JOURNAL OF CHEMICAL EDUCATION

Article

pubs.acs.org/jchemeduc

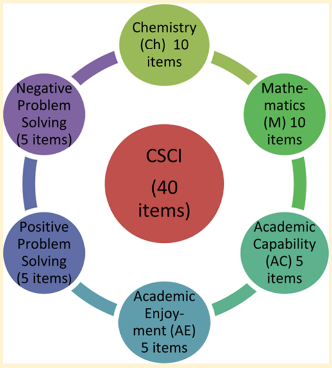
Exploring the Structure and Function of the Chemistry Self-Concept Inventory with High School Chemistry Students

Sara E. Nielsen and Ellen Yezierski*

Department of Chemistry and Biochemistry, Miami University, Oxford, Ohio 45056, United States

Supporting Information

ABSTRACT: Though the Chemistry Self-Concept Inventory (CSCI) was developed to study one aspect of the affective domain in college chemistry students, the instrument on which it was based, the Self-Description Questionnaire III, was developed for use with late adolescents. As such, we explored data generated from administering the CSCI to high school students in seven teachers' classrooms in four schools. This paper presents findings which describe the validity and the reliability of the data obtained from the CSCI when used with high school students. The validity of the data is supported by the results of an exploratory factor analysis, relationships between subscales and chemistry knowledge, and comparisons among different groups of students. Evidence supporting the reliability of the data is provided by internal consistency values for each of the subscales. The results presented in this paper demonstrate the usefulness of the CSCI with a high school population. This instrument may be used by teachers to better understand students' beliefs about themselves with respect to chemistry learning. Implications for research include the utility of the CSCI with high school chemistry students.



KEYWORDS: *High School/Introductory Chemistry, Chemical Education Research, Graduate Education/Research,, Learning Theories*

FEATURE: Chemical Education Research

The fields of educational and social psychology outline the importance of student attitudes and beliefs in learning.[1−6] The field of chemistry education research has employed some of these theories, and the affective domain is considered to be one of the domains of chemistry learning.[7] Chemistry-focused instruments that measure some part of the affective domain include the Colorado Learning Attitudes about Science Survey for chemistry (CLASS-Chem),[8] CHEMX,[9] the Attitude toward the Subject of Chemistry Inventory (ASCI),[10] the Chemistry Attitudes and Experiences Questionnaire (CAEQ),[11] the Chemistry Laboratory Anxiety Instrument (CLAI),[12] and the Chemistry Self-Concept Inventory (CSCI).[13] While this sampling is not intended to be an exhaustive list of chemistry-specific measures of the affective domain, the aforementioned instruments have been shown to measure several important affective constructs, including self-efficacy and self-concept. The attention to instrument development in the affective domain underscores the importance of assessing student beliefs and/or attitudes and understanding the differences among the various affective constructs.

Three widely studied affective constructs are self-esteem, self-efficacy, and self-concept. Self-esteem is a global construct that describes a person's attitudes toward him/herself.[14] However, self-esteem is not often studied in the chemistry context, as it is considered more overarching than topic-specific. On the other end of the specificity spectrum is self-efficacy, which is considered to be task-specific.[4,14] Self-efficacy refers to a person's belief that s/he can accomplish a specific task, such as balancing a chemical equation, solving a stoichiometry problem, or following a laboratory procedure. It is considered to be future-oriented and malleable. Somewhere between the task-specificity of self-efficacy and the global nature of self-esteem is self-concept. Self-concept is a person's domain-specific belief about his or her abilities.[2,5] Students form their judgments about their abilities in a subject, such as chemistry, based on their past performances in that area. The frame of reference of the student also plays a role in his/her self-concept. Self-concept is considered to be more resistant to change than self-efficacy, though some studies have created interventions to target student self-concept.[15]

Perhaps one of the most important features of self-concept is its relation to the cognitive domain. In chemistry it has been shown that chemistry self-concept is correlated with student achievement,[13,16] and may even help predict chemistry achievement.[17] In the broader field of educational psychology, the predictive power of self-concept on achievement has also been illustrated.[18−20] Though self-concept has often been studied with students of all ages in the educational psychology literature, it has less often been studied with high school students in the chemistry education literature. As most students who take general chemistry at the postsecondary level have been exposed to chemistry in high school, we considered it to be important to better understand the chemistry self-concept of high school students. To that end, we designed a longitudinal repeated-measures study to examine the relationships among high school chemistry classroom environment and culture,

DOI: 10.1021/acs.jchemed.5b00302

Table 1. Data Collection Timeline

Instrument	Round					
	0 (Aug−Sep)	1 (Oct−Nov)	2 (Dec)	3 (Jan−Feb)	4 (Mar−Apr)	5 (May)
CSCI	X	X	X	X	X	X
MOSART	X					X

student chemistry self-concept and student chemistry achievement. Presented here is a critical precursor to the larger study that evaluates how the CSCI is functioning in our sample. Bauer's 2005 work[13] includes the CSCI. As the predecessor to the CSCI, the Self-Description Questionnaire III (SDQ III)[21] was designed for use with late adolescents including secondary students, using the CSCI with our sample was appropriate. However, since the CSCI was developed for use with postsecondary students, it was necessary to demonstrate the validity and the reliability of the data produced by the CSCI with our population of interest.

RESEARCH QUESTIONS

The research questions guiding this study are

1. What is the internal structure of the CSCI when used with high school chemistry students?
2. What evidence exists to support the claim that the CSCI produces reliable data when used with high school chemistry students?
3. What evidence exists to support the claim that the CSCI produces valid data when used with high school chemistry students?

