2013

Using Structural Equation Modeling To Understand Chemistry Faculty Familiarity of Assessment Terminology: Results from a National Survey

Jeffrey R. Raker
Iowa State University

Mary Emenike
Rutgers University - New Brunswick/Piscataway

Thomas Holme
Iowa State University, taholme@iastate.edu

Using Structural Equation Modeling To Understand Chemistry Faculty Familiarity of Assessment Terminology: Results from a National Survey

Abstract
Chemistry departments have felt pressure in recent years to produce quality data on student achievement of learning outcomes. External (e.g., accreditation agencies) and internal (e.g., academic deans) entities are demanding regular review of student achievement. It is thus necessary for the chemistry community to develop valid and reliable instruments to assess student learning. With chemistry faculty members' integration into assessment practices, it is important that these faculty members have a sufficient understanding of the quality of and limitations to the interpretation of assessment data. As part of a larger national survey, 1505 chemistry faculty members from a diverse array of postsecondary institutions and teaching experience responded to a series of questions regarding their familiarity with assessment terminology. Advanced confirmatory factor analysis was conducted via structural equation modeling to represent an overall structure of faculty members' assessment knowledge.

Keywords
first-year undergraduate/general, interdisciplinary/multidisciplinary, testing/assessment

Disciplines
Educational Assessment, Evaluation, and Research | Higher Education | Other Chemistry | Science and Mathematics Education

Comments

JOURNAL OF CHEMICAL EDUCATION
Article
pubs.acs.org/jchemeduc

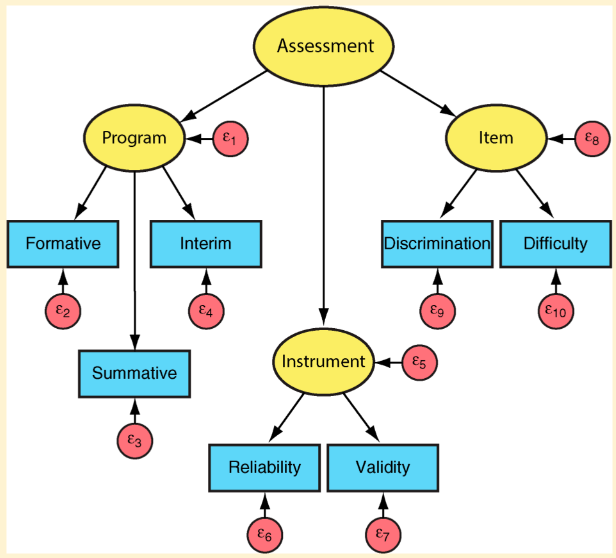
Using Structural Equation Modeling To Understand Chemistry Faculty Familiarity of Assessment Terminology: Results from a National Survey

Jeffrey R. Raker,[†] Mary E. Emenike,[‡] and Thomas A. Holme[†,*]

[†]Department of Chemistry, Iowa State University, Ames, Iowa 50011, United States
[‡]Rutgers Learning Centers, Rutgers University, New Brunswick, New Jersey 08901, United States

S *Supporting Information*

ABSTRACT: Chemistry departments have felt pressure in recent years to produce quality data on student achievement of learning outcomes. External (e.g., accreditation agencies) and internal (e.g., academic deans) entities are demanding regular review of student achievement. It is thus necessary for the chemistry community to develop valid and reliable instruments to assess student learning. With chemistry faculty members' integration into assessment practices, it is important that these faculty members have a sufficient understanding of the quality of and limitations to the interpretation of assessment data. As part of a larger national survey, 1505 chemistry faculty members from a diverse array of postsecondary institutions and teaching experience responded to a series of questions regarding their familiarity with assessment terminology. Advanced confirmatory factor analysis was conducted via structural equation modeling to represent an overall structure of faculty members' assessment knowledge.



KEYWORDS: *First Year Undergraduate/General, Interdisciplinary/Multidisciplinary, Testing/Assessment*

INTRODUCTION

There are many stakeholders associated with higher education, and as a result, there are demands for assessment of colleges and universities from a variety of perspectives. While most states instituted reporting of statistical data from institutions of higher education over a decade ago,[1] many faculty have only started to hear calls for enhanced assessment efforts relatively recently—and most often these calls are associated with accreditation processes.[2,3] Perhaps not surprisingly, survey research has revealed a range of views about assessment, not only among faculty, but also students.[4,5]

These observations are important because faculty attitudes play a key role in the advancement of assessment efforts. When faculty view assessment as a threat to their decision making in teaching, the chances for positive change are diminished.[6] A common way to combat this concern is via faculty development efforts.[7] One potential barrier to faculty development is specialized language usage or jargon. Green found that faculty members appear to have different expectations for the literature in their own academic field compared to literature related to teaching and learning.[8] Nonetheless, sustained professional development appears to be able to improve the understanding of assessment over time.[7]

