
Chemistry Education Research and Practice
www.rsc.org/cerp

Latent constructs of the Students' Assessment of Their Learning Gains instrument following instruction in stereochemistry

Abstract

Pedagogical practitioners who emphasise active learning in undergraduate chemistry courses widely use the Student Assessment of Learning Gains (SALG) instrument to measure students' perceptions of their gains in knowledge and skills in chemistry. Although numerous studies have reported SALG results in support of successful pedagogical interventions, a comprehensive construct-verified version measuring students' perceptions of their chemistry learning is lacking. This paper aims to identify latent constructs of the SALG instrument that was administered in Process Oriented Guided Inquiry Learning (POGIL) classes by using exploratory and confirmatory factor analyses. When the SALG was administered on two separate occasions with two different groups of students following four weeks of instruction on topics in stereochemistry, the results revealed a four-factor structure consisting of 32 items that included Active Learning, Concept Learning, Resources, and Process Skills. These findings demonstrate an approach to collect evidence to support the match between intended constructs and measured variables in light of a targeted pedagogical intervention.

Introduction

There has been an increasing interest in research-based learner-centred teaching approaches aimed at improving students' chemistry learning outcomes. Abundant research literature is available acknowledging the outcomes of the implementation of such practices via student evaluations of faculty and courses (Danielle, & Janice, 2012; Fairweather, 2008; Mataka & Kowalske, 2015; Smith, Douglas, & Cox, 2009; Weaver, Russell, & Wink, 2008). These student evaluations are considered helpful to the academic community for a wide variety of purposes such as identifying components of effective teaching and areas of instruction in need of improvement, and also recognising excellence in teaching (Wachtel, 2006). Researchers studying the effectiveness of teaching innovations focus increasingly on the measurement and evaluation of student perceptions of their learning (Arjoon, Xu, & Lewis, 2013) in conjunction with their cognitive achievement in the disciplinary area (Anaya, 1999; Bowman, 2013). The affective dimensions of pedagogical inventions are generally assessed either through questionnaires or interviews. In questionnaire/interview situations, students are presented with items asking about how each component of the instruction helped their learning and they judge each item using a pre-determined numerical scale or respond to it verbally (Schunk, 1992). However, students vary in their ability to accurately identify the extent to which various learning experiences from a pedagogical intervention positively influenced their learning (Bowman, 2011). Moreover, the nature and quality of learning experiences or learning gains expressed by the students are sensitive to the items of the instruments and their factorial structure (American Psychological Association, 1999). Therefore, it is of prime importance to continue to conduct research on instruments meant for students' self-evaluations mainly to provide some evidence of reliability, validity, and utility. In the case of adapted survey instruments where the researchers or practitioners make changes to its constituent items, the need for the demonstration of evidence of reliability and validity (Malhotra & Grover, 1998) is helpful to maximise the usefulness of the instrument in any study.

The objective of this research was to provide evidence of validity and reliability of the Students' Assessment of Learning Gains (SALG) – an instrument used for students to assess their perceptions of learning in chemistry classes – in the context of teaching and learning stereochemistry. Despite its wide usage, and availability of several versions, only the version of SENCER-SALG (Weston, Seymour, & Thiry, 2006) had been construct validated to reveal its factorial structure. In their report as part of science education for new civic engagement and responsibilities (SENCER) project, Weston et al. hinted on the validation of SALG data although no information was provided on the implied statistical procedures. Furthermore, convergent validity has not been previously determined. Of note is that Moody and Sindre (2003) revealed that SALG was withdrawn from their research study because it lacked evidence for its theoretical construct. The need for the establishment of a statistically-evident factorial structure was of a particular concern in this research because an instrument with evidence of validity provides a more coherent set of measures. With a factorial structure obtained through a systematic statistical technique, the SALG instrument can provide assurance to the administrators that the psychometric inferences from the data of students' self-measures of learning gains will be reliable and valid for the context within which it has been used. Further, the administrators or instructors can use these factor scores as *units of analysis* in various statistical tests in order to make judgements on the impact of their teaching.

The primary research questions that guided this research were:

1. Within the context of teaching and learning stereochemistry, using Exploratory Factor Analysis (EFA), what are the latent constructs of the SALG instrument?
2. Within the context of teaching and learning stereochemistry, does Confirmatory Factor Analysis (CFA) of the SALG instrument identify latent constructs consistent with the exploratory model?

The above research questions were carefully planned as they helped the researchers explore data obtained from a pedagogical intervention known as process oriented guided inquiry learning – POGIL (Moog, Creegan, Hanson, Spencer, Straumanis & Bunce, 2008) in order to provide answers in support of SALG's factor composition and their further verification. An overarching question that is often posed to any administrator of SALG is, "does SALG measure what it is supposed to measure". Therefore, this question sets an expectation to report factor analyses results. In short, "Factor analysis is intimately involved with questions of validity….. Factor analysis is at the heart of the measurement of psychological constructs" (Nunnally, 1978, pp. 112-113). Consequently, four latent constructs namely *active learning, concept learning, resources, and process skills* were found to underlying the version of SALG used in this study.

Background

Developed by Seymour, Wiese, Hunter, and Daffinrud (2000), the SALG instrument garners information related to the content, pedagogical approach, learning activities, grading and assessment procedures, resources, and student engagement in terms of workload and pace of learning. The SALG instrument is usually used at the end of the semester to measure the students' self-perception of their learning gains and their progress towards course learning gains. Instructors also may use the instrument halfway through the course to enable them make informed course corrections. Additionally, instructors also use a baseline or introductory survey to sense the position of the students with respect to the desired learning goals.

Further, the SALG data collected during pre-course and post-course evaluations are considered helpful in providing a snapshot (Middlecamp, Jordan, Shachter, Kashmanian, & Lottridge, 2006) of students' skills and attitudes before and after an intervention. The SALG instrument and various other traditional end-of-course questionnaires differ from each other in that the former primarily focuses on students' self-reporting of their learning gains in the specific activities or course elements, whereas the latter provides students with the opportunity to rate the instructors' teaching competencies, practices and resources.

The SALG instrument has been used in a number of studies related to student learning of university level chemistry that primarily focused on active learning and student-centred pedagogies. Seymour (2002) administered paper-pencil and online SALG instruments in a multi-institutional study both at the end and mid-points of the semester to measure the students' perception of their learning. Middlecamp et al. (2006) developed and used online SENCER-SALG as a tool for assessing how SENCER courses were successfully influencing student learning. Hopkins and Samide (2013) used SALG in their inquiry-based laboratory curriculum to teach general chemistry. The students reported their scores on their pre-class knowledge and subsequent gains in knowledge after the inquiry-based laboratory instruction. SALG was also used as a post-course survey of inquiry-based instruction by Prescott (2013) in a general chemistry course for non-majors. Similarly, Walker and Sampson (2013) used SALG survey to determine how students viewed argument-driven inquiry instruction in the chemistry laboratory.

The SALG instrument is used to gauge students' perceptions of skills, understanding, and attitudes towards teaching or laboratory courses (Herreid, 2013; Seymour, 2002; van Rooji, 2009; Yadav, 2011). Carroll (2010) inferred that a combination of SALG and student achievement tests could offer curriculum practitioners a powerful triangulation on measures and causes of student learning. Straumanis and Simons (2008) used SALG as an indicator of growth of students' process skills in non-didactic organic chemistry classes and reported that non-didactic responses were higher than those in the control didactic group. According to Seymour et al. (2000), SALG provides average scores and standard deviations for responses to each statement and requests that students include written explanations for their responses to each main question.

