ISSN 1648-3898

PSYCHOMETRIC RECONSTRUCTION OF THE INQUIRY SCIENCE TEACHING STRATEGIES (ISTS) INSTRUMENT

Sebastian Szyjka,
Frackson Mumba

Abstract. *This study re-evaluated the psychometric properties of the 40-item Inquiry Science Teaching Strategies instrument [ISTS] (Lazarowitz & Lee, 1976). The ISTS was administered to 201 elementary pre-service teachers at a university in the Midwest of United States during the 2009-2010 school terms. Initial CFA results failed to confirm the 3-factor structure of the original ISTS. A statistically significant goodness-of-fit statistic (χ^2 [737, N = 201] = 1394.61, p = 0.00) supported by two fit indices (GFI = 0.69, CFI = 0.45) indicated non-tenability of the measurement model. Subsequent PCA retained half of the components under an obliquely rotated 3-factor solution. Following the removal of 22 items, final CFA indicated non-tenability under slightly improved fitting (χ^2 [132, N = 201] = 236.93, p = 0.00, GFI = 0.88, CFI = 0.83). It is recommended that existing items be revised and new ones appended using Exploratory Factor Analysis with a different sample.*

Key words: *attitudes towards inquiry-based science, confirmatory factor analysis, evaluation of psychometric properties, instrument development, principal component analysis.*

Sebastian Szyjka
Western Illinois University, Macomb, Illinois, USA
Frackson Mumba
Southern Illinois University, Carbondale, Illinois, USA

Introduction

Inquiry-based science teaching has been shown to promote the acquisition of knowledge as well as improve the understanding of the social practices behind empiricism. Consequently, the world has witnessed the proliferation of inquiry-based science teaching as a method to advance learning at both the elementary and secondary levels (Bowen & Bencze, 2009). Reform documents produced all over the globe indicate that inquiry-based science teaching must be further developed and diversified at such levels (e.g., American Association for the Advancement of Science [AAAS], 1993; Council of the Ministers of Education of Canada [CMEC], 1997; European Commission [EC], 2004; National Center for Educational Research and Development [NCERD], 1997; National Research Council [NRC], 1996). For example, European countries have advanced a number of professional development initiatives designed to more adequately underscore inquiry science teaching through unique pedagogical strategies and perspectives – some of which are believed to increase student motivation and attainment (Osborne & Dillon, 2008). In describing international issues with the enactment of inquiry-based science teaching among countries such as the United States, Lebanon, United Kingdom, Israel, Venezuela, Australia and Taiwan, Abd-El-Khalick et al. (2004) attribute overreliance on didactic teaching, engagement in a disproportionate amount of verification-type laboratory activities, teaching irrelevant historical facts or rote information, as well as using mandated textbooks to follow rigid scope and sequences

ISSN 1648–3898

as factors hindering the realization of true inquiry-based science instruction. They uncovered that the relative success of science education often depends on environments comprised of properly defined inquiry and mindsets among teachers and students that facilitate positive dispositions.

Defining inquiry can be categorized as either teaching and learning science through inquiry or science as inquiry (Chiappetta & Koballa, 2006). Learning science through inquiry affords students the opportunity to partake in inquiry investigations designed to assist them in achieving greater levels of conceptual understanding and knowledge. This includes the development of a wide variety of inquiry skills such as identifying problems for study, formulating hypotheses, designing experiments, collecting and analyzing data, and interpreting results in order to draw more meaningful conclusions (AAAS, 1993). On the contrary, teaching science as inquiry assists students in understanding how scientific knowledge is advanced. This includes viewing science as a series of non-temporal methods through which knowledge is acquired. Hence, it is important for researchers in science education to carefully define inquiry, considering how its characteristics can not only shape the instructional events that take place in the classroom, but also shape the measures designed to assess any number of constructs related to this skill. Being that inquiry can vary from context to context, such as in the case of being highly structured or prescriptive to being more open-ended and unstructured (Lunetta, Hofstein, & Clough, 2007), it is critical that researchers sufficiently investigate the relationship between how teachers understand inquiry with their affinity towards it. To achieve this, measures must be capable of determining the extent of like or dislike toward specific aspects of inquiry. This would include understanding the perceived levels of necessary interaction required among teachers and students during the process as well as how supplemental materials or resources could play a role in learning. Knowing that attitudes in science are not comprised of any single or individual construct, but rather of several inter-related constructs in varying amounts and intensities, can be useful when informing evaluation and subsequent development of measures designed to assess attitudes towards science (Osborne, Simon, & Collins, 2003) or attitudes towards the processes of inquiry.