METHODS

Before the study was conducted, Institutional Review Board approval was obtained and all of the high school teachers and principals consented to participate in the study. All of the data presented are from students who have assented to participate and whose parents/guardians have also given consent.

The study was designed to have a duration of one academic year. Two cycles of data collection occurred over two consecutive academic years at four high schools with seven full time high school teachers, four student teachers, and 515 high school students. One of the limitations of conducting a longitudinal study in the high school setting is varying attendance and participation. As such, some of the statistical tests performed do not include all 515 students.

An important feature to note in our sample is the presence of academic tracking. All of the participating high schools had varied academic tracks for their chemistry classes. Factors that determined the academic track in which a student was placed included grades in previous science classes, selection by teacher, and self-selection by the student. Included in the study are four different types of chemistry classes: (1) Advanced Placement Chemistry (AP),[22] (2) Honors Chemistry (Hon), (3) College Prep Chemistry (CP), and (4) Chemistry in the Community (Com).[23] The AP students were taking a full-year chemistry course for the second time whereas the other three tracks consisted of students taking a full-year chemistry course for the first time. The Chemistry in the Community class consisted of a specialized curriculum designed for students less interested in a science career. The Honors and the College Prep classes generally worked from the same, more traditional curriculum, though the Honors students worked at a faster pace. The data from all four academic tracks were combined for the factor analysis to increase the sample size and to better understand the internal structure of the CSCI with a wide range of high school chemistry students.

Two of the instruments used in this study are the CSCI and the Misconceptions-Oriented Standards-Based Assessment Resources for Teachers (MOSART)[24] test. The MOSART Chemistry test was designed to elicit the chemistry misconceptions of high school students and was used as a pre-post measure of chemistry achievement for all of the students in our sample. Table 1 shows the timeline of data collection across an academic year. The rounds of data collection were spaced roughly five to six weeks apart.

The data analyzed in this paper were collected at the end of the first semester for all of the students. The decision to use this round of data (herein referred to as Round 2) stemmed from theoretical and practical considerations. One of the assumptions made in earlier self-concept studies was that student self-concept would not be well developed until partway through the academic year.[16] This suggests that the best time to assess the internal structure of the instrument would be with data from late in the academic year. However, as the year progressed, the number of students dropped considerably, and it has been suggested that small sample sizes can jeopardize the validity of factor analysis results.[25] One observation regarding sample size is that larger ratios of subjects to variables produce more accurate factor solutions.[25,26] A ratio of 10 subjects for every variable has been suggested. Using Round 2 data allowed us to preserve a large enough number of students to perform an exploratory factor analysis while still allowing for time for student chemistry self-concept to develop. For all of the analyses conducted, $N = 431$, which surpasses the 10:1 ratio as the CSCI contains 40 items. Survey responses were manually entered, and Microsoft Excel was used to reverse the scores for the negatively worded items. Statistical tests were performed with Microsoft Excel, R, and SPSS 19.

Reliability and Validity

When used in the context of measurement, reliability is often described as a measure of the consistency and stability of responses.[27,28] As self-concept has been shown to be a stable construct in the absence of interventions designed to improve it,[15,29] we only expected to see change in students' self-concept scores if the environment was aligned with practices shown to improve self-concept. Though the data are not presented here, we have evidence that suggests the classroom environments were not structured in a way that would lead us to believe that there would be change in students' self-concept scores. We will present evidence that demonstrates the internal consistency of student responses, as well as the stability of student responses throughout the academic year. The essential idea of validity in measurement is the extent to which the interpretations and conclusions are supported by evidence.[27,30] The ways in which we have established the validity of our data and interpretations is through the relationship of internal structure and variable relations with theory. The internal consistency of the data was measured by Cronbach's α, and the stability of scores over time