Descriptive statistical analysis and response frequencies of SALG are widely used to interpret students' responses (Douglas, & Chiu, 2009; Heady, 2002; Johnson, Corazzini, & Shaw, 2011; Keeney-Kennicutt, Gunersel, & Simpson, 2008; Keeves, 1995; van Rooji, 2009) to each or a set of the Likert scale questions in an effort to provide a glimpse of students' perception of course implementation. Heady (2002) administered the SALG survey successively to two student cohorts over two years in introductory biology classes to find out what helps students to learn. The study compared the mean values for all of the student responses to the items of SALG. Johnson et al. (2011) had used an on-line SALG survey containing 9 measurement *domains* to explore the virtual learning environment in nursing education. Their SALG instrument contained 4-point Likert type items organised as *domains* but information in support of *domain* composition was not available. In an another study on the effectiveness of project management methodology in a psychology class, van Rooij (2009) administered a 20-item SALG survey and presented comparative mean scores of students' SALG responses in project management methodology and traditional project scaffolding. Keeney-Kennicutt, Gunersel, and Simpson (2008) used a web-based SALG instrument to investigate general chemistry students' perceptions of an educational web-based tool called, calibrated peer review. The results of the trend analysis included the percentage values of students' responses to the 5-item SALG survey. A review of the above studies revealed researchers' attempts to establish validity and reliability of SALG data by comparing: *i)* student's SALG responses with their interview data; *ii)* mean values and other measures of learning like achievement. Comparing mean values from SALG with other measures of learning may not provide greater correspondence because students' perceptions and achievement may not be the same (Poe, 1969).

Theoretical Background on the SALG Survey

The original SALG had no theoretical evidence in support of its design and also subsequent users were less vigilant about the important aspect of theory that guided the design of the instrument. Based on the nature of the items used, it appears that the instrument may have been informed by the sociological theory of Merton (1968). Merton inferred that empirical uniformities can be derived from logically interconnected propositions and attributed manifest and latent functions to social processes. Accordingly, Seymour, Wiese, Hunter, and Daffinrud (1997) identified two characteristic features from effective teachers (1993) that were found to be relevant for the SALG instrument: 1) regular evaluation of teaching practice in the form of assessment and feedback to understand whether such practices are beneficial to students' learning, and 2) familiarity of students' academic preparation, knowledge, and abilities and fine-tuning teaching strategies to enhance students' learning gains. The SALG instrument also reflects the characteristics recommended by Kuh (2001) for student self-assessment surveys. These characteristics include students' awareness of the information they are asked for, precision and clarity of the questions, and question items focusing on meaningful activities that could evoke thoughtful response from the students. According to Kuh, the information, ideas and the language presented in the instruments is relevant to the learning context of the students. According to Seymour (2000), the flexibility of adapting SALG in between different disciplines of science is dependent on the extent of cohesiveness of various course elements such as goals of class or laboratory activities, curriculum, resources used and tested.

Construct Validity

At present, there is no evidence in support of construct validity for any of the adapted versions of SALG used in active learning pedagogical implementations. Construct validity indicates whether or not the instrument actually measures the construct under investigation (Coll, Dalgety, & Slater, 2002) and it is inclined to the nature of items in the modified SALG. Cronbach and Meehl (1959) warranted that construct validation is to be identified by the orientation of the investigator rather than by a specific investigatory approach. The construct validation procedures are theoretically based and include establishing the convergent validity and discriminant validity of the measure (Agarwal, 2013). Factorial analysis is often used to assess construct validity and is often achieved by including information from the items (observed variables) of instrument in as few derived factors as possible to keep the solution understandable (Gorsuch, 1983). Convergent validity is established when the variables that tap the same construct are correlated with each other, whereas discriminant validity is established when variables that tap different constructs are not correlated with each other (Hsiao, Wu, & Yao, 2014).

Design and Procedures

The research reported in this article was part of a major research project approved by the Human Research Ethics Committee (HREC) of the investigators' university that focused students' perceptions of learning chemistry in a student-centred intervention like POGIL. A short description of POGIL philosophy is included in the following section. The research design for the major study consisted of a post-positivist paradigm using a quasi-experimental design (Creswell, 2003) with quantitative data collected using SALG. Post-positivist research is commonly aligned with quantitative methods of data collection and analysis. Post-positivist paradigm emphasises well-defined concepts and variables, controlled conditions, precise instrumentation and empirical testing (Weaver & Olson, 2006). Post-positivism was considered appropriate for this study as it offered the researchers an impersonal position to make context-dependent generalisations (Cooper, 1997) using methods that minimise the susceptibility of participants. Further, the scope of this article restricted the authors to avail only quantitative data.

Pedagogical Context

The POGIL teaching-learning method has shown to be effective in chemistry major courses at several institutions in the United States. More recently, in Australia, Active Learning in University Science (ALIUS), a collaborative project of six Australian universities, uses POGIL as a model of teaching innovation to engage students in large first year chemistry classes (Bedgood Jr. et al., 2009). Consequently, the research study was undertaken at a large tertiary institution in Australia where POGIL has been actively practiced in selected first year chemistry courses.

POGIL is a student-centred instructional approach where students work in small groups with the instructor acting as a facilitator. In a POGIL classroom, students work in learning teams using specially designed activities that promote mastery of the discipline content and the development of skills necessary for scientific inquiry. POGIL practitioners have used SALG and published their results of student engagement, their perceptions of the value of small group learning and the perceived growth in process skills. The major characteristic features of the POGIL model are concept learning, development of process skills, active engagement and use of resources or activities. Henceforth, it is desirable to identify the comprehensive measurement scales for the SALG instrument in order to make it more reflective of the characteristics of the pedagogical approach for which it is used.

Disciplinary Context

Stereochemistry is an important aspect of organic chemistry that primarily includes the study of relative spatial arrangement of atoms within molecules and the study of stereochemical requirements and outcomes of chemical reactions. Topics in Stereochemistry, taught as part of a chemistry course to first year students, included: chirality, stereocenter, stereoisomers, molecular orientation at stereo-carbons, and identifying chiral molecules on the basis of plane of symmetry, non-superimposable mirror image formation, and ability to estimate the possible number of stereoisomers from a stereocenter of the molecule, distinguishing isomers, S_N1 and S_N2 reactions, curved arrow processes, and nucleophilic substitution reactions. A modified POGIL approach in the form of embedded mini-lectures, small group POGIL discussions, followed by clicker questions was utilised.

Sample

The sample comprised a cohort of first year chemistry students enrolled during 2011 and 2012, referred to as Group 1 (n = 114) and Group 2 (n = 154), respectively. Most of the students were in Engineering, Science, and Pharmacy programmes opting to study chemistry during first and second semesters. The majority of the students (domestic and international) were school leavers, however, non-traditional students included mature age learners and students with vocational qualifications comprised a minority of the population. Students' participation in the study was voluntary and only self-selected students were invited to complete the paper-pencil SALG during chemistry workshops/tutorial sessions. The topic being studied, over a period of four weeks, was stereochemistry.

The SALG instrument

The paper-pencil SALG instrument was administered to Groups 1 and 2 in 2011 and in 2012 to obtain data for the exploratory and confirmatory analyses, respectively. The SALG instrument, comprising 62, 5-point Likert scale items, was administered during the end of second semester 2011 for exploratory factor analysis (n = 114). Based on the results, the instrument was refined and the 44 item 5-point Likert scale SALG instrument was administered to Group 2 students during the end of second semester of 2012 for confirmatory factor analysis (n = 154). The outline of the development and validation of the SALG instrument is shown in Figure 1. In addition to the Likert scale items, SALG also included items that were aimed at seeking students' written responses on various aspects of the POGIL class (but not reported in this article).

For establishing **convergent validity** of SALG, the factor loadings and internal consistency reliability measures were computed for the 2011 data. Brown (2006) suggested a strong interrelation of different measures of theoretically similar or overlapping constructs for convergent validity. The explored model was then fitted to the 2012 data and fit statistics examined with reference to established criteria (Hu & Bentler, 1999).