In consideration of how inquiry-based science is defined, should appear in various teaching and learning contexts, as well as be represented in various measures, studies have pointed out continual disconnects between inquiry-based science and the accepting of practices related to its implementation (Costenson & Lawson, 1986; Damnjanovic, 1999; Lazarowitz & Lee, 1976; Lazarowitz, Barufaldi, & Huntsberger, 1978; Windschitl, 2002). Reasons for inadequate attention to inquiry-based practice include time constraint issues, problems with classroom management and funding (Costenson & Lawson, 1986; Windschitl, 2002), as well as low levels of confidence associated with self-efficacy (Smolleck, Zembal-Saul, & Yoder, 2006). Other areas that are thought to contribute to this disconnect include limited science subject matter knowledge and a lack of science pedagogical content knowledge (Shulman, 1986). Several researchers have addressed these areas by developing instruments designed to measure affective constructs related to the practices of inquiry (Damnjanovic, 1999; Lazarowitz & Lee, 1976; Smolleck et al., 2006). Although there are a cache of instruments designed for this purpose, some have examined constructs not directly related to the majority of the process skills inherent of inquiry (Damnjanovic, 1999) or have employed outmoded statistical procedures for the establishment of validity and reliability (Lazarowitz & Lee, 1976).

Despite the work by many, issues continue to surface regarding the psychometric validation of instruments developed prior to the advent of advanced statistical procedures for establishing sound validity and reliability. For example, in Lazarowitz and Lee's (1976) attempt to develop an instrument to measure inquiry attitudes among secondary science teachers, there exist some questionable elements of their establishment of validity and reliability – primarily because of the limited types of procedures for analysis available at the time as well as a restricted number of respondents used to demonstrate certain types of validity (e.g., construct – known groups). Hence, the purpose of this study was to re-evaluate the psychometric properties of the 40-item Inquiry Science Teaching Strategies (ISTS) instrument in accordance to more recent methods of instrument validation and offer a reconfigured instrument assuming the utilization of more current statistical procedures.

Methodology of Research

ISTS Instrument and Administration

The ISTS instrument was developed by Lazarowitz and Lee (1976) to measure inquiry attitudes of secondary science teachers. On the grounds that science curricula advocate various forms of inquiry-based instruction, as well as citing the importance of investigating variables related to positive attitudes towards inquiry strategies, the ISTS was constructed and subsequently used (Lazarowitz et al., 1978). The ISTS measures three constructs related to inquiry-based learning: Laboratory Investigations (LAB), Textbooks Used (TEXT) and Classroom Teacher-Student Interaction (INT) (see Table 1 for list of items and their designated constructs). Items were selected in accordance to relationships present during classroom inquiry-based teaching and were divided to represent both positive and negative statements towards various inquiry-based and non-inquiry-based teaching scenarios. A five point Likert-scale indicating strongly agree, agree, undecided, disagree, or strongly disagree designated the range of possible responses. Respective values of 5 to 1 were valued for positively worded items and 1 to 5 for negatively worded items.

Content validity was established after the ISTS was reviewed by a number of faculty members in science education. Following this process, items were reworded for readability. Item analysis consisted of calculating bi-serial correlations, which led to the finalization of a 40-item instrument. Construct validity was established using the technique of known groups (Shaw & Wright, 1967), a procedure designed to compare the performance of two groups when known differences on the construct being measured are apparent. Purportedly, results from this technique demonstrated that the ISTS could discriminate various criteria established within the constructs of the instrument. Lazorawitz and Lee reported alpha coefficients of 0.54 (N = 7), 0.48 (N = 16), 0.76 (N = 47), 0.69 (N = 44), and 0.85 (N = 735) for groups 1 – 5, respectively. Overall differences among these groups was statistically significant (F (4, 848) = 27.72, $p <$ 0.01), indicating the ability of the instrument to determine what extent of the population held favorable attitudes towards inquiry teaching strategies. However, no evidence was offered regarding a related feature of validation – the psychometric appraisal of the instrument's dimensional structure. According to more recent methods of validation, this type of evaluation yields more discriminable dimensions for the constructs being measured (Lichtenstein et al., 2008), resulting in more empirically robust indicators of reliability and validity for the overall instrument as well as the constructs subsumed within.

The 40-item ISTS requires respondents to indicate their extent of agreement to statements on a 5-point Likert-type response scale with a mid-point of undecided. Any references made to "secondary science courses" or "secondary science teachers" found in the ISTS instrument were changed to instead represent "elementary science courses" or "elementary science teachers". This modification more adequately resembles the context of elementary pre-service science teacher education without disturbing the inherent characteristics of the original ISTS instrument. Any male biased language was made to instead represent both male and female genders. The ISTS was distributed to respondents on a single sheet of paper where 20 items appeared on each side; answer sheets supplemented the presentation package.

Table 1. Item listing according to 3-factor structure for the ISTS.

Laboratory Activities (LAB)
1 A science teacher should encourage students to critically analyze their own conclusions.
5 Experimental results which cannot be interpreted show that the experiment is not appropriate for secondary science courses.
6 Students are often capable of designing valid experiments.
8 If unexpected results are obtained, they should be included in the analysis of laboratory work.
9 Each day's lesson should be based on the previous lessons.
11 Conflicting data can lead to a useful post-laboratory discussion.