Some efforts to understand the role of assessment in chemistry have been reported. A group of several authors[9] noted ways in which assessment efforts could be more regularly incorporated into education innovation efforts. Pienta reflected recently about participants at a chemistry education conference that "a few too many people seemed to be saying that they didn't know much about assessment, how to do it, or even why they should."[10] Similarly, Bretz, in a recent editorial,[11] anecdotally described a lack of assessment knowledge for faculty from all chemistry subdisciplines. These observations about assessment familiarity in the chemistry academy suggest that more quantitative results may be desirable. Understanding the current state of assessment terminology via a generalizable national survey holds the promise of informing professional development efforts to improve the use of assessment in educational innovation. In so far as chemistry faculty members are representative of higher education faculty more generally, lessons learned from this study may also be applicable in the

Published: May 30, 2013

broader context of professional development within higher education.

In this report, results are described from the application of structural equation modeling (SEM) to data from a national survey conducted to understand chemistry faculty perceptions and use of assessment. The intended audience for this study is twofold. The first aim is that such modeling can serve to inform the broader chemistry community and provide individual faculty members with a comparative reference for their own assessment knowledge and practices. Second, the nature of the SEM analysis itself is likely to be of interest to individual chemistry faculty and chemistry education research faculty members seeking advanced methodologies for exploring survey data. Thus, additional information about the nature of the method and how to judge aspects such as goodness of fit is included in order to facilitate this second goal more completely.

■ METHODOLOGY

Survey Design

In the fall of 2009, four focus groups were held at regional American Chemical Society (ACS) meetings to gather information about chemistry faculty members' use and understanding of assessment instruments and techniques. Approximately 40 faculty members in total participated in the four different focus groups. One theme from these focus groups, as described by the participants, was confusion related to assessment terminology. Participants noted specific jargon that they had heard but were unfamiliar with, or jargon with which they were comfortable but found their chemistry colleagues were unfamiliar. Likewise, focus group leaders contributed additional terminology to gage the participants' familiarity.

An online pilot survey was subsequently developed and administered to 24 chemistry faculty members in the spring of 2010. These faculty members included tenured, tenure-track, and nontenure-track professors and instructors at two-year, four-year, and doctoral colleges and universities. Participants were provided with an opportunity to comment and provide feedback on each of the survey questions and question categories.

The final survey was constructed from these suggestions. Survey questions included prior experience with assessment, department-level assessment, use of ACS standardized examinations, professional development related to assessment, and familiarity with assessment terminology. (It is this last question category, assessment terminology familiarity, that is the focus of this paper. See Emenike, Schroeder, Murphy, and Holme[12] for a discussion of the entire survey and an overall summary of the data.) Demographic data on participants' sex, years teaching chemistry, primary area of specialization (e.g., inorganic, analytical), and institution type were additionally gathered. Institutional type was defined by the highest chemistry degree awarded at the participant's institution: two-year (associate's degree), four-year (master's degree), doctoral (doctoral degree).

Familiarity with Assessment Terminology

Thirteen assessment terms were identified during the development of the survey (see Table 1). These terms emerged during the participant focus groups and, in the opinion of the authors, were conducive to successful development and implementation of individual, departmental, and university-wide assessment initiatives. Survey participants were asked to use a Likert-scale to rate how familiar they were with each of the 13 terms. "Familiarity" was chosen as a measure to determine which terms related to assessment are perceived by the participants to be "jargon". Participants could respond with these statements:

> "I have never heard this term before."
> "I have heard this term before but do not know what it means."
> "I have heard this term before but am not confident I know what it means."
> "I have heard this term before and have a sense of what it means."
> "I am completely familiar with this term and know what it means."

Table 1. Assessment Terms

Theoretical Groupings	Terms
Program Assessment	Formative assessment
	Summative assessment
	Interim assessment
Instrument Assessment	Assessment validity
	Assessment reliability
Item Assessment	Item response theory
	Item difficulty
	Item discrimination
General Statistics	Linear correlation coefficient
	Cronbach's α
	ANOVA
	Factor analysis
	Variance

The 13 terms were divided into four theoretical groupings for analysis of the "familiarity" data:

1. Program assessment
2. Instrument assessment
3. Item assessment
4. General Statistics terms

Formative, summative, and interim assessment are grouped as *Program* assessment; these terms refer to how homework, quizzes, and exams are utilized to measure learning. Validity and reliability are terms that refer to the development of assessment *Instruments*; these terms refer to the truthfulness of the instrument and the reproducibility of results. Item response theory, item difficulty, and item discrimination refer to the development and analysis of assessment *Items*. Finally, linear correlation coefficient, Cronbach's α, ANOVA, factor analysis, and variance are *General Statistics* terms. These four groupings will serve as a theoretical framework for conducting initial data analyses, a necessary precursor to proposing and testing structural equation models.