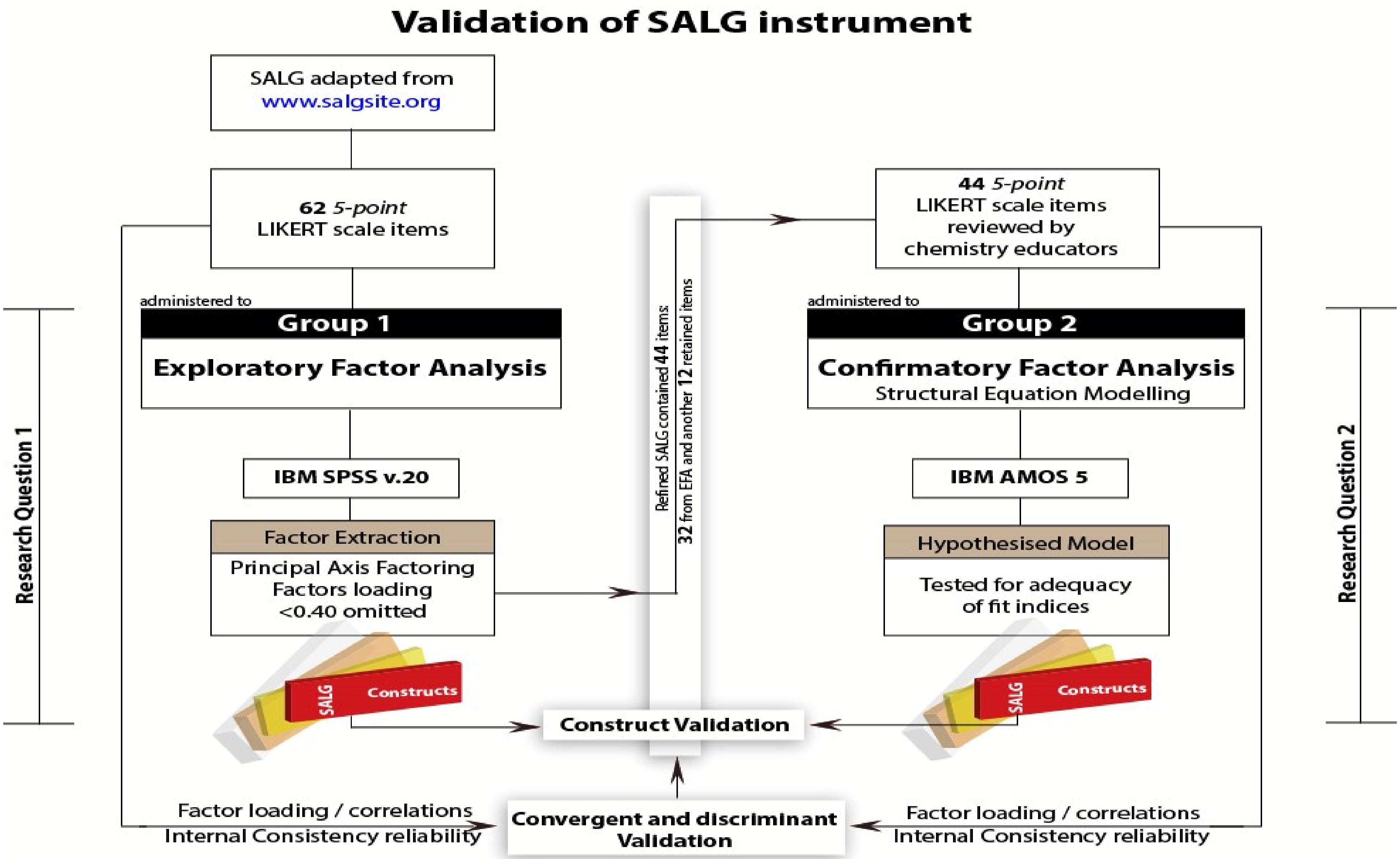


Figure 1: The process for collecting evidence for determining the SALG Instrument validity from administration in this study

Analysis and Findings

The most commonly used procedure for psychometric evaluation of questionnaires is *factor analysis* which can be performed in two ways: *exploratory factor analysis* (EFA) and *confirmatory factor analysis* (CFA). In EFA, researchers follow data reduction procedures (reducing large set of variables to a manageable number) and further explore the data for the appropriate number of common factors that can reasonably serve as indicators of a set of measured variables; in CFA, a pre-specified factor solution is evaluated (Brown, 2006).

Construct Validity: Exploratory factor analyses (EFA)

Exploratory factor analyses (EFA) is generally employed in the process of scale development and construct validation (Brown, 2006) and is a data-driven approach to see the relevant common factors emerging from it (Johnson & Stevens, 2001) and also investigate the relationship between manifest variables and factors (Everitt & Hothorn, 2011).

The purpose of EFA in this study involved the first research question to investigate the latent constructs encompassing the SALG instrument. Subsequently, the EFA, performed on all 62 items of SALG, used a principal axis factoring analysis with varimax rotation procedure using SPSS version 20 to extract four sets of factors from a total of 32 items. Appendix 2 summarizes the results of EFA carried out in this study. Varimax rotation was chosen due to its easiness to interpret the factors as it maximises variance between factors (Foster, Barkus & Yavorsky, 2006). The feasibility of factor analysis was determined by examining the Kaiser–Meyer–Olkin (KMO) measure of sampling and Bartlett's test of sphericity. The KMO measure of sampling adequacy was 0.785, indicating that the data were appropriate for exploratory factor analysis (Tabachnick & Fidel, 1989). The acceptable limit of the KMO measure is .50 (Kaiser, 1974). Bartlett's test of sphericity indicated that χ^2 = 2196.521 which was statistically significant ($p<0.001$) indicating that correlation matrix is significantly different from an identity matrix. Identity matrix is usually obtained in situations where there is no existence of correlation between any of the variables. Items loading on more than one factor with a loading score of equal to or less than 0.40 on each factor were eliminated from the analysis to indicate sufficient loading (Hinkin, 1998). Table 1 shows the results of the varimax rotation and the factors obtained after EFA.

Following the exploratory factor analysis, factor loadings indicate how strongly each item was related to a particular factor. Eigenvalues showed the relative importance of each factor, and the cumulative variance was used to check whether a sufficient number of factors have been retained. The eigenvalue for each factor was greater than 1, as per Kaiser criterion (Kaiser, 1960) and the cumulative variance for all the four factors was 45.79%. This showed that four factors can explain over 45 per cent of the total variability in the 32 items. After consideration of the intent of the items clustered on each factor, the derived four factors were labelled as Active Learning (18 items), Concept Learning (7 items), Resources (4 items) and Process Skills (3 items). Representative items of each factor are listed in Table 2. The Cronbach alpha reliability coefficients, shown in Table 4, are highly satisfactory (Arjoon, Xu, & Lewis, 2013) being greater than 0.7.

Table 1. Factor loading, eigenvalue and percentage of variance for SALG (Group 1, 2011) (*n* = 114)

Item Number	Factor Loadings			
	Active Learning	Concept Learning	Resources	Process Skills
1	.45			
2	.55			
3	.60			
4	.42			
5	.48			
6	.52			
7	.64			
8	.61			
9	.64			
10	.49			
11	.62			
12	.71			
13	.41			
14	.52			
15	.55			
16	.61			
17	.56			
18	.61			
19		.55		
20		.52		
21		.77		
22		.77		
23		.72		
24		.58		
25		.56		
26			.80	
27			.87	
28			.89	
29			.46	
30				.89

31				.68
32				.80
% Variance	18.12	10.69	9.33	7.65
Eigenvalue	9.46	2.86	2.64	1.90
Cumulative % Variance	18.12	28.82	38.15	45.79

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.

Table 2. A sample of representative items in each latent factor

Factor	Item: As a result of your work in this class, what gains did you make in the following:
Active Learning	Participating in group work
	Listening to discussions
Concept Learning	Understanding and classifying chiral-achiral molecules
	Understanding and distinguishing isomers
Resources	Mini-lectures helped my learning
	Clicker questions helped my learning
Process Skills	Skill of argument use of evidence
	Skill of identifying data pattern

Construct Validity: Confirmatory factor analysis (CFA)

Confirmatory factor analysis (CFA) was used to determine the **construct validity**, namely, whether the factor structure resulting from the exploratory factor analysis (EFA) could be consistent with the data obtained from another similar group – in this case the Group 2 cohort during semester 2 in 2012 when the refined SALG was used. The refined SALG contained 44 items, of which 32 were from the four-factored EFA and another 12 that were retained because of their relevancy to the intervention was completed by 154 students. An outline on the development and administration of the SALG instrument is presented in the right hand side of Figure 1.