ISSN 1648–3898

Laboratory Activities (LAB)
12 In an investigation students should know from the beginning the steps they will perform.
15 A student is successful in a laboratory experiment if his results are similar to the class results.
17 At the end of an experiment, the science teacher should analyze the results to help students understand them.
21 Science teachers should make clear in advance all the problems which arise in the performance of a laboratory experiment.
22 In a science course students should develop skill in interpreting data.
24 Students will perform experiments successfully when the teacher presents an overall explanation of the subject to be investigated.
25 A primary role of the secondary science teachers is to design the investigations to be done.
26 Science teaching should enable students to identify the assumptions made in a given investigation.
27 By presenting an acceptable rule to students, the teacher avoids the risk of having them arrive an incorrect one.
30 Unexpected results should be considered as a part of laboratory work.
31 Science teachers should formulate hypotheses when questions are raised by students.
32 Students should be asked to prepare the equipment needed for laboratory work.
34 In science class, students should learn to make careful and relevant observations.
35 In general, it is not practical for students to test their own hypothesis.
36 Teachers, not students, should formulate the problems to be taught in science.
37 Experimental results that differ greatly from what is expected should not be considered.
39 An examination at the end of a science course should ask students to solve problems that they have not seen before.
40 Differences in data can lead to the proposal of alternative procedures in laboratory work.
Textbooks Used (TEXT)
2 Students should be guided to include articles from different scientific journals in their notebooks.
3 A secondary science course should include more learning material than a teacher intends to use.
7 Questions which are integrated in the text are confusing to students and should be omitted.
14 A secondary science course should have laboratory experiments integrated with the text material.
18 Textbooks should contain subject matter which could be covered in one academic year.
19 Scientific journals and reference books should be available for students to use while performing experiments.
28 A science teacher should prevent his students from trying to critique scientific material before they master it.
33 A textbook should contain both the problems to be studied and the answers.
Classroom Teacher-Student Interaction (INT)
4 A science teacher should immediately correct a wrong answer given by a student.
10 Students will learn better when their curiosity is aroused before a subject is studied.
13 It is desirable to present to students science questions to which answers are not necessarily known.
16 For each new topic, generalizations should be presented before examples and illustrations of the generalizations are provided.
20 One of the roles of the classroom teacher is to present learning situations in such a way that students will raise questions.
23 Unstructured activities in the laboratory work may often lead to exciting kinds of science experiences for both the teacher and the student.
29 A science teacher should be receptive of any reasonable answer a student gives.
38 Teaching that introduces a scientific problem should sooner or later lead to its solution.

Sample

The ISTS was administered to a convenience sample of 201 elementary pre-service teachers enrolled in science methods courses during the fall and spring semesters of the 2009-2010 school year at a university located in the upper Midwest of the United States. Acquisition of additional participants was not possible because of the finite nature of student enrollment in the program. Of the 201 participants, approximately 88% were female; all of which were juniors or seniors ranging in age from 20 to 26. While socioeconomic data for the sample is not reported here, the ethnic distribution of the population from which the sample was drawn is as follows: Non Resident Alien 3%; Black Non-Hispanic 11%; Hispanic 2%; Asian/Pacific Islander 1%; American Indian/Alaskan Native 1%; White Non-Hispanic 74%; Race Unknown 7%. The instructors teaching the methods courses agreed to administer the instrument to participants. To ensure that respondents completed all items found on the original ISTS, proctors of the instrument were instructed to promote the completion of all items presented in each distributed package. This study was approved by the appropriate University Institutional Review Board.

Protocol of Analyses

A three-phase data analysis progression necessitated the use of Confirmatory Factor Analysis [(CFA) phase-one] followed by Principal Component Analysis [(PCA) phase-two]. CFA culminated the third phase based on PCA outcomes. Characteristically, CFA is intended for testing statistical postulates that juxtapose theoretical and observed factor structures within an instrument. The utility of CFA for the researcher is to hypothesize a statistical relationship that ultimately accepts the null outcome (desired non-significance), indicating that congruence has been achieved between an originally developed measure (i.e., the theoretical factor structure originally crafted by the author of an instrument using preconceived constructs) with that of the one administered to a similar sample (observed) (Blunch, 2008; Byrne, 2001). As a result, statistically tested relationships can be more closely examined as the validity and reliability of an instrument is determined. Structural Equation Modeling (SEM) using IBM AMOS 19 was used for testing the statistical significance of the hypothesized factor structures originally set forth by the authors of the ISTS.