Survey Participants

An invitation to participate in the study was sent to approximately 14,000 chemistry faculty members across the United States of America in the summer of 2010. "Faculty members" included tenured, tenure-track, and nontenure-track professors and instructors. After preliminary analyses, it was determined that the participation of two-year institution faculty was low; the full survey was readministered to faculty members of the 2-Year College Chemistry Consortium (2YC3). Final participation included 1546 faculty members (approximately 10% response rate); data were scanned for duplicate entries with such data being removed. A Fisher Exact Test[13] comparison of the percentage institutional type of the population versus the respondent sample returned a nonsignificant p-value ($p = 0.1116$), suggesting that the sample is representative of the population. Only 1505 participants responded to all 13 items of the familiarity with assessment terminology question and thus are included in the data analysis within this paper.

Participants' characteristics were as follows:

- 63% male, 35% female, with 2% preferring not to say/blank/other
- The average number of years teaching chemistry was 15 years
- Areas of specialization included 28% organic chemistry, 18% inorganic chemistry, 17% physical chemistry, 13% analytical, 10% chemistry education, 9% biochemistry, 6% other/blank
- 51% were from four-year institutions, 28% were from doctoral institutions, 21% were from two-year institutions

Data Analysis: Structural Equation Modeling

Structural equation modeling was used as the statistical analysis technique for understanding faculty members' familiarity with assessment terminology; SEM models were analyzed with STATA Version 12 software.[14] Broadly defined:[15]

> *[S]tructural equation modeling (SEM) is a statistical methodology that takes a confirmatory (i.e., hypothesis-testing) approach to the analysis of a structural theory bearing on some phenomenon.*

The goal of SEM is to elucidate causal relationships between variables, either measured (i.e., observed) or latent, within a data set. As is true in any statistical method, the goal of SEM models is to describe the data with as good of fit as possible.[16] Models define theoretically appropriate relationships between the measured and latent variables; "good models" are defined as explaining the variability in the data set better than when no model is defined, in other words, with no defined relationships.[17]

For SEM, any proposed model must fulfill minimum requirements on several "goodness-of-fit" (GOF) statistics. Essentially, all GOF measures check the SEM model versus the null hypothesis, which in this case is a model that restricts covariance of any modeled variables to be equal to the sample covariance.[16] The null hypothesis overspecifies the relationships present in the data. The main GOF statistic is the likelihood ratio chi-squared (χ^2), which measures the cumulative difference between variance predicted by the model and that predicted by the null hypothesis.[17] For models with a high number of degrees of freedom, Kline[17] has suggested a more appropriate measure of the χ^2 statistic: the ratio of the χ^2 value to the number of degrees of freedom should be less than or equal to three. The Comparative Fit Index (CFI) and Tucker–Lewis Index (TLI) measure how well the data fit the overidentified model better than a model in which variables have no relationship; the TLI adjusts for model complexity, that is, the number of defined relationships. CFI and TLI values should approach unity, with ≥0.95 values considered good.[16,18] Finally, the Coefficient of Determination (CD) is an R^2 statistic for the overidentified model; as with R^2 in linear regression modeling, values approaching unity are desirable. If a "majority of the [GOF] indexes indicate a good fit, then there is probably a good fit."[18] For each model described here, the full suite of GOF statistics described above will be given.

In SEM models, path coefficients estimate the strength of the relationship between the connected variables. Standardized path coefficients will be reported for each model (see Figure 1, an example SEM), while unstandardized path coefficients can be found in the Supporting Information. Reported standardized path coefficients can be interpreted in the same way as standardized linear regression coefficients; one standard deviation increase in the independent variable results in a standard deviation increase in the dependent variable equal to the value of the path coefficient (e.g., 0.95 and 0.73 in Figure 1). "SEM estimates all coefficients in the model simultaneously. Thus, one is able to assess the significance and strength of a particular relationship in the context of the complete model."[16]

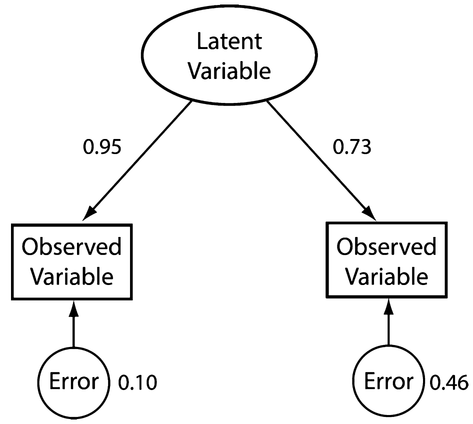


Figure 1. An example SEM.