DOI: 10.1021/acs.jchemed.5b00302
J. Chem. Educ. XXXX, XXX, XXX−XXX

Table 2. Chemistry Self-Concept Inventory Factor Loadings

		Factor					
Item No.	Item Statements, by Subscales	1	2	3	4	5	6
	Mathematics Self-Concept						
1	I find many math problems interesting and challenging.	0.46	−0.01	0.37	−0.12	0.12	−0.22
5	I have hesitated to take courses that involve math.	0.72	0.10	0.11	−0.09	−0.14	0.04
9	I have generally done better in math courses than in other courses.	0.95	−0.03	−0.12	−0.05	−0.01	−0.17
13	Math makes me feel inadequate.	0.65	0.04	0.00	0.13	−0.09	0.13
17	I am quite good at math.	0.88	−0.09	−0.14	0.14	0.15	−0.12
21	I have trouble understanding anything based on math.	0.72	0.05	−0.02	0.13	−0.07	0.10
25	I have always done well in math classes.	0.88	−0.12	−0.15	0.17	0.08	−0.21
29	I never do well on tests that require math reasoning.	0.61	0.06	−0.04	0.14	−0.04	0.27
33	At school, my friends always come to me for help in math.	0.68	−0.02	−0.03	0.03	0.14	−0.14
37	I have never been very excited about math.	0.75	0.00	0.33	−0.29	−0.09	0.10
	Chemistry Self-Concept						
4	I have never been excited about chemistry.	−0.06	0.70	0.16	0.02	−0.07	0.00
8	I participate confidently in discussions with school friends about chemical topics.	−0.03	0.57	−0.01	−0.05	0.31	−0.07
12	I find chemistry concepts interesting and challenging.	−0.06	0.53	0.31	−0.02	0.17	−0.23
16	When I run into chemical topics in my courses, I always do well on that part.	−0.05	0.79	−0.14	0.09	0.26	−0.20
20	I would hesitate to enroll in courses that involve chemistry.	0.05	0.77	0.14	−0.03	−0.20	0.00
24	I am quite good at dealing with chemical ideas.	−0.02	0.82	−0.06	0.04	0.18	−0.12
28	Chemistry intimidates me.	0.02	0.72	−0.12	0.13	−0.11	0.16
32	I have always had difficulty understanding arguments that require chemical knowledge.	0.06	0.69	0.02	−0.07	−0.15	0.20
36	I have always done better in courses that involve chemistry than in most courses.	0.01	0.77	−0.13	−0.17	0.09	−0.17
40	I have trouble understanding anything based on chemistry.	−0.05	0.74	−0.05	0.22	−0.11	0.09
	Academic Enjoyment Self-Concept						
2	I enjoy doing work for most academic subjects.	0.04	−0.01	0.72	−0.13	0.18	−0.11
6	I hate studying most academic subjects.	0.00	−0.03	0.73	0.00	−0.06	0.10
10	I like most academic subjects.	−0.06	−0.07	0.76	0.20	0.10	−0.16
22	I'm not particularly interested in most academic subjects.	−0.01	−0.01	0.76	0.08	−0.07	0.08
30	I hate most academic subjects.	−0.16	−0.01	0.77	0.27	−0.10	0.10
	Academic Capability Self-Concept						
14	I have trouble with most academic subjects.	0.17	−0.03	0.05	0.67	−0.06	0.06
18	I'm good at most academic subjects.	0.10	−0.03	0.09	0.71	0.09	−0.11
26	I learn quickly in most academic subjects.	0.25	0.08	0.07	0.35	0.24	0.01
34	I get good marks in most academic subjects.	−0.06	−0.07	0.08	0.79	0.19	−0.23
38	I could never achieve academic honors, even if I worked harder.	0.09	0.05	0.06	0.52	0.00	0.11
	Positive Problem Solving Self-Concept						
7	I am good at combining ideas in ways that others have not tried.	0.14	0.03	−0.10	−0.08	0.77	0.50
15	I enjoy working out new ways of solving problems.	0.19	0.09	0.27	−0.11	0.53	0.17
23	I have a lot of intellectual curiosity.	−0.09	0.00	0.18	0.23	0.56	0.22
31	I am an imaginative person.	−0.21	−0.05	−0.05	0.10	0.63	0.49
39	I can often see better ways of doing routine tasks.	0.06	0.02	−0.04	0.11	0.51	0.22
	Negative Problem Solving Self-Concept						
3	I am never able to think up answers to problems that have not already been figured out.	0.08	0.04	0.01	0.15	0.20	**0.55**
11	I wish I had more imagination and originality.	−0.10	−0.06	−0.06	−0.16	0.34	**0.78**
19	I'm not much good at problem solving.	0.33	0.12	0.03	0.07	0.21	**0.38**
27	I am not very original in my ideas, thoughts, and actions.	−0.11	−0.04	0.08	−0.03	0.41	**0.84**
35[a]	I would have no interest in being an inventor.	0.00	0.41	−0.01	−0.21	0.19	**0.30**

[a]Though this item has a higher loading on factor 2 than on factor 6, based on its wording we believe it belongs on factor 6.

was measured by a correlation between scores at several time points. The internal structure of the instrument was assessed through factor analysis. We examined variable relations through correlations among subscales and chemistry achievement, as well as comparisons among student groups.

Exploratory Factor Analysis

The following sections outline our procedures which were carried out according to the Costello and Osbourne's[25] guidelines for best practices for factor analysis with psychological data.

Though an exploratory factor analysis (EFA) had already been conducted on the CSCI,[13] our sample is younger and more heterogeneous in ability than the sample used for the original EFA. Also, the students who participated in the initial testing of the CSCI were all enrolled in college chemistry classes. This differs from our study as it is likely that some of the students in our study will not study chemistry in college. Kline[26] notes that the factors obtained are affected by the sample used for the analysis. Since the two samples are quite different, it was likely that the factor structures might also differ. Furthermore, different exploratory factor analysis techniques

DOI: 10.1021/acs.jchemed.5b00302
J. Chem. Educ. XXXX, XXX, XXX−XXX

were used on the CSCI with college students, the CSCI with our sample, and the SDQ III, which could also affect the factor structure obtained. A more detailed description of previous factor analysis decisions can be found in the Supporting Information. All things considered, there was not a strong enough theoretical basis to test different models with a confirmatory factor analysis.

When performing an EFA, it is important to consider the most appropriate extraction and rotation techniques. Some commonly used extraction techniques include principal components, maximum likelihood (ML), and principal axis factoring (PAF). Principal components analysis is not, strictly speaking, a factor analysis technique. With factor analysis there is the assumption that there are underlying latent factors with which the items are aligned. Principal components analysis has no such assumption and is a data reduction technique. Maximum likelihood is often used as an extraction but one of the key assumptions is that the data meet the requirement of multivariate normality. We assessed our data for multivariate normality using an SPSS macro[31] and found that the data violated that assumption. Principal axis factoring is an extraction method that does not require data to meet the assumption of multivariate normality. As such, we decided that PAF was the most appropriate technique for our data. Other important considerations when performing a factor analysis are the values for the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) and Bartlett's test of sphericity. Higher values of KMO are preferred,[32] and our value was 0.923. Bartlett's test of sphericity tests the hypothesis that our data resemble an identity matrix. The p-value was less than 0.05, indicating that we can reject the null hypothesis and the item data are indeed correlated.