On the other hand, PCA is commonly employed as a means of developing an instrument when it is not necessarily clear which items correspond to a given construct or the researcher is unsure of which items should load on specific factors. Moreover, PCA can be used to further clarify any reasonable patterns already in existence among pre-determined factors. Outcomes of the PCA procedure often resemble those of factor analysis (e.g., principal axis factoring or maximum likelihood methods), yielding nearly analogous results with less complexity in the calculation. In part, determining which items load more effectively on potential or unknown constructs is the primary objective. This procedure can extend to situations involving the pairing down of items to refine previously designed instruments in anticipation of achieving more psychometric robustness. In this study, PCA articulated a 3-factor extraction rotated to an oblique (Direct Oblim with Kaiser Normalization) solution. This method accounts for correlations among rotated factors. Taken together, these and other considerations for analysis (Field, 2005) with corresponding protocol included the following: (1) Examination of Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy a statistic that indicates whether data is appropriate for component analysis. Minimum values of at least 0.50 are required, while those between 0.70 and 0.90 are considered good to excellent; (2) Examination of Bartlett's Test of Sphericity – a statistic indicating that there are relationships among the variables to be included in analysis. Statistical significance below the 0.05 level is required; (3) Examination of the proportion of variance criterion – initial eigenvalues indicating percentage variance as well as cumulative percentage. These percentages explain the total variance of each component comprising a particular factor. The researcher set a minimum critical value of approximately 70% for retaining a collective set of items; (4) Interpretation of the pattern matrix for emergent factors following preliminary extraction. This matrix contains figures regarding the unique contribution of each variable to a corresponding factor. A loading criterion to suppress absolute values of less than 0.40 was set in order to display only coefficients of significance at that level during factor loading.

PCA was arranged to address anticipated CFA outcomes of an undesirable rendering. Based on this consideration, PCA was an *ad hoc* procedure – viewed as a means to abridge the original ISTS structure for improved utility and psychometric fidelity only when CFA criteria were not met. As such, CFA outcomes would have to yield statistically significant values opposed to the "good fit" standards set forth by many: (1) Chi-square goodness-of-fit-test (Blunch, 2008), with a low statistic indicating good fit. The Chi-square statistic would also have to be non-significant at the 0.05 level to indicate a good fit. In essence, the Chi-square goodness-of-fit test employed in the initial evaluation of the ISTS statistically determines whether there were similarities in observed and expected data set matrices. Typically in a Chi-square goodness-of-fit test with this purpose, one hopes to find non-significant lower Chi-square values, indicative of congruence between the original and current instrument under evaluation. A low Chi-square value reveals little difference between the theoretical model and the current data. Achieving statistically significant Chi-square values (higher) indicate non-congruence between observed and expected data set matrices. Ulti-

ISSN 1648–3898

mately, the Chi-square goodness-of-fit tests for instruments are interpreted as tests of dependence, rather than independence; (2) Goodness-of-fit index [GFI] (Jöreskog & Sörbom, 1988), with values approaching 1 indicating a good fit; (3) Comparative fit index [CFI] (Bentler, 1990), with values approaching 1 indicating a good fit; (4) Root mean square error approximation [RMSEA] (Brown & Cudeck, 1993), with values below 0.05 indicating a good fit, while values no greater than 0.08 are indicative of acceptable fit.

Ultimately, the three-phase data analysis progression was brought to fruition with a final CFA procedure designed to re-evaluate the ISTS according to newly reconstructed factors and calculated factor loadings. In consideration of assessing psychometric validity and reliability, both the CFA and PCA procedures served as a means to improve both the understanding of the original constructs, offer a simplified instrument structure, and suggest revised constructs to load items based on factor calculation outcomes. For additional information concerning the implementation of CFA and PCA data analysis procedures see Blunch (2008), Byrne (2001), DeCoster (1998), and Field (2005).

Results of Research

Confirmatory Factor Analysis (Phase-One)

The initial CFA arrangement hypothesized a 3-factor structure original to Lazarowitz and Lee's ISTS, loading 8 items on the TEXT factor, 8 items of on the INT factor and 24 items on the LAB factor. Figure 1 illustrates the CFA 3-factor structure, item pairing with each respective construct, factor loadings and the explained variance of each item on its factor. Latent variable structures resembling the ISTS include a single testable measurement model with correlated factors. Exogenous variables are represented as ovals, while endogenous variables are represented as rectangles. Associated items for each construct are imbedded in the rectangles. For example, the construct "Classroom Teacher-Student Interaction" is supported by items INT4, INT10, INT13, INT16, INT20, INT23, INT29 and INT38. Small circles converging to each rectangle are measurement error terms (e.g., err1-err40) associated with the observed variable. One-way arrows pointing in the direction of the rectangles denote random and unique measurement error impact. Directional arrows represent unique impact that one variable has on another, while double-headed arrows represent correlations between variable pairs. The arrows pointing in the direction of the rectangles are standardized correlation coefficients, or factor loadings (Byrne, 2001).