Subsequently, CFA was conducted using a Structural Equation Modelling (SEM) approach to test the four factor model derived from the EFA (Figure 2). Unlike other statistical procedures meant for establishing construct validity of factors, SEM offers the researcher the ability to use multiple measures to represent constructs and test their relationship with other constructs addressing the issue of specific errors of measurement (Weston & Gore, 2006). All SEM measurement models are tested against a host of fit indices to evaluate their representation of relationship among constructs and observed variables. For this study, the four factor model of EFA was applied to Group 2 data (n=154) in an effort to answer the second research question. From the proposed four factor model, using AMOS v20 software, a χ^2 = 619.406, *df* = 385, and p <0.001 were obtained indicating that the model can be estimated and tested. The proposed four factor measurement model was identified fulfilling the two recommended general requirements (Hu & Bentler, 1999); first, the number of pieces of information in the model shall be at least as large as the number of parameters to be estimated. The four factor model contained 80 distinct parameters to be estimated and 465 distinct pieces of information; thereby meeting the first requirement. The second requirement; every latent variable including the residual terms must be assigned a scale; all the latent variables in the model and the errors terms has a scale assigned to each one of them. Though the χ^2 statistics obtained from the four factor model appeared significant, finding an exact fit to the data is rare (Weston & Gore, 2006). Owing to the limitations of χ^2 statistics, this statistics is not used as the sole index of overall model fit (Brown, 2006). Therefore, other fit indices were explored to determine whether the model fit is acceptable. Further the fit indices χ^2/df = 1.60, CFI = 0.92, RMSEA = 0.06, TLI = 0.91, and SRMR = 0.07 appeared to have met the adequacy criteria of goodness-of-fit. Appendix 3 summarizes the results of CFA carried out in this study. Hu and Bentler (1999) suggested the following cut-off limits for achieving a reasonably good fit between the target model and the observed data: (1) CFI values greater than 0.90; (2) RMSEA values close to 0.06 or lower; (3) TLI values close to 0.95 or lower; (4) SRMR values close to 0.80 or lower.

For improving the model fit, the model modification indices were used. Two of the Active Learning items (15 and 18), '*grading system what I need to work*' and '*Willingness to seek help from others (instructor, peers, tutor) when working on academic problems*' did not fit the CFA model. The items 15 & 18 are yet useful (if retained) to the researchers/instructors in order to capture students' perceptions of their learning in POGIL because these items inquire the usefulness of seeking help from instructors/tutors/peers and their understanding of the grading system used in workshops. Similarly, the items numbered 24, 25 and 29 also were not included in the CFA *measurement model* due to the fact that, the planning of analysis is driven principally by theoretical relationships of observed (items) and unobserved variables (latent constructs) (Schreiber, 2008). Further the use of correlated measurement errors is allowed when such practices are theoretically or methodologically justifiable (Shah & Goldstein, 2006). As an example of a substantive evidence for modifications in model, items *14* and *16* convey nearly the same and serve as mutual controls on the consistency of answer, hence they were correlated. A similar example existed for items *3* and *9*. The items of the four factor measurement model - for active learning, resources and process skills are relevant to all teaching of chemistry by student-centred approach - are presented in Table 3. The concept learning in stereochemistry is specific to the study. When using a different topic, the concept learning outcomes need to be changed accordingly.

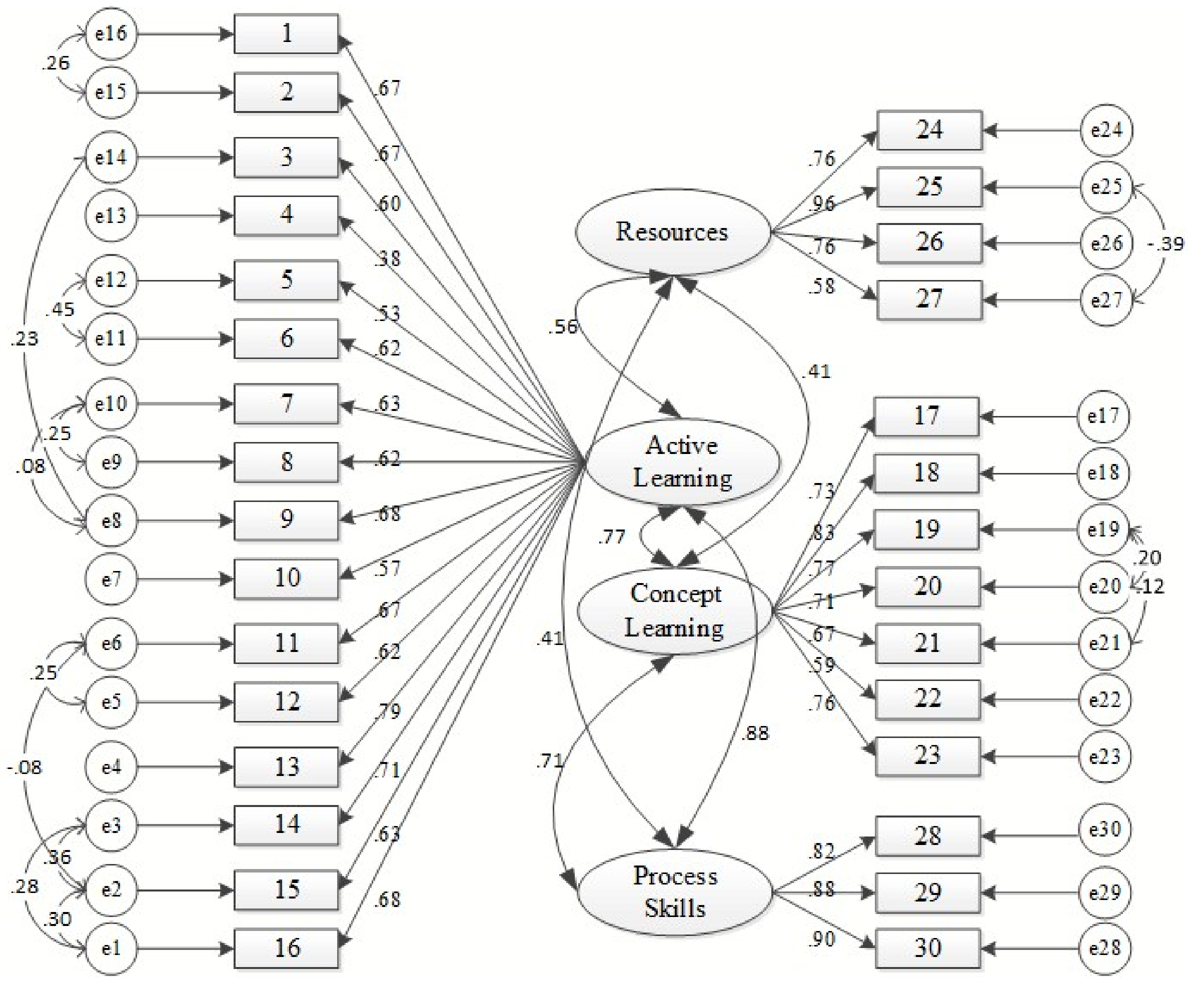


Figure 2. The four factor measurement model of SALG

Table 3. Items of the four factor measurement model

Item Number *CFA*	*Item Number EFA*	Item
Active Learning		
1	*2*	Attending class
2	*1*	Pace of class
3	*3*	Working with peers
4	*4*	Working with peers outside the class
5	*5*	Explanation of instructor for involving small groups
6	*6*	Explanation of focus on topics presented
7	*10*	Participating in class discussions
8	*11*	Listening to discussions
9	*12*	Participating in group work
10	*13*	Class activities help learning
11	*14*	Number and spacing of tests
12	*16*	Feedback on my work received during and after tutorials
13	*17*	Connecting key ideas to other knowledge
14	*7*	Confidence that you understand the material
15	*8*	Confidence in ability to do POGIL activities
16	*9*	Comfort level involving complex ideas
Concept Learning		
17	*20*	S_N1 S_N2 reaction mechanism
18	*21*	Distinguishing different types of isomers
19	*22*	Classifying chiral-achiral molecules
20	*23*	Identifying stereocentres in molecules
21	*19*	Attractive forces between molecules and the effect on physical properties
22	..	Applying curved arrow conventions to describe bond forming and bond breaking processes
23	..	The reactions of alkyl halides, nucleophilic substitution reactions
Resources		
24	*26*	Mini lectures
25	*27*	Posted Pencasts
26	*28*	Pencasts solutions of homework problems

27	..	Clickers
Process Skills		
28	*31*	Identifying patterns in data
29	*30*	Recognising a sound argument and appropriate use of evidence
30	*32*	Develop logical argument

Convergent validity

Convergent validity was assessed by examining the inter-correlations of the constructs of SALG. As shown in Table 7, the correlations among the four constructs were found to be statistically significant.