The two main categories of rotation techniques are orthogonal rotations, which require the factors to be uncorrelated, and oblique rotations, which allow (but do not require) the factors to be correlated. When used with college students, nearly all of the factors on the CSCI were correlated with one another. This led us to choose an oblique rotation technique. Two commonly used oblique rotations are direct oblimin and promax. They are both widely used and both produce similar results. When exploring different data analysis decisions we performed separate EFAs with direct oblimin and promax, and as both analyses suggested the same groupings of items, we felt that either approach was appropriate. In the end, we chose to use a promax rotation with the standard κ value of 4.

When determining how many factors should be retained, there are several different techniques that may be used. Retaining all factors with an eigenvalue greater than 1.0 is the default method in SPSS, but this is considered to be an inaccurate method for EFA.[25] Scree plots may also be used but are not considered to be as accurate as parallel analysis. Parallel analysis is a technique that compares actual data to simulated data.[26] All factors from the actual data that have lower eigenvalues than those of the factors in the simulated data are considered noise and are not retained. As such, we conducted a parallel analysis to determine the number of factors to be retained. A more thorough description of parallel analysis can be found in the Supporting Information.

RESULTS AND DISCUSSION

Exploratory Factor Analysis

The results of the parallel analysis suggested that six factors were most appropriate for our sample. Therefore, for our analysis, we fixed the number of factors in SPSS at six. Table 2 shows the six factor solution. The statements are organized by factor and item number. The commonly accepted value for a salient factor loading is 0.3.[26] As such, the minimum requirement for an item to be considered as belonging on a factor was a loading of 0.3 or greater. In some cases, items showed loadings greater than 0.3 on multiple factors; in those cases, we generally assigned the item to the factor on which it had the highest loading. The only anomaly is that Item 35 has a higher factor loading on the Chemistry factor than on the Negative Problem Solving factor. On the basis of the wording of the item alone, it does not fit well with the Chemistry items. Likewise, when considering that on the SDQIII this item factored with other problem solving items,[21] the history of this item also suggests that it does not belong on the Chemistry factor. Furthermore, the factor loadings for Item 35 are similar in magnitude on both factors. Taking all of this into account, we placed Item 35 on the Negative Problem Solving factor.

The factor structure of the CSCI when used with our sample is different from that of the CSCI when used with college students and of the SDQ III. Comparisons between our factor structure and those of the CSCI with college students and the SDQ III can be found in the Supporting Information. While the split between academic capability and enjoyment was seen with the CSCI on college students,[13] it was not seen with the SDQ III.[21] The split in our sample is also slightly different from the one seen with college student data. Also, the items from the Problem Solving factors in our high school data loaded on multiple factors, while the CSCI college student data and SDQ III data were on one factor. Though the way our data factored is not the same as previous instruments, it does have some support from the literature. Other studies have found that students can differentiate between academic capability (or competency) and enjoyment (or affect)[33] and that negatively and positively worded items often factor separately.[34]

Though the first four factors have a mixture of positively and negatively worded items, one likely reason that only the Problem Solving subscales factored separately is due to the lack of cohesion of the items on those subscales. For the first four subscales, all of the items are worded similarly and focus on a singular, unifying construct. The Problem Solving subscales, however, are composed of items that are often quite different from one another, with items ranging from imagination to problem solving. High school students may not associate these items with one another as strongly as they do with the items contained in the Chemistry or Mathematics subscales. This may affect the reliability of the data produced by these subscales.

Reliability and Validity

The internal consistency, one measure of reliability, was measured by Cronbach's α. Table 3 shows the values for each subscale. Values larger than 0.7 are generally considered acceptable.[35] While developers of content tests, especially concept inventories, have often found α values below 0.7, the intent of the instruments is not necessarily to have a test that shows high internal consistency. In fact, the argument has been made that for concept inventories, high α values may indicate redundant questions.[36] For instruments measuring affective constructs, of which the CSCI is one, demonstrating high

DOI: 10.1021/acs.jchemed.5b00302
J. Chem. Educ. XXXX, XXX, XXX−XXX

Table 3. Cronbach's α Values

Subscale	Cronbach's α
Chemistry	0.91
Mathematics	0.93
Academic Capability	0.83
Academic Enjoyment	0.87
Positive Problem Solving	0.72
Negative Problem Solving	0.63

internal consistency within a scale is a commonly accepted practice.

The values in Table 3 indicate that most of the subscales produce internally consistent data. As suggested above, the Problem Solving subscales do not provide as reliable data as the other subscales. Though the Positive Problem Solving subscale does meet the 0.7 guideline, the Negative Problem Solving subscale does not. The number of items on a subscale and the interrelatedness of the items may affect the Cronbach α value.[37] Though the Problem Solving factors each only have five items, it does not seem that the number of items is limiting the α value. The Academic Capability and Academic Enjoyment factors also have five items each. In this case, it seems that the lack of similarity in item wording is leading to subscales that do not produce internally consistent data. Since the two Problem Solving subscales are not as strongly related to school, nor do they provide as reliable data as the other four subscales, they are not included in further analyses.