The measurement model did not fit the data well, with the calculated Chi-square statistic (χ^2 [737, N = 201] = 1394.61, p = 0.00) indicating implied covariance and sample covariance as statistically significant, yet unacceptable in terms of fit with the data. This signified a substantial departure from the theoretical arrangement with the observed data set. Further evidence corroborated poor-fit results (GFI = 0.69, CFI = 0.45), with the exception of the RMSEA, where a calculated value of 0.07 was attained.

As part of psychometric evaluation and to further examine the internal consistency of the three constructs, reliability statistics were calculated. Results indicate poor and untenable correlations. The coefficients for "Textbooks Used" (0.09) and "Classroom Teacher Student Interaction" (0.36) were low and unacceptable according to determined standards. "Laboratory Activities" (0.64) was also moderate-low. "Textbooks Used" and "Laboratory Activities" (0.68) was moderate-low, while "Laboratory Activities" and "Classroom Teacher-Student Interaction" (0.71) was moderate. Overall, reliability of the 40-item ISTS was 0.73.

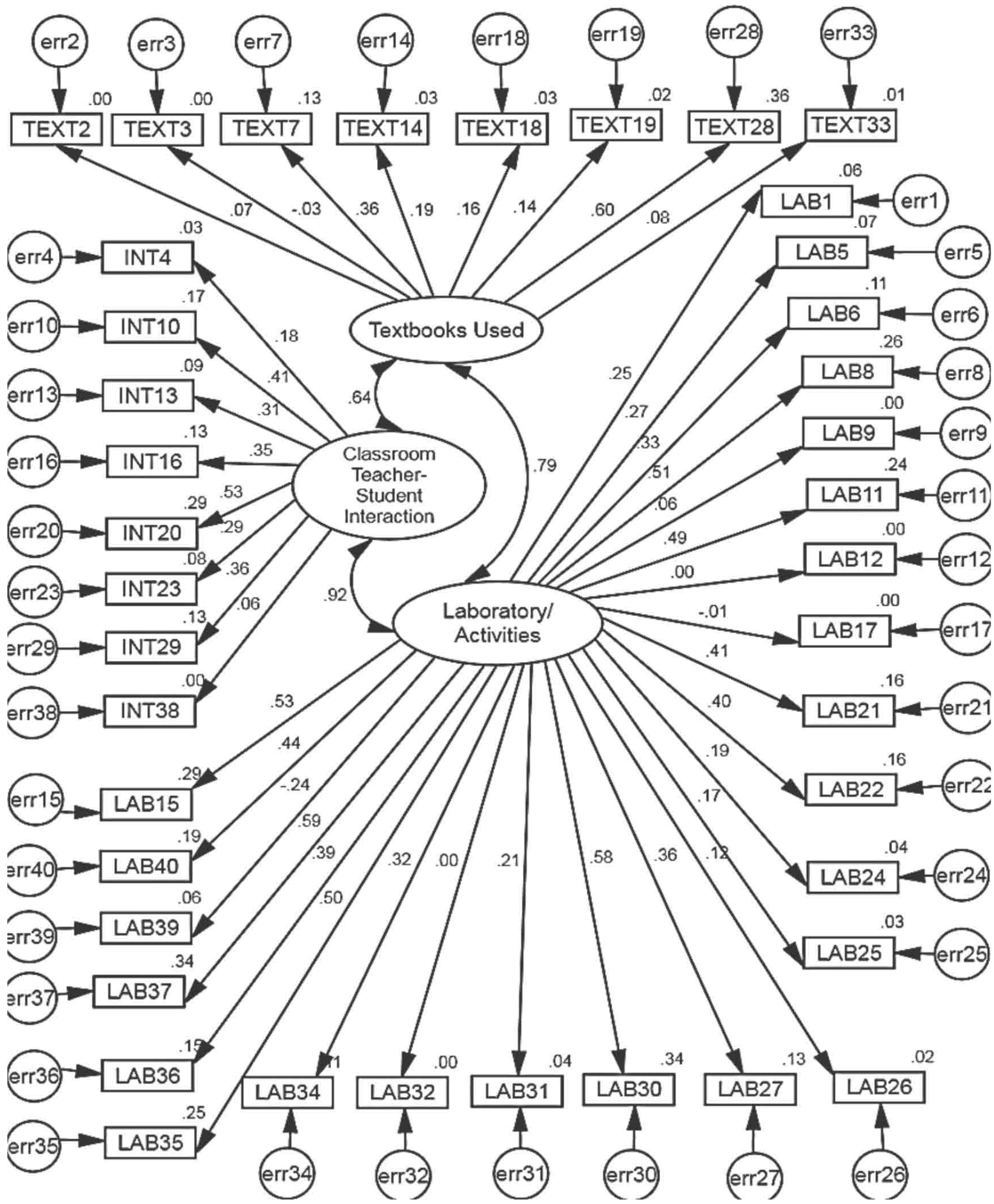

Figure 1: Confirmatory factor analysis of ISTS.