Internal consistency reliability was established by calculating the Cronbach's alpha coefficient for each factor. The guidelines (Cohen, Mannion, & Morrison, 2000; Nunnally, 1978) indicate that an alpha coefficient of 0.70 is adequate for an instrument in the early stage of development; a coefficient of at least 0.80 is adequate for a more developed instrument. The results portrayed in Table 4 show that the Cronbach's alpha coefficient for each factor was above 0.80, affirming the reliability of the scales of SALG. Based on the analysis of the data, the factor loadings and internal consistency measure confirmed the convergent validity of the SALG questionnaire used in this study.

Table 4. Internal consistency reliability (Cronbach's alpha) for the SALG scales

Factor	Number of items	Cronbach's Alpha
Active Learning	18	0.90
Concept Learning	7	0.84
Resources	4	0.81
Process Skills	3	0.89

The Cronbach alpha internal consistency reliability of the items of SALG after CFA was calculated and the values are presented in Table 5. These values are highly satisfactory with similar values to those presented in Table 4.

Table 5. Internal consistency reliability of SALG scales after CFA using EFA scales

Factor	Number of items	Cronbach's Alpha
Active Learning	16	0.92
Concept Learning	7	0.89
Resources	4	0.82
Process Skills	3	0.90

This four factor CFA – SEM analysis would appear to demonstrate Brown's (2006) criteria for **convergent validity** with strong interrelation of different measures of theoretically similar or overlapping latent constructs (see Table 6).

Table 6: Pearson correlation coefficient values of four factors of the SALG instrument

	Active Learning	Concept Learning	Resources	Process Skills
Active Learning		0.69	0.54	0.77
Concept Learning			0.42	0.66
Resources				0.41
Process Skills				

p< 0.01 level (2-tailed)

Discriminant validity

Discriminant validity according to Brown (2006) is expressed by results showing that indicators of theoretically distinct constructs are not highly inter-correlated. He further argued that, factor correlations above 0.80 imply overlap of items and point towards poor discrimination validity. The discriminant validity of the items of the instrument was assessed by comparing the construct correlations with the square root of the average variance extracted (AVE). Fornell and Larcker (1981) specify that discriminant validity is achieved when the square root of the AVE of a construct is larger than its correlation with other constructs. The square roots of the AVE were calculated and are represented in bold on the main diagonal of Table 7. The off diagonal elements represent the correlations among the latent variables. The results reported in Table 7 confirm that the discrimination validity was achieved by all scales. As shown in the Table 7, the correlations ranged from 0.17 to 0.51, providing further evidence in support of the discriminant validity.

Table 7: Inter construct correlations and square roots of average variance extracted for the SALG scales

	Active Learning	Concept Learning	Resources	Process Skills
Active Learning	**0.78**			
Concept Learning	0.45	**0.82**		
Resources	0.31	0.17	**0.89**	
Process Skills	0.51	0.41	0.35	**0.94**

p< 0.01 level (2-tailed)

Note. Square root of average variance extracted (AVE) is shown on the diagonal of the matrix

Discussion and Conclusion

Based on the responses following instruction in stereochemistry, four factors containing 32 items were extracted from the SALG instrument during EFA. Appendix 1 provides a complete list of factors and their corresponding items. The factor analysis of the data obtained from 114 students from Group 1 resulted in a four factorial structure of the SALG instrument; Active Learning, Concept Learning, Resources, and Process Skills. Since the study had occurred at only one institution, it was always difficult to acquire the desired sample size. However, considering the criteria of *variable to sample* ratio (Bentler & Chou, 1987; Conway & Huffcutt, 2003), the sample for EFA *(n = 114)* had ratio less than 5:1 and the sample for CFA *(n = 154)* had ratio greater than 5:1. The internal consistency reliability was highly satisfactory where each factor had a Cronbach's alpha coefficient value greater than 0.80. For CFA, the explored four factor model was fitted to the data obtained from Group 2, 2012 (*n* = 154) using a measurement model of structural equation modelling (SEM); the fit statistics met the criteria of a good fit. The Cronbach's alpha internal consistency reliability values of the SALG constructs after CFA were highly satisfactory (>0.80). The findings give support to Hong, Purzer, and Cardella's (2011) suggestion that, for adapted instruments, the CFA be used to test the fit of the factor structure from a sample different to the EFA.

The findings from the sophisticated use of EFA and CFA indicate that the SALG questionnaire has high convergent and discriminant validity when used with these first year chemistry classes learning stereochemistry. Therefore, data collected using this survey is likely to be valid and reliable in this study context. Although the results of this study need to be replicated with large samples across a range of chemistry units (by substituting items under subscale *concept learning*), and in different institutions, the four-factor model may provide POGIL practitioners with a useful approach to predict students' acceptance of the intervention when implemented in different cultural contexts. The causal relationships between the four subscales of SALG, when explored further, may provide opportunities for meaningful evaluation of students' perceptions of their learning gains in research-based student-centred pedagogies.

The data utilised in this study identified a fit between latent constructs and observed variables that are relevant to POGIL instruction. For example, Active Learning construct obtained after EFA contained 18 items (see Appendix 1) that are distinctly relevant to various elements of pedagogical intervention followed in this study. Similarly, the 8 items under Concept Learning were appropriate and broadly covered the disciplinary context.

According to Brown (2006), there are two categories of factors – overdetermined, and underdetermined, based on the number of strongly related indicators (items) in them. Conspicuously in this study, Active Learning, an overdetermined factor (factor with several indicators) and Process Skills, an underdetermined factor (factor with two or three indicators) have emerged. Though they have theoretical relevance to the pedagogical intervention that the study had followed, it becomes more practical for the purpose of interpretability if no such skewed breakdown of factors occurred. A large sample replication study is recommended to overcome to further ensure recoverability of the proposed model.

Limitations

The SALG instrument has shown both good reliability and validity for measuring students' perceptions of their learning in active learning classrooms; in this case in teaching and learning the topic of stereochemistry using a modified POGIL approach. However, despite its rigor and depth to the interpretation of results, the research based on self-report data has potential for continuous errors in self-assessments to confound the results (Beghetto, 2007; Dunning, Heath, & Suls, 2004). Despite its flexibility for adoption or adaptation, the various versions of SALG need to be consistently verified for underlying constructs in order to enhance their interpretability of the generated data. The data sets used for EFA and CFA emerged from a relatively smaller sample size which may lead to unstable solutions. Smaller samples increase the likelihood of obtaining underdetermined factors. The original SALG developed by Seymour et al. (2000) had 5 Likert scale points. However, chemistry educators who have successively used SALG contained 6 Likert scale points. Considering the variation in the number of Likert scale points used, the internal consistency reliability of SALG instrument may (Chang, 1994) or may not (Cummins & Gullone, 2000) vary. The instrument could be enriched by reducing the number of items on the first factor *active learning* or alternatively creating a new scale *comfort and confidence* using items *12* to *16* to avoid overemphasis on modification indices while attempting to improve the model fit. The findings from this study may not be generalizable to other contexts due to the fact that SALG is flexible in terms of its constituent items thus making this outcome more context-dependent. Depending on the conceptual area taught, the items will change as required. This study specifically deals with the teaching and learning of stereochemistry. The scope of the article restricted authors to the measurement model only, hence the fit of structural models to the observed data is not discussed. Above all, a considerable theoretical and statistical sophistication is required for the pedagogical practitioners intending to evaluate the impact of their implementations.