The last general understanding of structural equation modeling necessary for interpreting models is the error values for each of the exogenous (i.e., internal) variables (see Figure 1 for an example). Reported as a latent variable within a circle, an error value is understood as the fraction of the corresponding variable's variability that is not explained within the model. For example, a value of 0.46 would mean that 46% of the variability of that variable is not explained by the overall model. Acceptable limits are determined through inspection and in the context of the overall model; low values are desirable.

■ RESULTS AND DISCUSSION

Summary Statistics

Prior to presenting any structural equation models, it is helpful to understand the "raw" participant familiarity responses to the 13 assessment terms. Table 2 reports the percentage of each response for the 13 terms; mean values were calculated by setting the "never heard" response equal to one, "no meaning" response equal to two, ..., "completely familiar" response equal to five. While mean values are not considered an appropriate statistical description of ordinal data, such values are necessary for conducting structural equation modeling on ordinal data and are thus reported. (Standard deviations and correlation coefficients, two additional and necessary general statistics for computing structural equation models, are reported in the Supporting Information.)

Two observations should be made of Table 2. First, there is very little uniformity of response distributions for the 13 assessment terms; that is, familiarity with one of the terms is not necessarily conducive to qualitatively predicting familiarity on another term. *Linear correlation coefficient* and *variance* are

Table 2. Percent Response to Familiarity with Assessment Terminology

Terms	Participant-Selected Responses, %					
	Never Heard	Heard, No Meaning	Heard, Not Confident	Heard, Sense of Meaning	Completely Familiar	Mean Total
Formative assessment	27	13	18	22	20	2.9
Summative assessment	28	12	16	23	21	3.0
Interim assessment	33	11	20	24	12	2.7
Assessment validity	12	9	24	39	16	3.4
Assessment reliability	12	10	23	39	16	3.4
Item response theory	56	15	15	10	5	1.9
Item difficulty	12	8	18	37	25	3.5
Item discrimination	29	13	20	24	14	2.8
Linear correlation coefficient	14	8	14	23	40	3.7
Cronbach's α	78	9	6	4	4	1.5
ANOVA	38	14	13	17	17	2.6
Factor analysis	28	17	21	20	14	2.7
Variance	7	10	19	29	36	3.8

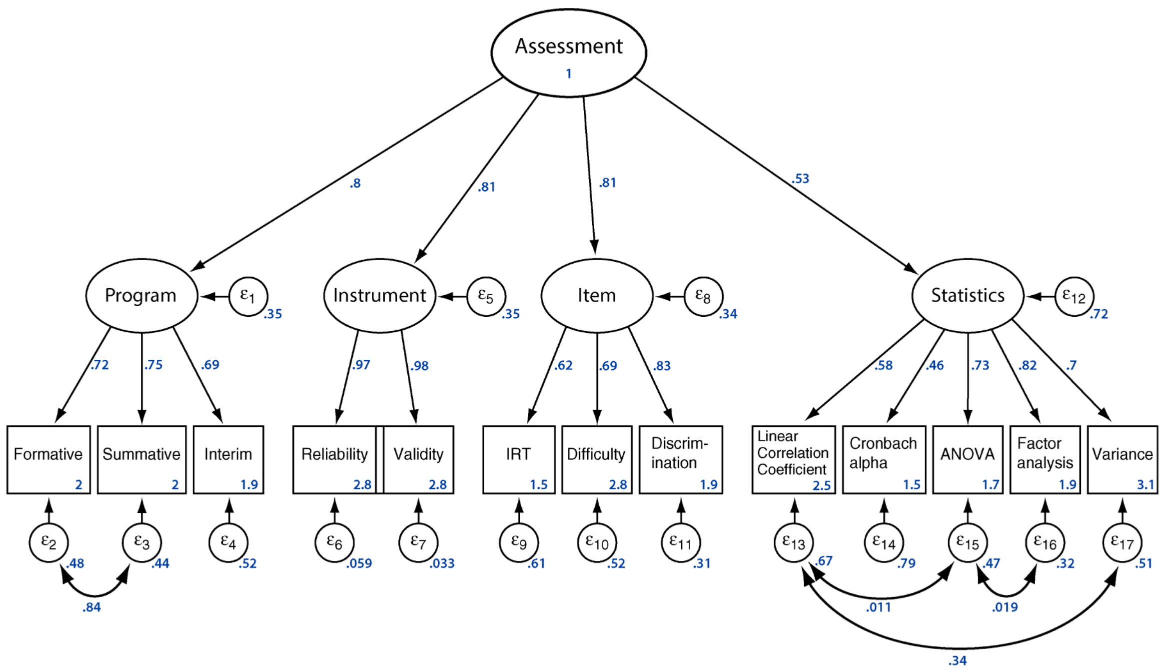


Figure 2. Model 1: Select error variables allowed to covary, standardized path coefficients.