Principal Component Analysis (Phase-Two)

Because of the unacceptable fit of the 3-factor measurement model, PCA was initiated to continue the process of psychometrically refining the 40-item ISTS. Based on the assumption that multiple constructs defined by Lazarowitz and Lee are correlated, the factor models were poised for oblique rotation during analysis. Once these and other SPSS analysis criteria were defined, results indicated the data set was acceptable for processing. Both the KMO (0.70) and Bartlett's Test (χ^2 [780, N = 201] = 1836.40, p = 0.00) specified the appropriate relationships among the necessary variables for analysis. Examination of the proportion of variance criterion indicated that approximately 18 components be

ISSN 1648–3898

retained according to the pre-set minimum critical value of 70%. Remaining components failed to load into any meaningful configuration. This means that half of the original ISTS items were discarded prior to the final extraction. Table 2 shows the full extraction of the pattern matrix. Two double-loaded factors were discarded; loadings smaller than 0.40 do not appear in the table. Resulting factors were interpreted as "Procedural" (Factor 1), "Teacher Actions" (Factor 2), and "Process Skills" (Factor 3). Factor loadings ranged from 0.41 to 0.73.

Table 2. Principal component analysis pattern matrix (N = 201).

ITEM	Component		
	FACTOR 1 (Procedural)	FACTOR 2 (Teacher Actions)	FACTOR 3 (Process Skills)
LAB11	0.73		
INT10	0.63		
LAB30	0.61		
LAB8	0.60		
INT20	0.60		
INT29	0.52		
LAB9	0.48		-0.42
LAB22	0.44		
LAB34	0.41		
LAB31		0.61	
LAB24		0.61	
INT16		0.59	
LAB21		0.56	
LAB27		0.55	
LAB17		0.49	
LAB15		0.44	0.41
LAB36			0.68
LAB5			0.55
LAB37			0.49
TEXT28			0.46

Confirmatory Factor Analysis (Phase-Three)

Final application of CFA on the reduced measurement model accounted for changes in latent construct dimensions and re-configuration of items according to previous analysis. Figure 2 illustrates the improvement of the reduced measurement model (χ^2 [132, N = 201] = 236.93, p = 0.00, GFI = 0.88, CFI = 0.83, RMSEA = 0.06). When compared to the original model (see Figure 1) examined during phase-one of analysis, this slightly improved fitting is nevertheless unacceptable because of its statistical significance.