Acknowledgements

We thank the students for their participation in this study and the Department of Chemistry for allowing the researchers to collect the data.

Research Ethics

This research has been reviewed and given approval by the institution's Human Research Ethics Committee (Approval number: SMEC-45-10).

Supplementary Information

Appendix 1. The SALG Instrument

Appendix 2. Details of EFA results

Appendix 3. Details of CFA results

References

Agarwal V., (2013). Investigating the convergent validity of organizational trust. *J. Commun. Manag.,* **17**(1), 24-39.

Americal Psychological Association., (1999). *Standards for Educational and Psychological Testing*. Washington, DC.: American Psychological Association.

Angelo T. A. and Cross K. P., (1993). *Classroom assessment techniques: A handbook for college teachers*, San Francisco, CA: Jossey-Bass.

Arjoon J. A., Xu X. and Lewis J. E., (2013). Understanding the state of the art for measurement in chemistry education research: Examining the psychometric evidence, *J. Chem. Educ.*, **90**(5), 536−545.

Bedgood Jr D. R., Yates B., Buntine M., Pyke S., Lim K., Mocerino M., Zadnik M., Southam D., Bridgeman A., Gardiner M. and Morris, G., (2009). ALIUS: Active learning in university science: Leading change in Australian science teaching. *Chem. Australia,* **77**(5), 18-19.

Beghetto R. A., (2007). Factors associated with middle and secondary students' perceived science competence, *J. Res. Sci. Teach.*, **44**(6), 800-814.

Bentler P. M. and Chou C.-P., (1987). Practical issues in structural modeling, *Sociol. Method. Res.*, **16**(1), 78-117.

Bowman N. A., (2011). Can college students accurately assess what affects their learning and development? *J. Coll. Student. Dev.,* **52**(3), 270-290.

Bowman N. A., (2013). Understanding and addressing the challenges of assessing college student growth in student affairs. *Res. Pract. Assess.,* **8**, 1-14.

Brown T. A., (2006). *Confrimatory factor analysis for applied research*, New York: The Guildford Press.

Carroll S. B., (2010). in *Science education and civic engagement: The SENCER approach*, ed. Sheardy R. D., Oxford: Oxford University Press, vol. 1037, ch. 10, pp. 149-198.

Chang L., (1994). A psychometric evaluation of 4-point and 6-point Likert-type scales in relation to reliability and validity. *Appl. Psychol. Meas.*, **18**(3), 205-215. doi: 10.1177/014662169401800302

Cohen L., Mannion L. and Morrison K., (2000). *Research methods in education*. New York: RoutledgeFalmer.

Conway J. M. and Huffcutt A. I., (2003). A review and evaluation of exploratory factor analysis practices in organizational research. *Organ. Res. Method.*, **6**(2), 147-168.

Coll R., Dalgety J. and Salter D., (2002). The development of the chemistry attitudes and experiences questionnaire (CAEQ). *Chem. Educ. Res. Pract.*, **3**(1), 19-32. doi: 10.1039/B1RP90038B

Cooper M. M., (1997). Distinguishing critical and post-positivist research, *Coll. Compos. Commun.*, **48**(4), 556-561.

Creswell J. W., (2003). *Research design, qualitative, quantitative, and mixed methods approaches* Thousand Oaks, CA: Sage Publications, Inc.

Cronbach L. J. and Meehl P. E., (1955). Construct validity in psychological tests. *Psychol. Bull.*, **52**(4), 281-302.

Cummins R.A. and Gullone E., (2000). Why we should not use 5-point Likert scales: The case for subjective quality of life measurement. Proceedings, Second International

Conference on Quality of Life in Cities (pp.74-93). Singapore: National University of Singapore.

Danielle K. T. and Janice B., (2012). STEM practice and assessment: SENCER's influence on educators. In R. D. Sheardy & W. D. Burns (Eds.), *Science Education and Civic Engagement: The Next Level* (Vol. 1121, pp. 163-178). Washington, DC: American Chemical Society.

Douglas E. P. and Chiu C. C., (2009). Use of guided inquiry as an active learning technique in engineering, Palm Cove, Qld, Australia.

Dunning D., Heath C. and Suls J. M., (2004). Flawed self-assessment: Implications for health, education, and the workplace, *Psychol. Sci. Public Interest*, *5*(3), 69-106.

Everitt B. and Hothorn T., (2011). in *An Introduction to Applied Multivariate Analysis with R*, New York: Springer ch. 5, pp. 135-161.

Fairweather J., (2008). *Linking evidence and promising practices in science, technology, engineering, and mathematics (STEM) undergraduate education*, Board of Science Education, National Research Council, The National Academies, Washington, DC.

Fornell C. and Larcker D. F., (1981). Evaluating structural equation models with unobservable variables and measurement error *J. Marketing Res.*, **18**(1), 39-50.

Foster J. J., Barkus E. and Yavorsky C., (2006). *Understanding and using advanced statistics*, Thousand Oaks, CA: SAGE Publications.

Gorsuch R. L., (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Lawrance Erlbaum Associates, Inc.

Heady J. E., (2002). in *Innovative Techniques for Large Group Instruction*, Arlington, VA: National Science Teachers Association Press, pp. 57-62.

Herreid C. F. and Schiller N. A., (2013). Case studies and the flipped classroom, *J. Coll. Sci. Teaching*, **42**(5), 62-66.

Hinkin T., (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organ. Res. Method.,* **1**(1), 104-121.

Hong T., Purzer, S. and Cardella M. E., (2011). A psychometric re-evaluation of the design, engineering and technology (DET) survey, *J. Eng. Educ.*, **100**(4), 800-818.

Hopkins T. A. and Samide M., (2013). Using a thematic laboratory-centered curriculum to teach general chemistry, *J. Chem. Educ.*, **90**(9), 1162-1166.

Hu L. and Bentler M., (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives, *Struct. Equation Model. Multidscip. J.*, **6**(1), 1-55.

Johnson B. and Stevens J. J., (2001). Exploratory and confirmatory factor analysis of the school level environment questionnaire (SLEQ), *Learn. Envir. Res.*, **4**, 325-344.

Johnson C. M., Corazzini, K. N. and Shaw R., (2011). Assessing the feasibility of using virtual environments in distance education, *Knowl. Manag. E-Learn. Int. J.*, **3**(1), 5-16.

Hsiao Y.-Y., Wu, C.-H. and Yao, G., (2014). Convergent and discriminant validity of the WHOQOL-BREF using a multitrait-multimethod approach. *Soc. Indic. Res.*, **116**(3), 971-988. doi: 10.1007/s11205-013-0313-z

Kaiser H. F., (1960). The application of electronic computers to factor analysis, *Educ. Psychol. Meas.*, **20**(1), 141-151.

Kaiser H. F., (1974). An index of factorial simplicity. *Psychometrika*, **39**(1), 31-36.

Keeney-Kennicutt W. L., Gunersel, A. B. and Simpson N. J., (2008). Overcoming student resistance to a teaching innovation, *Int. J. Scholarship. Teach. Learn.*, **2**(1), 1-26.

Keeves J. P., (1995). The contribution of IEA research to Australian Education. In Reflections on Educational Achievement, Bos, W., Lehmann, R. H., Eds.; Waxmann: New York.

Kuh G. D., (2001). The national survey of student engagement: Conceptual framework for overview of psychometric properties, http://nsse.indiana.edu/2004_annual_report/pdf/2004_Conceptual_Framework.pdf 2001 (accessed July 2014).

Malhotra M. K. and Grover, V., (1998). An assessment of survey research in POM: from constructs to theory. *J. Oper. Manag.*, **16**(4), 407-425.

Mataka L. M. and Kowalske M. G., (2015). The influence of PBL on students' self-efficacy beliefs in chemistry. *Chem. Educ. Res. Pract.*, **16**(4), 929-938. doi: 10.1039/C5RP00099H

Merton R. K., (1968). *Social theory and social structure*, New York: Free Press.