To analyze the stability of the students' self-concept scores throughout the academic year, we compared students' scores from Round 1 (mid-to-late October) with their scores from Round 2 (mid-December). These time points were chosen for several reasons. Consistent with earlier assumptions that students do not have a reliable self-concept until they have experience with a class, we chose rounds of data at which it could reasonably be argued that students have enough experience to form a reliable self-concept. Another factor in our decision was the limit of sample size. As previously stated, the sample size steadily decreased throughout the academic year. To best demonstrate the stability of scores over time, we felt a larger sample size would be appropriate. Furthermore, while there have been studies that assessed stability over time intervals as long as five months,[29] we felt that limiting the time period of the test-retest interval would lessen the effect of instruction on students' score stability. This is consistent with recommendations by Haertel about measuring stability.[38] Figure 1 shows the distribution of chemistry self-concept scores at Rounds 1 and 2.

Figure 1 shows a fairly strong relationship between Rounds 1 and 2, which demonstrates that the self-concept scores of students are stable between administrations. We chose to display the data with a scatterplot as recent work in the Yezierski group[39] has called into question the use of correlation coefficients alone to demonstrate test-retest reliability. Our data show not only a relationship between administrations, but also low systematic error.

We assessed the variable relations by examining the correlations among the subscale scores. Table 4 gives the values for the Spearman's ρ correlations. In the paper on the development of the CSCI, it was shown that CSCI subscale scores are correlated with one another as well as with chemistry content measures. To support the claim that the data in our study are valid, we would expect similar correlation results

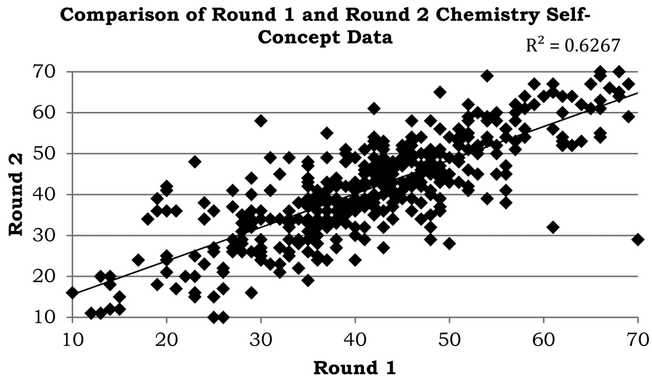


Figure 1. Distribution of chemistry self-concept scores at rounds 1 and 2.

among the subscales and between the subscales and the MOSART test. The values in Table 4 demonstrate that one criterion for the claim that the data are valid has been met.

We would expect that students with different levels of chemistry knowledge and different experiences would have different subscale scores. The possible range for the Chemistry and Mathematics subscales is 10−70 and the possible range for the Academic Capability and Academic Enjoyment subscales is 5−35. The data are presented as a sum of all of the items on the subscale. Figure 2 shows the distribution of Chemistry self-concept scores; the distributions for the other subscales can be found in the Supporting Information. Table 5 shows the subscale means and standard deviations for each of the subscales. The numeric labeling system for the different student groups in Table 5 is not intended to have an ordinal connotation. It is merely provided as a simple way to indicate the groups that are statistically significantly different from one another. The superscript letters indicate the groups that are different based on an analysis of variance test (ANOVA). Assumptions underlying the ANOVA were met (a Levene's test with $p > 0.05$, and adequate sample size for violations of normality). To control the familywise error rate for multiple comparisons, a Bonferroni correction was used. Effect sizes for each statistically significant difference in group mean shown in Table 5 may be found in the Supporting Information.

When the CSCI was used with postsecondary students, it was shown that different groups that were expected to have different chemistry self-concept scores (General Chemistry students, peer leaders, and Chemistry majors) indeed differed.[13] Since the four groups in our study have different backgrounds and preparation in science, we expected the scores for the four groups of students in our study to differ. However, an interesting and unexpected trend emerged from the data. For the Chemistry and Academic Enjoyment subscales, there are only two distinct groups, and for the Mathematics and Academic Capability subscales, there are only three distinct groups. The scores for all of the subscales indicate that the students in the Honors classes and the students in the AP classes are not significantly different from one another. One possible explanation for this trend is the reflected-glory effect, which has been demonstrated in secondary settings, including schools with academic tracking.[40,41] The reflected-glory effect posits that students in groups that are perceived highly due to the accomplishments or qualities of the members of the group will have higher academic self-concepts than would otherwise be predicted. This effect may be contributing to the similarity of the self-concept scores between the AP and Honors students, as

DOI: 10.1021/acs.jchemed.5b00302
J. Chem. Educ. XXXX, XXX, XXX−XXX

Table 4. Spearman's ρ Correlation Values

Subscales	Mathematics	Academic Capability	Academic Enjoyment	Positive Problem Solving	Negative Problem Solving	MOSART post-test
Chemistry	0.43[a]	0.47[a]	0.44[a]	0.42[a]	0.38[a]	0.35[a]
Mathematics		0.63[a]	0.49[a]	0.38[a]	0.26[a]	0.28[a]
Academic Capability			0.58[a]	0.45[a]	0.33[a]	0.33[a]
Academic Enjoyment				0.40[a]	0.22[a]	0.20[a]

[a] $p < 0.001$.