the two most familiar terms; this result is not surprising given the role both terms have in statistical analysis of chemical data in analytical and physical chemistry. *Item response theory* and *Cronbach's α* are the two most unfamiliar terms; greater than half the sample report having "never heard" these two terms. The second observation is that the proposed theoretical groupings of terms show no particular uniformity.

Structural Equation Models

The advantage to applying structural equation modeling to survey data is that understanding beyond individual survey items can be garnered. Variability in individual items (e.g., familiarity) can be described with simpler analyses. In this investigation, for example, familiarity with assessment terminology is not isolated. When necessary, chemistry educators can access definitions to any one of the 13 terms in this study; however, this ultimately ignores how people learn, and connect new learning with previous learning. Clustering of information is more cognitively efficient. The participants in this study are considered to be expert learners; therefore, it is important to evaluate the underlying structures of their learning, using that understanding to develop more holistic materials for professional development. Structural equation modeling provides the analytical tool to arbitrate between differing models that explain response variability that can in turn provide guidance for the creation of such professional development materials.

Structural equation modeling analysis of faculty familiarity with assessment terminology was thus used to test the hypothesis that a cognitive structure of assessment knowledge for chemistry faculty can be inferred from the needs assessment data. This hypothesis arose from the four theoretical groupings of the 13 jargon terms, as previously discussed, and from observations made by the facilitators of the initial workshops in which the structure of the needs assessment was established. Essentially, understanding over the sample (as opposed to any individual in the sample) appears to be hierarchal. Some faculty members have only a very broad understanding of assessment terminology, while others have more nuanced understanding in addition to this broad view. Within this overall hierarchy, structural equation models were proposed, relating the

dx.doi.org/10.1021/ed300636m | J. Chem. Educ. 2013, 90, 981−987

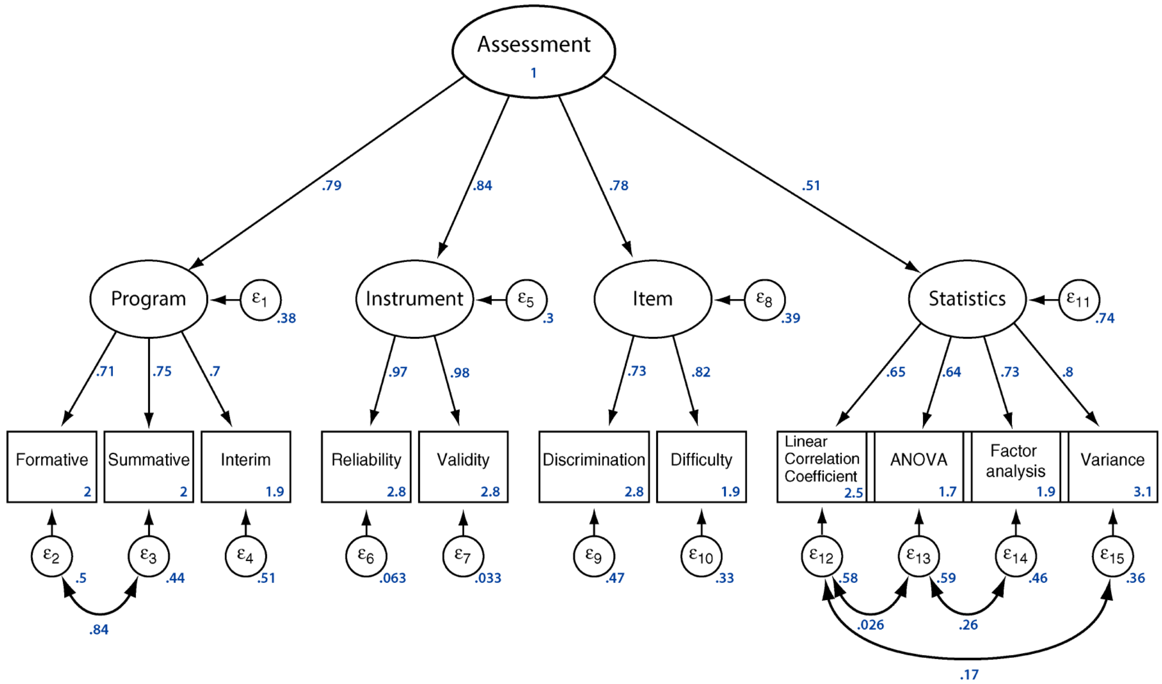


Figure 3. Model 2: Select error variables allowed to covary, standardized path coefficients.

observed familiarities and latent understandings of assessment representing the four theoretical groupings. Visual depiction of these models show observed variables in squares and latent variables in ovals. Analysis of several test structural equation models led to three progressively more parsimonious models.