Middlecamp C. H., Jordan T., Shachter A. M., Kashmanian O. K. and Lottridge S., (2006). Chemistry, society, and civic engagement (Part 1): The SENCER project, *J. Chem. Educ.*, **83**(9), 1301.

Moody D. L. and Sindre G., (2003). Evaluating the effectiveness of learning interventions: An information systems case study.

Moog R. S., Creegan J. F., Hanson M. D., Spencer N. J., Straumanis A. and Bunce M. D. (2009). POGIL: Process-oriented guided-inquiry learning. In N. Pienta, M. M. Cooper, & T. J. Greenbowe (Eds.), *Chemists' Guide To Effective Teaching* (Vol. 2, pp. 90-101). Upper Saddle River, NJ: Prentice Hall.

Nunnally J. C., (1978). *Psychometric Theory* (2nd ed.). New York: McGraw-hill.

Poe C. A., (1969). Convergent and discriminant validation of measures of personal needs. *J. Educ. Meas.*, **6**(2), 103-107. doi: 10.2307/1434286

Prescott S., (2013). Green goggles: Designing and teaching a general chemistry course to nonmajors using a green chemistry approach, *J. Chem. Educ.*, **90**(4), 423-428.

Schreiber J. B., Nora A., Stage F. K., Barlow E. A. and King, J., (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review, *J. Educ. Res.*, **99**(6), 323-338.

Schunk D. H., (1992). Theory and research on student perceptions in the classroom. In D. H. Schunk & J. L. Meece (Eds.), *Student perceptions in the classroom* (pp. 3-24). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Seymour E., Wiese D., Hunter A. and Daffinrud S., (1997). Student assessment of learning gains (SALG) CAT . http://www.flaguide.org/extra/download/cat/salg/salg.pdf 1997

Seymour E., Wiese D., Hunter A. and Daffinrud S.M., (2000) presented in part at the National Meeting of the American Chemical Society, San Francisco, CA.

Seymour E., (2002). Tracking the processes of change in US undergraduate education in science, mathematics, engineering, and technology, *Sci. Ed.*, **86**(1), 79-105.

Shah R. and Goldstein S. M., (2006). Use of structural equation modeling in operations management research: Looking back and forward. *J. Oper. Manag.*, **24**, 148-169.

Smith K. A. Douglas T. C. and Cox M. F., (2009). Supportive teaching and learning strategies in STEM education, *New Dir. Teach. Learn.*, **117**, 19-32.

Straumanis A. and Simons E. A., (2008). A milti-institutional assessment of the use of POGIL in organic chemistry. In R. S. Moog & J. N. Spencer (Eds.), *ACS Symposium Series 994: Process-Oriented Guided inquiry Learning* (pp. 226-239). Washington, DC: American Chemical Society.

Tabachnick B. and Fidel F., (1989). *Using multivariate statistics*, New York: Harper Collins Publishers.

Van de Vijver F., (2011). Capturing bias in structural equation modelling. In E. Davidov, P. Schmidt & J. Billiet (Eds.), *Cross-cultural analysis methods and applications* (pp. 3-34). New York: Taylor & Francis Group, LLC.

van Rooij S. W., (2009). Scaffolding project-based learning with the project management body of knowledge (PMBOK), *Comput. Educ.*, **52**(1), 210-219.

Wachtel H. K., (1998). Student evaluation of college teaching effectiveness: A brief review., *Assess. Eval. High. Educ.*, **23**(2), 191-212.

Walker J. P. and Sampson V., (2013). Argument-driven inquiry: Using the laboratory to improve undergraduates' science writing skills through meaningful science writing, peer-review, and revision, *J. Chem. Educ.*, **90**(10), 1269-1274.

Weston R. and Gore P. A. (2006). A brief guide to structural equation modelling. The Counselling Psychologist, **34**(5), 719-751.

Weaver K. and Olson J. K., (2006). Understanding paradigms used for nursing research. *J. Adv. Nurs.*, **53**(4), 459-469.

Weaver G. C., Russell, C. B. and Wink, D. J., (2008). Inquiry-based and research-based laboratory pedagogies in undergraduate science, *Nat. Chem. Biol.*, **4**(10), 577-580.

Weston T., Seymour E. and Thiry H. (2006). Evaluation of science education for New Civic Engagements and Responsibilities (SENCER) Project. Washington, DC: The National Center for Science and Civic Engagement.

Yadav A., Subedi D., Lundeberg M. A. and Bunting, C. F., (2011). Problem-based learning: Influence on students' learning in an electrical engineering course *J. Eng. Educ.*, **100**(2), 253-280.

Appendix 1: SALG Instrument

Student Assessment of their Learning Gains

Instructions:
Please fill in one box only for the following statements corresponding to the scale indicated by completely filling the circle ● with a blue or black ballpoint pen.

Active Learning
HOW MUCH did each of the following aspects of the class HELP YOUR LEARNING?

		No help	A little help	Moderate help	Much help	Great help
1	Attending class	1	2	3	4	5
2	The pace of class	1	2	3	4	5
3	Working with peers	1	2	3	4	5
4	Working with peers outside the class	1	2	3	4	5
5	Explanation of instructor for involving small groups	1	2	3	4	5
6	Explanation of why the class focus on topics presented	1	2	3	4	5
7	Participating in class discussions during class	1	2	3	4	5
8	Listening to discussions	1	2	3	4	5
9	Participating in group work during class	1	2	3	4	5
10	Class activities help learning	1	2	3	4	5
11	The number and spacing of tests	1	2	3	4	5
12	The feedback on my work received during and after tutorials	1	2	3	4	5
13	Connecting key ideas to other knowledge	1	2	3	4	5
14	Confidence that you understand the material	1	2	3	4	5
15	*Grading system what I need to work*	1	2	3	4	5
16	Comfort level in working complex ideas	1	2	3	4	5
17	Confidence in ability to do POGIL activities	1	2	3	4	5
18	*Willingness to seek help from others (instructor, peers, tutor) when working on academic problems*	1	2	3	4	5

Concept Learning
As a result of your work in this class, what GAINS DID YOU MAKE in your UNDERSTANDING of each of the following?

19	SN1 SN2 reaction mechanism	1	2	3	4	5
20	Distinguishing different types of isomers	1	2	3	4	5
21	Classifying compounds as chiral/achiral	1	2	3	4	5
22	Identifying StereoCentres in molecules	1	2	3	4	5
23	Attractive forces between molecules and effect on physical properties	1	2	3	4	5
24	Applying curved arrow conventions to describe bond forming and bond breaking processes	1	2	3	4	5
25	The reactions of alkyl halides and nucleophilic substitutions	1	2	3	4	5

Resources
HOW MUCH did each of the following aspects of the class HELP YOUR LEARNING?

26	Mini lectures posted on the Blackboard	1	2	3	4	5
27	Pencasts posted on the Blackboard	1	2	3	4	5
28	Pencasts solutions homework problems	1	2	3	4	5
29	Clickers during instruction	1	2	3	4	5

Process Skills
As a result of your work in this class, what GAINS DID YOU MAKE in the following SKILLS?