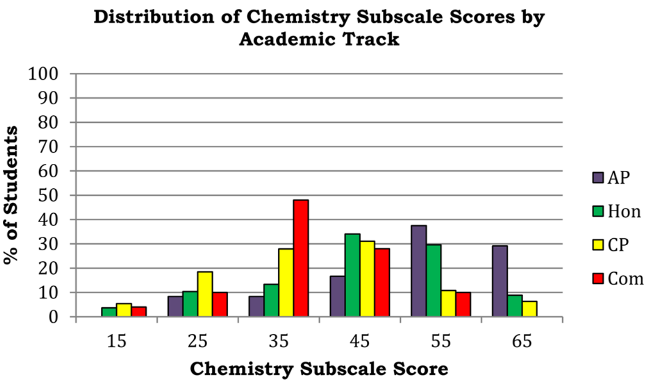


Figure 2. Distribution of chemistry self-concept scores for student groups.

the self-concept scores of the Honors students are enhanced by the glory of being in an Honors class.

The other unexpected trend, the similarity of the CP and Com scores on two of the subscales, cannot be explained by the reflected-glory effect. It may, however, be attributed to the Big-Fish-Little-Pond effect (BFLPE) as described by Marsh and Parker.[42] The BFLPE demonstrates that students' frame of reference and the comparisons they make affect what they believe about their own abilities. Students in groups perceived as having lower status have higher self-concept scores than would otherwise be predicted, whereas students in groups perceived as having higher status have lower scores than would otherwise be predicted. Recent work on the BFLPE in math and science has demonstrated this effect and has underscored the importance of the social comparisons students make and their frame of reference.[43] Scores on the MOSART pre and post tests indicate that the students in the Com group have lower chemistry content knowledge than the students in the CP group. As such, students in the Com group may perceive their group to have a lower status than the CP group, leading an outside observer to predict lower chemistry self-concept scores for the Com students. The BFLPE attributes the difference between the predictions and the results to the frame of reference of the students. The students in the Com group are likely aware of the perceived status of their group, so when they compare themselves to the other members of the group, they may consider themselves to be as good as or better than their peers, the proverbial big fish in the little pond. This may lead them to rank themselves higher than outside observers would expect. Students in groups that are perceived to have a higher status may compare themselves with other members of the group and consider themselves to be average or lower than the other members of the group, the proverbial little fish in a big pond. The BLFPE may be an explanation for the similarity of the self-concept scores between the CP and Com groups. Though the data do not follow the trend exactly as predicted, when taking into account the possible frame of reference effects, the data may still be considered valid.

LIMITATIONS

There are some factors that limit the generalizability of the results presented here. The complexities of conducting research in high schools contribute to many of the limitations of this study. As the majority of student participants are minors, both student assent and parental consent are required. The additional step of acquiring parental consent may limit the number of participating students. Another complication is the use of class time to administer the surveys. Student absences are common, and as every minute of instructional time is valuable, teachers often cannot administer the survey multiple times per round of data collection. This can lead to reduced sample sizes, especially as the academic year progresses. The classes with specialized curricula (AP and Com) often have fewer sections and, thus, fewer student participants than more general classes (Hon and CP). Furthermore, the sampling is not completely randomized, as the study was limited by geography and voluntary participation of teachers. Student self-selection may lead to higher (or, less likely, lower) average group self-concept scores. However, our study captured a variety of school environments and chemistry classes, which provides some evidence for the representativeness of our sample.

IMPLICATIONS

Implications for Research and Future Work

This study shows that students do form a reliable chemistry self-concept in high school, and it can be measured with the CSCI. Future studies should investigate the structure and profile of student chemistry self-concept in settings other than chemistry classes in public high schools in Southwest Ohio.

Table 5. CSCI Subscale Comparisons among Student Groups[a]

	Student Groups (N)			
Subscales	Level 1, AP (24)	Level 2, Hon (135)	Level 3, CP (222)	Level 4, Com (50)
Chemistry	51.5 ± 12.7[d,e]	44.8 ± 12.2[d,e]	38.8 ± 12.2[b,c]	38.0 ± 9.2[b,c]
Mathematics	57.8 ± 8.6[d,e]	53.6 ± 11.6[d,e]	47.0 ± 13.6[b,c,e]	40.5 ± 12.9[b,c,d]
Academic Capability	31.5 ± 4.2[d,e]	29.6 ± 4.1[d,e]	26.6 ± 5.4[b,c,e]	23.3 ± 6.3[b,c,d]
Academic Enjoyment	28.1 ± 5.7[d,e]	25.5 ± 6.5[d,e]	22.5 ± 6.5[b,c]	21.2 ± 5.4[b,c]

[a] Superscripts indicate the levels that are statistically significantly different from the mean listed. [b] AP. [c] Hon. [d] CP. [e] Com.

DOI: 10.1021/acs.jchemed.5b00302
J. Chem. Educ. XXXX, XXX, XXX−XXX

Possible settings may include private high schools, classes other than those dedicated to chemistry, or students of different age groups. Another aim of future studies may be to construct interventions meant to improve high school student chemistry self-concept. Furthermore, previous work has shown that subject-specific self-concept has an effect on selection of future coursework.[44] Further research into chemistry self-concept at both the secondary and postsecondary levels may lead to a better understanding of student behavior, the influence of the affective domain on chemistry learning, and ways to predict student persistence in chemistry coursework.

Implications for Practice

This work demonstrates that the CSCI is an instrument that can produce valid and reliable data with our sample of high school students. These findings suggest that the CSCI may be a valuable tool for high school teachers to use to assess the beliefs of their students. As the affective domain has been shown to be related to the cognitive domain,[5,13,17] it is important for teachers to have data on students' affective characteristics as well as their cognitive performance. College professors may also wish to use the CSCI as a measure of the beliefs of incoming students. This may help them better get to know and understand their students at the beginning of the year. Work on self-concept interventions in other disciplines[15] may be useful to inform ways that teachers and professors can improve their students' chemistry self-concept. Furthermore, self-concept has been shown to be related to important facets of mental health such as anxiety, depression, and attention problems.[45] As such, low self-concept scores can provide teachers with information about students who may need additional support.