The first proposed model sought to explain the overall structure of the 13 terms, and the error variables for several observed variables were allowed to covary (see Figure 2). This is the most complex model, representative of all observed measures and theoretical groupings. Each of the 13 jargon terms is represented as a rectangle because they are all measured observations. The four theoretical groups (i.e., program, instrument, item, and statistics) are represented as ovals because they are latent. In the same way, the overarching assessment construct is an oval. Directionality of the connecting arrows demonstrates the causality of the relationships between variables. Thus, a participant's composite familiarity of assessment is causally related to that participant's familiarity of program assessment terminology. In turn, a participant's familiarity of program assessment is causally related to that participant's measured familiarity of formative, summative, and interim assessment. Error measures, represented as circles, provide a measure of how much of the variable's variability is explained by the model; a double-headed arrow connecting to error variables is interpreted as the error of the two variables covaries. Such covariances are introduced while attempting to improve the overall goodness-of-fit for a given model; note that the connectivity of latent and observed variables remains unchanged.

Model 1 resulted in a χ^2 value of 548.68 (df = 57; $p <$ 0.0000), CFI and TLI values of 0.959 and 0.944, and a coefficient of determination of 0.859; while the CFI, TLI, and coefficient of determination are near or within acceptable limits, the χ^2 value is not. Adjusting the χ^2 value by dividing by the number of degrees of freedom[17] gives a resultant value of 8.99, which falls above the 3.00 limit, suggesting that the model is not a good fit for the data.

While the goodness-of-fit data for Model 1 do not suggest a good fit, it does provide some understanding of the relationships of the variables that informs development of further models. Pathway coefficients between latent variables and between latent and observed variables are high (greater than 0.46; with a majority greater than 0.70). Error variables for the endogenous variables range from 0.033 to 0.79; these can be translated as 3.3 to 79% of the variability of the corresponding variable is "not" accounted for in the model. Some specific aspects of this model, particularly the knowledge of terms related to assessment instruments, appears to be fit well, suggesting that this component is likely to be helpful in other models. Similarly, the covariance of error variables for *formative* and *summative assessment* (depicted by the curved double arrow on the lower left) can be explained by the connected theoretical relationship of the two variables. However, the covariances of General Statistics (similar curved double arrows near the bottom right) terms are more difficult to explain. Therefore, a more parsimonious model was sought to explain observed variability in familiarity.

One potentially helpful observation is that item response theory (IRT) and Cronbach's α (Cronbach alpha) were both rated quite low on the familiarity scale. With such generally low familiarity, there is an inherent lack of variability in responses that in turn appears to lead to nonideal behavior in the SEM models. Thus, a new model that removes these terms is posited. While removing these terms, it is not obvious that covariances of error should be included, but without them, GOF estimates are weaker. Model 2 is presented in Figure 3 and GOF results for this model include a χ^2 value of 185.88 (df = 36; $p = 0.0000$; χ^2 adjusted for df = 5.16), CFI and TLI values of 0.986 and 0.979, and a coefficient of determination of 0.855. The adjusted χ^2 value (5.16) is closer to the accepted threshold of three, but ultimately, this model cannot yet be categorized as a good fit.

By inspecting the standardized path coefficients for Model 2, it can be observed that the General Statistics (Statistics) latent variable does not load onto the overall Assessment Familiarity