30	Identifying patterns in data	1	2	3	4	5
31	Recognizing a sound argument and appropriate use of evidence	1	2	3	4	5
32	Developing a logical argument	1	2	3	4	5

Appendix 2: Exploratory Factor Analyses

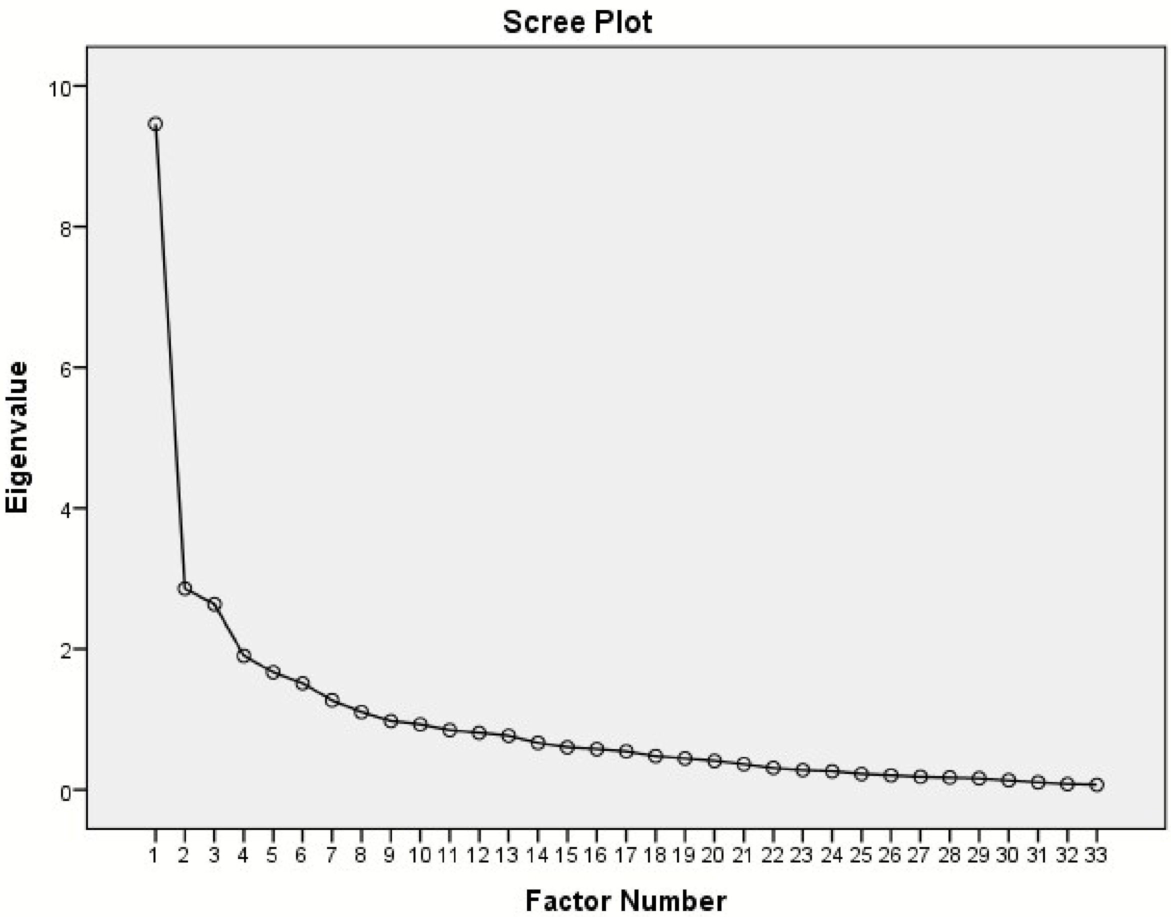
KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.785
Bartlett's Test of Sphericity	Approx. Chi-Square	2196.521
	df	528
	Sig.	.000

Total Variance Explained

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	9.457	28.658	28.658	8.931	27.064	27.064	5.981	18.123	18.123
2	2.856	8.655	37.314	2.427	7.354	34.418	3.529	10.693	28.816
3	2.635	7.985	45.299	2.211	6.701	41.119	3.079	9.330	38.146
4	1.902	5.765	51.064	1.543	4.675	45.794	2.524	7.648	45.794
5	1.670	5.061	56.125						
6	1.512	4.583	60.708						
7	1.269	3.846	64.553						
8	1.103	3.344	67.897						
9	.976	2.957	70.854						
10	.928	2.811	73.665						
11	.846	2.563	76.228						
12	.808	2.449	78.677						
13	.767	2.323	81.000						
14	.666	2.017	83.017						
15	.604	1.830	84.847						
16	.578	1.750	86.598						
17	.545	1.651	88.248						
18	.478	1.447	89.695						
19	.445	1.347	91.043						
20	.411	1.244	92.287						
21	.363	1.099	93.386						
22	.308	.934	94.320						
23	.279	.846	95.166						
24	.262	.795	95.960						
25	.224	.678	96.639						
26	.204	.618	97.257						
27	.184	.557	97.814						
28	.172	.522	98.336						
29	.161	.488	98.825						
30	.134	.405	99.229						
31	.105	.318	99.548						
32	.079	.240	99.788						
33	.070	.212	100.000						

Extraction Method: Principal Axis Factoring.


Scree Plot
Eigenvalue
10
8
6
4
2
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33
Factor Number

Rotated Factor Matrix[a]

Item No		Factor			
		Active Learning	Concept Learning	Resources	Process Skills
1	Pace_of_class_5.3	.447			
2	Attending_class_6.1	.546			
3	Working_with_Peers_10.3	.602			
4	Working_with_Peers_outside_10.4	.423			
5	Explanation_instructor_why_small_groups_9.2	.475			
6	Explanation_why_focus_topics_presented_9.3	.522			
7	Confidence_understanding_material_G3.4	.644			
8	Confidence_can_do_3.5	.618			
9	Comfort_lvl_complex_ideas_3.6	.636			
10	Participating_class_discussions_6.2	.487			
11	Listening_to_discussions_6.3	.619			
12	Participating_Group_Work_6.4	.706			
13	Class_Activities_help_lng_6.5	.412			
14	Number_spacing_tests_7.1	.519			
15	Grading_system_what_I_need_to_work_7.4	.552			
16	Feedback_on_my_work_tutorials_7.5	.612			
17	Connecting_key_ideas_other_knowledge_4.1	.561			
18	*Willing_help_others (teacher, peers, TA)_3.7*	.610			
19	Molecular_Forces_1.3.5		.546		
20	SN1_SN2_rxn_mechanism_1.3.9		.520		
21	Distinguishing_Isomers_1.3.10		.769		
22	Classifying_chiral_achiral_molecules_1.3.11		.770		
23	Identifyinig_StereoCentres_1.3.12		.717		
24	Ideas_VS_ideas_other_classes_1.4		.584		
25	Ideas_VS_Major_1.5		.558		
26	Mini_lectures_BlackBoard_8.4			.796	
27	Pencasts_8.5			.869	
28	Pencasts_solutions_HW_problems_8.6			.885	
29	Interacting_with_Instructor_office_hrs_10.2			.460	
30	Argument_use_of_evidence_G2.2				.888
31	Identify_Data_Pattern_G2.1				.680
32	Develop_logical_argument_2.3				.801

Extraction Method: Principal Axis Factoring.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 5 iterations.

Appendix 3: Confirmatory Factor Analyses

Notes for Model (Default model)

Computation of degrees of freedom (Default model)

Number of distinct sample moments:	465
Number of distinct parameters to be estimated:	80
Degrees of freedom (465 - 80):	385

Result (Default model)

Minimum was achieved
Chi-square = 619.406
Degrees of freedom = 385
Probability level = .000

Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	80	619.406	385	.000	1.609
Saturated model	465	.000	0		
Independence model	30	3327.417	435	.000	7.649

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	.098	.801	.760	.663
Saturated model	.000	1.000		
Independence model	.470	.171	.113	.160

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.814	.790	.920	.908	.919
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Parsimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI
Default model	.885	.720	.813
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

NCP

Model	NCP	LO 90	HI 90
Default model	234.406	170.360	306.367
Saturated model	.000	.000	.000
Independence model	2892.417	2712.586	3079.624

FMIN

Model	FMIN	F0	LO 90	HI 90
Default model	4.048	1.532	1.113	2.002
Saturated model	.000	.000	.000	.000
Independence model	21.748	18.905	17.729	20.128

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.063	.054	.072	.012
Independence model	.208	.202	.215	.000

AIC

Model	AIC	BCC	BIC	CAIC
Default model	779.406	820.062	1022.362	1102.362
Saturated model	930.000	1166.311	2342.183	2807.183
Independence model	3387.417	3402.663	3478.525	3508.525

ECVI

Model	ECVI	LO 90	HI 90	MECVI
Default model	5.094	4.676	5.564	5.360
Saturated model	6.078	6.078	6.078	7.623
Independence model	22.140	20.965	23.364	22.240

HOELTER

Model	HOELTER .05	HOELTER .01
Default model	107	112
Independence model	23	24