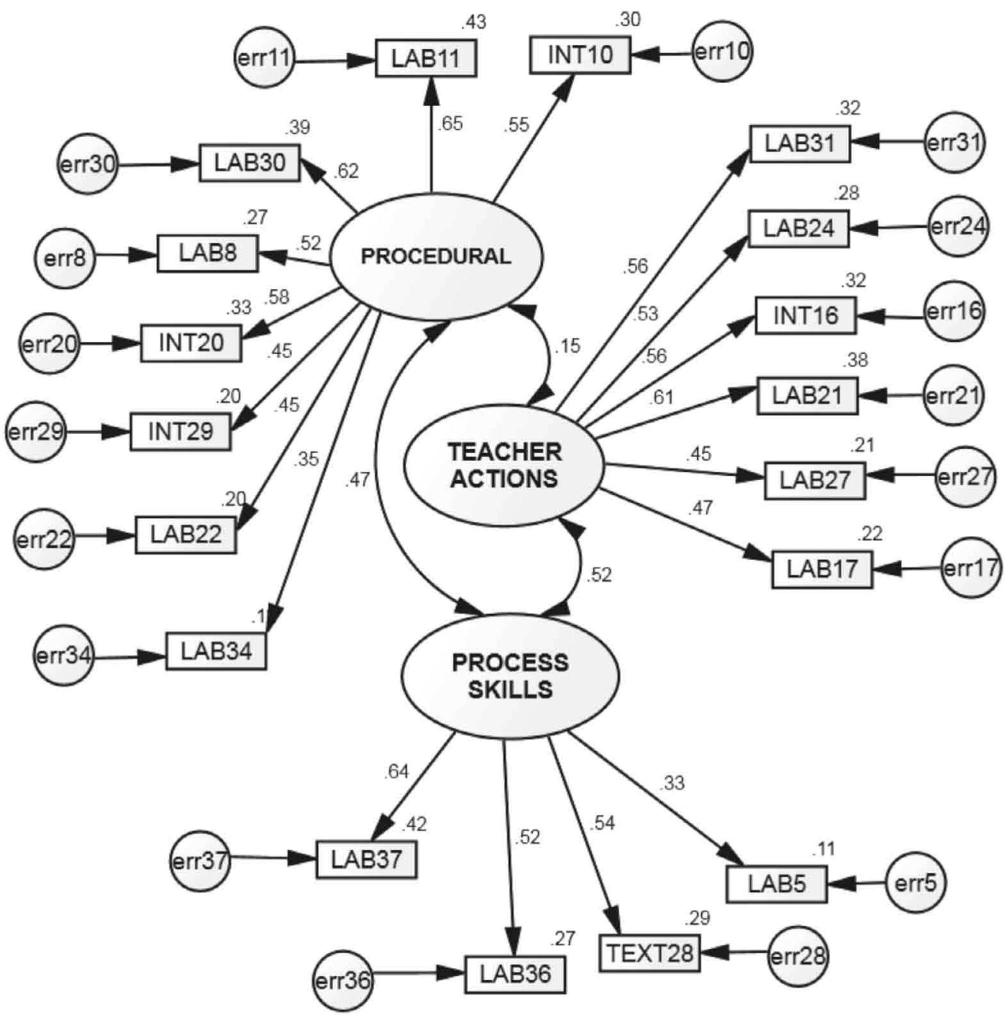

Figure 2: Reconstructed ISTS measurement model (Confirmatory Factor Analysis).

Alpha coefficient estimates were calculated on the entire sample for each reconstructed factor. The "Procedural," "Teacher Actions," and "Process Skills" factors had alphas of 0.74, 0.70 and 0.58, respectively.

Discussion

A three-phase data analysis progression including CFA and PCA was used to re-evaluate the psychometric properties of the Inquiry Science Teaching Strategies (ISTS) instrument. Based on a convenience sample of 201 elementary pre-service teachers in the upper Midwest of the United States, CFA analysis failed to confirm the original 3-factor structure of the ISTS. Instead, results indicated a marginally fitting 3-factor solution with three revised factors: "Procedural" (8 items); "Teacher Actions" (6 items); and "Process Skills" (4 items). Alpha coefficients for these factors were mediocre to weak, being 0.74, 0.70 and 0.58, respectively. PCA reduced the number of ISTS items from 40 to 18. Subsequent results of this study must be interpreted cautiously considering the possibility of potential limitations. First, a convenience sample of elementary pre-service teachers from the United States was used. The extent to which the results from this sample can generalize to other populations is not entirely clear. It is not known how representa-

ISSN 1648–3898

tive the pre-service teachers from the sample derived from this study compare to other populations of pre-service teachers in teacher education programs globally. Those interested in making comparisons can refer to the demographic information that was provided regarding the population from which the current sample was drawn. Furthermore, and based on the notion that a wide range of attitude instruments in science education measure similar constructs, there appears to be an abundance of studies relying on validation methods of limited value – namely content and face validity (Liu, 2009). Therefore, it is conjectured that other researchers, both in the United States and abroad, would encounter similar psychometric outcomes with the ISTS regardless of the demographic.

Second, it should be noted that prior to computer-generated calculation, the process involved in renaming the original three constructs was based on the most judicious human decision-making possible. For this reason, the constructs listed above should be interpreted as they might pertain to future efforts in revising factors or supplementing a new set of potential items based on results found in this study. Although computer-generated calculation (PCA) should never completely override or take the place of human judgment in construct development (Field, 2005), it is altogether possible that some of the items loaded on the revised constructs can be viewed as ambiguous or ill-suited. This appears to be an issue only with the "Procedural" construct. In this case, items loaded on this factor deserve further appraisal, if not a complete rewrite. This caveat should be taken into consideration in future instrument development utilizing the revised ISTS.

Results found as part of this study do not support psychometric authorization of using the initial ISTS for measuring "Classroom Teacher-Student Interaction," "Textbooks Used" or "Laboratory Activities." Data appears to support this claim, being that the instrument's factor structure is not in line with statistical outcomes suggestive of sophisticated methods involved with affective instrument development (Gable & Wolf, 1993), psychometric development guided by Structural Equation Modeling through confirmatory or exploratory methods (Blunch, 2008; Byrne, 2001), or studies aiming to re-evaluate the psychometric properties of instruments (Blalock et al., 2007; Lichtenstein et al., 2008; Smolleck et al., 2006).

This study ascertained some noteworthy construct dimensions. After removing a total of 22 items from the original ISTS, items loaded on the "Teacher Actions" and "Process Skills" factors appear, to some extent, represent respective constructs. For example, items under the "Teacher Actions" factor relate to areas that can be directly influenced by the teacher-facilitated inquiry experience. This requires a wide range of teacher directives dependent on the level of inquiry. It is recommended that additional items be added to reflect explicit teacher directions as well as more recent definitions of inquiry pertaining to the steps that teachers take in preparing for and implementing inquiry-based science teaching. Under the "Process Skills" factor, additional items should be added to more comprehensively reflect the range of science process skills inherent of inquiry-based science. Ultimately, determining the extent to which the 3-dimensions identified in this study contain useful evaluations of pre-service teacher attitudes towards inquiry will require further item revision utilizing appropriate international perspectives of what best constitutes positive affinity towards inquiry and the processes thereof (Abd-El-Khalick et al., 2004), additional instrument items developed according to researcher-solicited input from a sample of respondents, as well as further investigations with different samples derived from various national and international science teacher education program populations (Koballa & Glynn, 2007).

