a lower choice of surface approaches than low-achieving D/F students. The C students demonstrated an intermediary or transitional position between high- and low-achieving students, reporting use of deep approaches similarly to unsuccessful students (D/F) and surface approaches intermediate between high- and low-achieving students. These results suggest that the performance of C students in a general chemistry course may be largely related to their greater dependence on surface rather than deep study approaches. This study provides information on methodology (M-ASSIST), analysis (structured means modeling), and results (reported differential use of deep and surface approaches to studying used by different achievement groups in general chemistry), which should help support further research into student learning.

Limitations and Future Work

The limitations of this study include the fact that deep and surface study approaches were designated on the basis of students' self-report to an instrument (M-ASSIST) that has only been tested with the USNA general chemistry population. Additionally, course grades were used to group students because course grades were viewed as a more consistent assessment of student achievement over the entire semester rather than relying on a single exam score. However, course grades may not provide a complete measure of the depth of student knowledge and understanding. Additional work should be done to test the M-ASSIST with other student populations and to look for relationships between M-ASSIST responses and other measures of student knowledge and understanding.

Though students who passed the first semester course went on to take the second semester course, the analysis reported here treated the student response data from each semester as independent data sets in order to focus on reporting achievement group differences and to look for consistency in instrument functioning across semesters. Although there was some movement between grade categories by individuals between the two semesters, a majority of students earned the same grade in both semesters. Future work includes examining individual student responses across semesters to see how self-reported study approaches change over time and how this change may or may not be related to students' course grade. An in-depth examination of study approaches used by C students over time may reveal whether these students are transitioning between surface and deep approaches.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available on the ACS Publications website at DOI: 10.1021/acs.jchemed.7b00202.

> Typical schedule for USNA freshmen (plebes), visualization of 2nd semester student responses to individual M-ASSIST items, descriptive statistics for M-ASSIST items, R code for confirmatory factor analysis, item loadings from single group confirmatory factor analyses for both semesters, and R code for structured means modeling (PDF, DOCX)

DOI: 10.1021/acs.jchemed.7b00202
J. Chem. Educ. 2017, 94, 1415−1424

AUTHOR INFORMATION

Corresponding Author

*E-mail: bunce@cua.edu.

ORCID

Diane M. Bunce: 0000-0001-9069-3573
Regis Komperda: 0000-0003-4837-7141
Maria J. Schroeder: 0000-0003-4411-6294
Debra K. Dillner: 0000-0003-1446-890X
Shirley Lin: 0000-0001-8085-5006
Melonie A. Teichert: 0000-0002-9892-6631
JudithAnn R. Hartman: 0000-0003-0186-9405

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors wish to thank the students and faculty involved in the general chemistry course at the U.S. Naval Academy and the USNA James Kinnear Research Fund for support in completing this study as well as Dr. G. Hancock, Director of the Measurement, Statistics and Evaluation Program, University of Maryland, College Park, for his consultation on the structured means modeling analysis.