Though we collected CSCI scores at 6 time points throughout the academic year, the teachers never indicated that the time required to administer the assessment was an issue. However, if a shorter instrument is desired, it may be possible to assess chemistry and mathematics self-concepts with fewer items, or to eliminate the Problem Solving scales from the instrument. It is advisable to reestablish the reliability and the validity of the results with a shortened instrument.

Future Work

As this paper has demonstrated that the CSCI can offer reliable and valid data about high school students' self-concept, we plan on examining our data for longitudinal trends, as well as looking to find methods of grouping the students by their CSCI profiles. Qualitative analysis on the classroom environment, as well as student and teacher interviews, will also be undertaken.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available on the ACS Publications website at DOI: 10.1021/acs.jchemed.5b00302.

> Comparisons of the factor structures among the SDQ III, the CSCI when used with college students, and the CSCI when used with high school students; effect sizes, a more thorough description of Parallel Analysis, and the distributions of students' self-concept scores on the Mathematics, Academic Capability, and Academic Enjoyment subscales (PDF, DOCX)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: yeziere@miamioh.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the high school principals, teachers, and students for their willingness to participate in this study. We are grateful to the members of the Yezierski and Bretz research groups for their advice and support.

■ REFERENCES

(1) Rosenberg, M. *Society and the Adolescent Self-Image*; Princeton University Press: Princeton, 1965.
(2) Burns, R. B. *The Self-Concept in Theory, Measurement, Development, and Behaviour*; Longman: New York, 1979.
(3) Bandura, A. *Social Learning Theory*; Prentice-Hall: Englewood Cliffs, NJ, 1977.
(4) Schunk, D. H. Self-Efficacy and Academic Motivation. *Educ. Psychol.* **1991**, *26*, 207−231.
(5) Marsh, H. W.; Martin, A. J. Academic Self-Concept and Academic Achievement: Relations and Causal Ordering. *Br. J. Educ. Psychol.* **2011**, *81*, 59−77.
(6) Novak, J. D. Human Constructivism: A Unification of Psychological and Epistemological Phenomena in Meaning Making. *International Journal of Personal Construct Psychology* **1993**, *6* (2), 167−193.
(7) Bretz, S. L. Novak's Theory of Education: Human Constructivism and Meaningful Learning. *J. Chem. Educ.* **2001**, *78*, 1107−1116.
(8) Adams, W. K.; Wieman, C. E.; Perkins, K. K.; Barbera, J. Modifying and Validating the Colorado Learning Attitudes about Science Survey for Use in Chemistry. *J. Chem. Educ.* **2008**, *85* (10), 1435−1439.
(9) Grove, N.; Bretz, S. L. CHEMX: An Instrument to Assess Students' Cognitive Expectations for Learning Chemistry. *J. Chem. Educ.* **2007**, *84* (9), 1524−1529.
(10) Bauer, C. F. Attitude towards Chemistry: A Semantic Differential Instrument for Assessing Curriculum Impacts. *J. Chem. Educ.* **2008**, *85* (10), 1440−1445.
(11) Dalgety, J.; Coll, R. K.; Jones, A. Development of Chemistry Attitudes and Experiences Questionnaire (CAEQ). *J. Res. Sci. Teach.* **2003**, *40* (7), 649−668.
(12) Bowen, C. W. Development and Score Validation of a Chemistry Laboratory Anxiety Instrument (CLAI) for College Chemistry Students. *Educ. Psychol. Meas.* **1999**, *59* (1), 171−185.
(13) Bauer, C. F. Beyond "Student Attitudes": Chemistry Self-Concept Inventory for Assessment of the Affective Component of Student Learning. *J. Chem. Educ.* **2005**, *82* (12), 1864−1870.
(14) Bong, M.; Skaalvik, E. M. Academic Self-Concept and Self-Efficacy: How Different Are They Really? *Educ. Psychol. Rev.* **2003**, *15* (1), 1−40.
(15) O'Mara, A. J.; Marsh, H. W.; Craven, R. G.; Debus, R. L. Do Self-Concept Interventions Make a Difference? A Synergistic Blend of Construct Validation and Meta-Analysis. *Educ. Psychol.* **2006**, *41* (3), 181−206.
(16) Nieswandt, M. Student Affect and Conceptual Understanding in Learning Chemistry. *J. Res. Sci. Teach.* **2007**, *44* (7), 908−937.
(17) Lewis, S. E.; Shaw, J. L.; Heitz, J. O.; Webster, G. H. Attitude Counts: Self-Concept and Success in General Chemistry. *J. Chem. Educ.* **2009**, *86* (6), 744−749.
(18) Marsh, H. W. Causal Ordering of Academic Self-Concept and Academic Achievement: A Multiwave, Longitudinal Panel Analysis. *J. Educ. Psychol.* **1990**, *82* (4), 646−656.

(19) Marsh, H. W.; Yeung, A. S. Causal Effects of Academic Self-Concept on Academic Achievement: Structural Equation Models of Longitudinal Data. *J. Educ. Psychol.* **1997**, *89* (7), 41−54.