dx.doi.org/10.1021/ed300636m | J. Chem. Educ. 2013, 90, 981−987

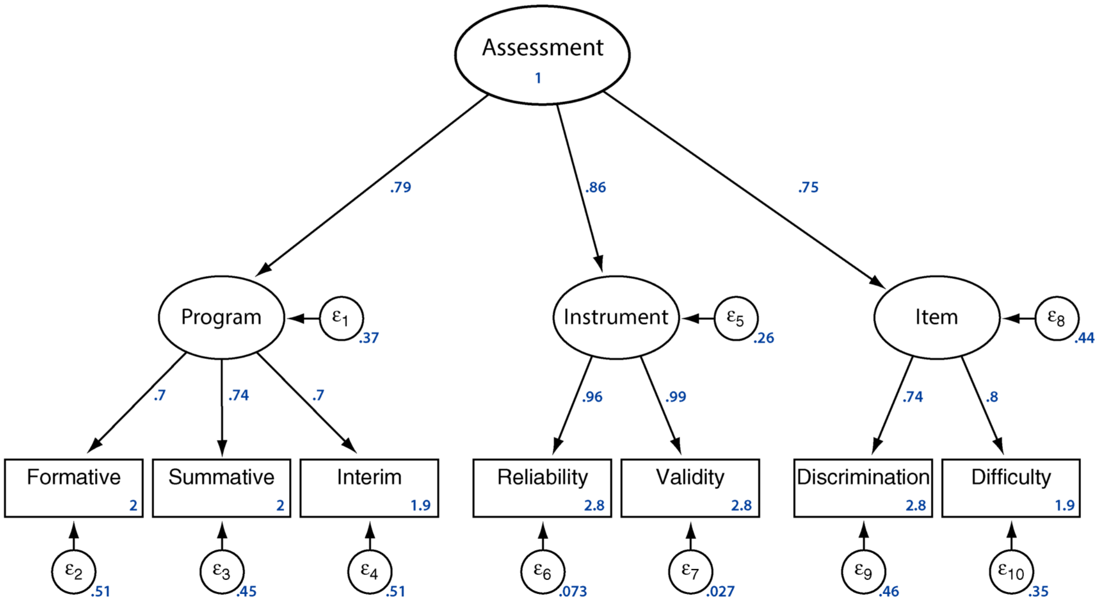


Figure 4. Model 3: Select error variables allowed to covary, standardized path coefficients.

variable in the same manner as the Program, Instrument, and Item Assessment variables (0.51 versus 0.79, 0.84, and 0.78). One might imagine that familiarity of chemists with terms from general statistics need not vary in any way similarly to their understanding of educational assessment as a result of the use of general statistics in other aspects of their professional lives. Further evidence that this aspect of instructor familiarity is different lies in the error of the General Statistics variable, which is 0.737. This means that 73.7% of the variability of the General Statistics variable is unaccounted for in Model 2. Thus, while familiarity with statistics can play a role in things such as professional development projects directed toward chemistry instructors, it does not appear to help in the understanding of the structure of knowledge that chemists have for educational assessment.

In so far as the goal of this particular SEM model exercise is to infer an understanding of how chemists build their knowledge structure about educational assessment, an additional model is built that eliminates the General Statistics aspect. Model 3 (see Figure 4) resulted in a χ^2 value of 29.46 (df = 10; p = 0.0011; χ^2 adjusted for df = 2.95), CFI and TLI values of 0.998 and 0.995, and a coefficient of determination of 0.854. Thus, Model 3 is a model in which acceptable values for the goodness-of-fit statistics are achieved.

While additional adjustments can provide other acceptable models, Model 3 suggests an overall cognitive structure of faculty knowledge of assessment that is comprised of three theoretical groupings: Program Assessment, Instrument Assessment, and Item Assessment. The overall Assessment factor is composed more of Instrument Assessment (path coefficient = 0.86) than Item and Program Assessment (0.75 and 0.79, respectively). Error values for the three theoretical groupings are inversely proportional to their path coefficients to the overall Assessment factor.

Looking at this model, it might be argued that familiarity with *Formative* and *Summative Assessment* as a function of institution type should be included in this model. Such a model would arise from the hypothesis that some institutions types may have more of a focus on assessment than others. Additionally, the relationship between submitting National Science Foundation TUES and CCLI grants for educational innovation projects might also contribute to an understanding of program assessment terms. Such grants may be more commonly associated with doctoral and four-year institutions, so once again, additional models could be hypothesized that provide interesting structures related to faculty knowledge of assessment. Familiarity with *Item Difficulty* and *Discrimination* could be affected by the use of ACS Examinations, given that difficulty and discrimination are included in nationally normed data. These influences (i.e., institution type and ACS Exam usage) could be built into further models; however, the proposed variable relationships in Model 3 sufficiently account for observed variability and thus offer a potentially useful perspective of the structure of faculty assessment knowledge.

■ CONCLUSION

This project has attempted to identify areas of assessment, including terminology, in which chemistry instructors perceive some lack of knowledge. Measures of familiarity derived from self-reported data about assessment terminology may provide a template for identifying areas for targeted professional development on assessment issues. Professional development inevitably competes for time with many other components of academic life. Therefore, an initial understanding of the most common structure of current knowledge about assessment terminology affords the prospect that professional development can be more efficiently targeted. This prospect is advanced by the analysis of a national survey of chemistry instructors' knowledge using a structural equation model for the survey data. The resultant SEM model suggests three assessment term pairs that when discussed in tandem will increase faculty members' overall familiarity with assessment. The ACS Exams Institute has devised grant proposals that leverage this knowledge in the hopes of providing professional development materials that might be of the most use to the chemistry education community.