REFERENCES

(1) Ausubel, D. P. *Educational Psychology: A Cognitive View*; Holt, Reinhart and Winston: New York, 1968.
(2) Smith, E. L. A Conceptual Change Model of Learning Science. In *The Psychology of Learning Science*; Glynn, S. M., Yeany, R. H., Britton, B. K., Eds.; Lawrence Erlbaum Associates: Hillsdale, NJ, 1991; pp 43–63.
(3) Svensson, L. *Study Skill and Learning*; Acta Universitatis Gothoburgensis: Gothenburg, 1976.
(4) Svensson, L. On Qualitative Differences in Learning: III. Study Skill and Learning. *Br. J. Educ. Psychol.* **1977**, *47*, 233–243.
(5) Chan, J. Y. K.; Bauer, C. F. Learning and Studying Strategies Used by General Chemistry Students with Different Affective Characteristics. *Chem. Educ. Res. Pract.* **2016**, *17* (4), 675–684.
(6) Sinapuelas, M. L. S.; Stacy, A. M. The Relationship Between Student Success in Introductory University Chemistry and Approaches to Learning Outside of the Classroom. *J. Res. Sci. Teach.* **2015**, *52* (6), 790–815.
(7) Marton, F.; Säljö, R. On Qualitative Differences in Learning: I. Outcome and Process. *Br. J. Educ. Psychol.* **1976**, *46*, 4–11.
(8) Chin, C.; Brown, D. E. Learning in Science: A Comparison of Deep and Surface Approaches. *J. Res. Sci. Teach.* **2000**, *37* (2), 109–138.
(9) Entwistle, N. J.; Peterson, E. R. Conceptions of Learning and Knowledge in Higher Education: Relationships with Study Behaviour and Influences of Learning Environments. *Int. J. Educ. Res.* **2004**, *41* (6), 407–428.
(10) Biggs, J. What Do Inventories of Students' Learning Processes Really Measure? A Theoretical Review and Clarification. *Br. J. Educ. Psychol.* **1993**, *63* (1), 3–19.
(11) Tait, H.; Entwistle, N. J.; McCune, V. ASSIST: A Reconceptualisation of the Approaches to Studying Inventory. In *Improving Students as Learners*, Rust, C., Ed.; Oxford Brookes University, The Oxford Centre for Staff and Learning Development: Oxford, 1998.
(12) Biggs, J. *Study Process Questionnaire Manual*; Australian Council for Educational Research: Melbourne, 1987.
(13) Pintrich, P. R.; Smith, D. A. F.; Garcia, T.; McKeachie, W. J. *A Manual for the Use of the Motivated Strategies for Learning Questionnaire (MSLQ)*; National Center for Research to Improve Postsecondary Teaching and Learning, Ann Arbor, MI, 1991.
(14) Entwistle, N. J.; McCune, V. The Conceptual Bases of Study Strategy Inventories. *Educ. Psychol. Rev.* **2004**, *16* (4), 325–345.
(15) Tomanek, D.; Montplaisir, L. Students' Studying and Approaches to Learning in Introductory Biology. *CBE Life Sci. Educ* **2004**, *3* (4), 253–262.
(16) Marton, F.; Säljö, R. Approaches to Learning. In *The Experience of Learning*, Marton, F., Hounsell, D., Entwistle, N. J., Eds.; Scottish Academic Press: Edinburgh, 1984; pp 36–55.
(17) Christian, K.; Talanquer, V. Content-Related Interactions in Self-Initiated Study Groups. *Int. J. Sci. Educ.* **2012**, *34* (14), 2231–2255.
(18) Ye, L.; Shuniak, C.; Oueini, R.; Robert, J.; Lewis, S. Can They Succeed? Exploring at-Risk Students' Study Habits in College General Chemistry. *Chem. Educ. Res. Pract.* **2016**, *17* (4), 878–892.
(19) Brown, A. Metacognition, Executive Control, Self-Regulation, and Other More Mysterious Mechanisms. In *Metacognition, Motivation, and Understanding*, Weinert, F. E., Kluwe, R. H., Eds.; Lawrence Erlbaum: Hillsdale, NJ, 1987; pp 65–116.
(20) Lewis, S. E.; Lewis, J. E. Predicting at-Risk Students in General Chemistry: Comparing Formal Thought to a General Achievement Measure. *Chem. Educ. Res. Pract.* **2007**, *8* (1), 32–51.
(21) Bryer, J.; Speerschneider, K. *likert: Analysis and Visualization Likert Items, R package version 1.3.5*; https://CRAN.R-project.org/package=likert, 2016 (accessed Jun 2017).
(22) R Core Team. *R (version 3.3.2)*; R Foundation for Statistical Computing [Computer Software]: Vienna, Austria, 2014.
(23) Rosseel, Y. lavaan: An R Package for Structural Equation Modeling. *J. Stat. Softw.* **2012**, *48* (2), 1–36.
(24) Finney, S. J.; DiStefano, C. Non-Normal and Categorical Data in Structural Equation Modeling. In *Structural Equation Modeling: A Second Course*, Hancock, G. R., Mueller, R. O., Eds.; Quantitative Methods in Education and the Behavioral Sciences: Issues, Research, and Teaching; Information Age Publishing: Charlotte, NC, 2013; pp 439–492.
(25) Hu, L.; Bentler, P. M. Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives. *Struct. Equ. Model. A Multidiscip. J.* **1999**, *6* (1), 1–55.
(26) Schermelleh-Engel, K.; Moosbrugger, H.; Müller, H. Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. *Methods Psychol. Res. Online* **2003**, *8* (2), 23–74.
(27) Arjoon, J. A.; Xu, X.; Lewis, J. E. Understanding the State of the Art for Measurement in Chemistry Education Research: Examining the Psychometric Evidence. *J. Chem. Educ.* **2013**, *90* (5), 536–545.
(28) American Educational Research Association; American Psychological Association; National Council on Measurement in Education. *Standards for Educational & Psychological Testing*; American Educational Research Association: Washington, DC, 2014.
(29) Nielsen, S. E.; Yezierski, E. Exploring the Structure and Function of the Chemistry Self-Concept Inventory with High School Chemistry Students. *J. Chem. Educ.* **2015**, *92* (11), 1782–1789.
(30) Shedlosky-Shoemaker, R.; Fautch, J. M. Who Leaves, Who Stays? Psychological Predictors of Undergraduate Chemistry Students' Persistence. *J. Chem. Educ.* **2015**, *92* (3), 408–414.
(31) Salta, K.; Koulougliotis, D. Assessing Motivation to Learn Chemistry: Adaptation and Validation of Science Motivation Questionnaire II with Greek Secondary School Students. *Chem. Educ. Res. Pract.* **2015**, *16* (2), 237–250.
(32) Hancock, G. R. Structural Equation Modeling Methods of Hypothesis Testing of Latent Variable Means. *Meas. Eval. Couns. Dev.* **1997**, *30* (2), 91–105.
(33) Steinmetz, H. Analyzing Observed Composite Differences across Groups: Is Partial Measurement Invariance Enough. *Methodology* **2013**, *9* (1), 1–12.
(34) Field, A. *Discovering Statistics Using IBM SPSS Statistics*, 4th ed.; SAGE: Los Angeles, 2013.
(35) Kline, R. B. *Principles and Practice of Structural Equation Modeling*, 3rd ed.; The Guilford Press: New York, 2011.