(20) Marsh, H. W.; Trautwein, U.; Ludtke, O.; Koller, O.; Baumer, J. Academic Self-Concept, Interest, Grades, and Standardized Test Scores: Reciprocal Effects Models of Causal Ordering. *Child Dev.* **2005**, *76* (2), 397−416.

(21) Marsh, H. W.; O'Neill, R. Self Description Questionnaire III: The Construct Validity of Multidimensional Self-Concept Ratings by Late Adolescents. *J. Educ. Meas.* **1984**, *21* (2), 153−174.

(22) AP Chemistry Course Overview. College Board. https://apstudent.collegeboard.org/apcourse/ap-chemistry (accessed Dec 2014).

(23) Chemistry in the Community. American Chemical Society. http://www.acs.org/content/acs/en/education/resources/highschool/chemcom.html (accessed Dec 2014).

(24) Misconceptions-Oriented Standards-Based Assessment Resources for Teachers. Harvard College. https://www.cfa.harvard.edu/smgphp/mosart/aboutmosart_2.html (accessed Dec 2014).

(25) Costello, A. B.; Osborne, J. W. Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. *Pract Assess., Res. Eval.* **2005**, *10* (7), 1−9.

(26) Kline, P. *An Easy Guide to Factor Analysis*; Routledge: New York, 1994.

(27) Messick, S. Consequences of Test Interpretation and Use: The Fusion of Validity and Values in Psychological Assessment. In *Problems and Solutions in Human Assessment: Honoring Douglas N. Jackson at Seventy*; Goffin, R. D., Helmes, E., Eds.; Kluwer Academic Publishers: Boston, MA, 2000; pp 3−20.

(28) 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*, 536−545.

(29) Marsh, H. W.; Smith, I. D.; Barnes, J.; Butler, S. Self-Concept: Reliability, Stability, Dimensionality, Validity, and the Measurement of Change. *J. Educ. Psych.* **1983**, *75* (5), 772−790.

(30) Kane, M. T. Validation. In *Educational Measurement*, 4th ed.; Brennan, R. L., Ed.; Praeger: Westport, 2006; pp 17−64.

(31) DeCarlo, L. T. On the Meaning and Use of Kurtosis. *Psychol. Methods* **1997**, *2*, 292−307.

(32) Kaiser, H. F. An Index of Factorial Simplicity. *Psychometrika* **1974**, *39* (1), 31−36.

(33) Marsh, H. W.; Craven, R.; Debus, R. Separation of Competency and Affect Components of Multiple Dimensions of Academic Self-Concept: A Developmental Perspective. *Merrill-Palmer Q.* **1999**, *45* (4), 567−601.

(34) Marsh, H. W.; Scalas, L. F.; Nagengast, B. Longitudinal Tests of Competing Factor Structures for the Rosenberg Self-Esteem Scale: Traits, Ephemeral Artifacts, and Stable Response Styles. *Psychol. Assessment* **2010**, *22* (2), 366−381.

(35) Gable, R. K. *Instrument Development in the Affective Domain*; Kluwer-Nijhoff Publishing: Boston, MA, 1986.

(36) Adams, W. K.; Wieman, C. E. Development and Validation of Instruments to Measure Learning of Expert-Like Thinking. *Int. J. Sci. Educ.* **2011**, *33* (9), 1289−1312.

(37) Tavakol, M.; Dennick, R. Making Sense of Cronbach's Alpha. *Int. J. Med. Educ.* **2011**, *2*, 53−55.

(38) Haertel, E. H. Reliability. In *Educational Measurement*, 4th ed.; Brennan, R. L., Ed.; Praeger: Westport, 2006; pp 65−110.

(39) Harshman, J.; Yezierski, E. J. Test-Retest Reliability of the Adaptive Chemistry Assessment Survey for Teachers: Measurement Error and Alternatives to Correlation. *J. Chem. Educ.*, submitted.

(40) Marsh, H. W.; Kong, C. K.; Hau, K. T. Longitudinal Multilevel Models of the Big-Fish-Little-Pond Effect on Academic Self-Concept: Counterbalancing Contrast and Reflected-Glory Effects in Hong Kong Schools. *J. Pers. Soc. Psychol.* **2000**, *78* (2), 337−349.

(41) Trautwein, U.; Ludtke, O.; Marsh, H. W.; Nagy, G. Within-School Social Comparisons: How Students Perceive the Standing of Their Class Predicts Academic Self-Concept. *J. Educ. Psych.* **2009**, *101* (4), 853−866.

(42) Marsh, H. W.; Parker, J. W. Determinants of Student Self-Concept: Is It Better to Be a Relatively Large Fish in a Small Pond Even if You Don't Learn to Swim as Well? *J. Pers. Soc. Psychol.* **1984**, *47* (1), 213−231.

(43) Chiu, M. S. The Internal/External Frame of Reference Model, Big-Fish-Little-Pond Effect, and Combined Model for Mathematics and Science. *J. Educ. Psych.* **2012**, *104* (1), 87−107.

(44) Marsh, H. W.; Yeung, A. S. Coursework Selection: Relations to Academic Self-Concept and Achievement. *Am. Educ. Res. J.* **1997**, *34* (4), 691−720.

(45) Marsh, H. W.; Parada, R. H.; Ayotte, V. A Multidimensional Perspective of Relations Between Self-Concept (Self Description Questionnaire II) and Adolescent Mental Health (Youth Self-Report). *Psychol. Assessment* **2004**, *16* (1), 27−41.

DOI: 10.1021/acs.jchemed.5b00302
J. Chem. Educ. XXXX, XXX, XXX−XXX