Consideration of the data by demographic groups in further studies may provide the possibility of even more targeted areas for improving faculty members' understanding of assessment terminology. Additionally, for the community of chemistry educators, future work exploring the impact of ACS Examinations usage and the composite culture of assessment may include specific components related to assessment terminology familiarity. Such work can help identify how the availability of a discipline-based assessment organization like ACS Exams can affect the broader educational community, even in the majority of disciplines that have no such organization.

■ ASSOCIATED CONTENT

Supporting Information

Descriptive statistics for familiarity with assessment terminology responses; correlation table for familiarity with assessment terminology responses; unstandardized path coefficients for the reported SEM Models. This material is available via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: taholme@iastate.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We would like to thank the 1500+ faculty members who graciously gave of their time to participate in focus groups and complete the pilot or final survey. Work on the development of the needs assessment survey by Kristen Murphy and Jacob Schroeder is gratefully acknowledged. Funding for this project was provided by the National Science Foundation (DUE-0920266).

■ REFERENCES

(1) Burke, J. C.; Minassians, H. *Performance Reporting: The Preferred "No Cost" Accountability Program*; Nelson A. Rockefeller Institute of Government: Albany, NY, 2001.

(2) Towns, M. H. *J. Chem. Educ.* **2010**, *87*, 91.

(3) Klein, S. P.; Kuh, G. D.; Chun, M.; Hamilton, L.; Shavelson, R. *Res. Higher Educ.* **2005**, *46*, 251.

(4) Brown, G. T. L. *Assess. Educ.: Princ. Policy Pract.* **2004**, *11*, 305.

(5) Brown, G. T. L. *Psychol. Rep.* **2006**, *99*, 161.

(6) Welsh, J. F.; Metcalf, J. *J. Higher Educ.* **2003**, *74*, 445.

(7) Haviland, D.; Shin, S.; Turley, S. *Innovative Higher Educ.* **2010**, *35*, 61.

(8) Green, D. A. *Int. J. Acad. Dev.* **2010**, *15*, 47.

(9) Holme, T.; Bretz, S. L.; Cooper, M.; Lewis, J.; Paek, P.; Pienta, N.; Stacy, A.; Stevens, R.; Towns, M. *Chem. Educ. Res. Pract.* **2010**, *11*, 92.

(10) Pienta, N. *J. Chem. Educ.* **2011**, *88*, 1199.

(11) Bretz, S. L. *J. Chem. Educ.* **2012**, *89*, 689−691.

(12) Emenike, M.; Schroeder, J. D.; Murphy, K.; Holme, T. *J. Chem. Educ.* **2013**, *90* (5), 561−567.

(13) Sheskin, D. J. *Handbook of Parametric and Nonparametric Statistical Procedures*; 4th ed.; Chapman & Hall/CRC: New York, 2007.

(14) StataCorp *Stata Statistical Software: Release 12*; StataCorp LP: College Station, TX, 2011.

(15) Byrne, B. M. *Structural Equation Modeling with AMOS: Basic, Concepts, Applications, and Programming*; Routledge, Taylor & Francis Group: New York, 2010.

(16) Dion, P. A. *J. Bus. Ethics* **2010**, *83*, 365.

(17) Kline, R. B. *Principles and Practice of Structural Equation Modeling*; 3rd ed.; The Guilford Press: New York, 2011.

(18) Schreiber, J. B.; Stage, F. K.; King, J.; Nora, A.; Barlow, E. A. *J. Educ. Res.* **2006**, *99*, 323.

(19) Xu, X.; Villafane, S. M.; Lewis, J. E. *Chem. Educ. Res. Pract.* **2013**, *14* (2), 188−200.

(20) Brandriet, A. R.; Ward, R. M.; Bretz, S. L. *Chem. Educ. Res. Pract.* **2013**, DOI: 10.1039/C3RP00043E.

■ NOTE ADDED IN PROOF

Recently, two manuscripts have been published within the field of chemical education research that use structural equation modeling methodologies. The first, from Xu et al.,[19] explores student attitudes, knowledge, and achievement in general chemistry. Their SEM models suggest that attitudes play an important role in student achievement in chemistry. The second, from Brandriet et al.,[20] explores the interplay between student thinking, feeling, and performance in the development of chemistry understanding. Their best SEM model suggests that all three dimensions contribute to meaningful student learning.