DOI: 10.1021/acs.jchemed.7b00202
J. Chem. Educ. 2017, 94, 1415–1424

(36) Sörbom, D. A General Method for Studying Differences in Factor Means and Factor Structure between Groups. *Br. J. Math. Stat. Psychol.* **1974**, *27* (2), 229−239.
(37) Chen, F. F. Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Struct. Equ. Model. A Multidiscip. J.* **2007**, *14* (3), 464−504.
(38) Thompson, M. S.; Green, S. B. Evaluating Between-Group Differences in Latent Variable Means. In *Structural Equation Modeling: A Second Course*; Hancock, G. R., Mueller, R. O., Eds.; Quantitative Methods in Education and the Behavioral Sciences: Issues, Research, and Teaching; Information Age Publishing, 2013; pp 163−218.
(39) Bunce, D. M.; Komperda, R.; Dillner, D. K.; Lin, S.; Schroeder, M. J.; Hartman, J. R. Choice of Study Resources in General Chemistry by Students Who Have Little Time To Study. *J. Chem. Educ.* **2017**, *94* (1), 11−18.
(40) Cook, E.; Kennedy, E.; McGuire, S. Y. Effect of Teaching Metacognitive Learning Strategies on Performance in General Chemistry Courses. *J. Chem. Educ.* **2013**, *90* (8), 961−967.
(41) Rickey, D.; Stacy, A. M. The Role of Metacognition in Learning Chemistry. *J. Chem. Educ.* **2000**, *77* (7), 915−920.
(42) Harper, G.; Kember, D. Interpretation of Factor Analyses from the Approaches to Studying Inventory. *Br. J. Educ. Psychol.* **1989**, *59*, 66−74.

DOI: 10.1021/acs.jchemed.7b00202
J. Chem. Educ. 2017, 94, 1415−1424