Conclusion

Overall, the three-phase data analysis progression used to re-evaluate and ultimately reconstruct the 40-item ISTS adhered to the recommended rigors of instrument development – attention to reliability and validity. Despite the shortfall in achieving intended, statistically non-significant outcomes, the psychometric properties were nonetheless re-examined under the direction of Confirmatory Factor Analysis, Principal Component Analysis, and general analysis of factor loadings. Hence, the restoration of the ISTS offers future researchers and science teacher educators a point of entry in working with a shortened attitudinal measure more pragmatic for the conditions of the modern day as well as a pathway in developing an affective instrument suitable for international audiences. Despite issues associated with instruments that are not entirely psychometrically reliable, the reconstructed ISTS does show promise because of a better resulting fit following item reduction. Further instrument modification in the form of adding items or making significant wording changes would require the reestablishment of reliability and validity with different samples.

References

Abd-El-Khalick, F., Boujaoude, S., Duschl, R., Lederman, N. G., Mamlok-Naaman, R., Hofstein, A., et al., (2004). Inquiry in science education: International Perspectives. *Science Education, 88*, 397-419.

American Association for the Advancement of Science (AAAS). (1993). *Benchmarks for science literacy*. New York: Oxford University Press.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238-246.

Blalock, C. L., Pruski, L. A., Toepperwein, M. A., Marshall, C. E., Owen, S.V., & Lichtenstein, M. J. (2007). In pursuit of validity: A comprehensive review of science attitude instruments 1935–2005. *International Journal of Science Education*, *30* (7), 961-977.

Blunch, N. J. (2008). *Introduction to structural equation modeling: Using SPSS and AMOS*. London: Sage Publications Ltd.

Bowen, G. M., & Bencze, L. (2009). Engaging preservice secondary science teachers with inquiry activities. In K. Tobin & W.-M. Roth (Eds.), *The World of Science Education: Handbook of Research in North America* (Vol. 1, pp. 587-609). Rotterdam, The Netherlands: Sense Publishers.

Brown, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In Bollen, K. A. & Long, J. S. (Eds.), *In Testing Structural Models* (pp. 136-162). Newbury Park, CA: Sage.

Byrne, B. M. (2001). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum Associates.

Chiappetta, E. L., & Koballa, T. R. (2006). *Science instruction in the middle and secondary schools: Developing fundamental knowledge and skills for teaching*. Upper Saddle River: NJ: Pearson Education, Inc.

Costenson, K., & Lawson, A. E. (1986). Why isn't inquiry used in more classrooms? *The American Biology Teacher, 48* (3), 150-158.

Council of Ministers of Education, Canada. (1997). *Common framework of science learning outcomes K to 12: Pan-canadian protocol for collaboration on school curriculum*. Toronto.

Damnjanovic, A. (1999). Attitudes toward inquiry-based teaching: Differences between pre-service and in-service teachers. *School Science and Mathematics, 99* (2), 71-76.

DeCoster, J. (1998). *Overview of Factor Analysis.* Retrieved February 6, 2011 from http://www.stat-help.com/notes.html.

European Commission. (2004). *Europe needs More Scientists: Report by the High Level Group on Increasing Human Resources for Science and Technology*. Brussels. European Commission.

Field, A. (2005). *Discovering statistics using SPSS* London: Sage.

Gable, R. K., & Wolf, M. B. (1993). *Instrument development in the affective domain: Measuring attitudes and values in corporate and school settings* (2 ed.). Boston: MA: Kluwer Academic Publishers.

Jöreskog, K. G., & Sörbom, D. (1988). *LISREL 7. A guide to the program and applications*. Chicago: SPSS Inc.

Koballa, T. J., & Glynn, S. M. (2007). Attitudinal and motivational constructs in science learning. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of Research on Science Education* (pp. 75-102). New York: Routledge.

Lazarowitz, R., Barufaldi, P. J., & Huntsberger, P. J. (1978). Student teachers' characteristics and favorable attitudes toward inquiry. *Journal of Research in Science Teaching, 15* (6), 559-566.

Lazarowitz, R., & Lee, A. E. (1976). Measuring inquiry attitudes of secondary science teachers. *Journal of Research in Science Teaching, 13* (5), 455-460.

Lichtenstein, M. J., Owen, S. V., Blalock, C. L., Liu, Y., Ramirez, K. A., Pruski, L. A., et al. (2008). Psychometric reevaluation of the scientific attitude inventory-revised (SAI-II). *Journal of Research in Science Teaching, 45* (5), 600-616.

Liu, X. (2009). Standardized measurement instruments in science education. In K. Tobin, & W. M. Roth (Eds.), *The World of Science Education: Handbook of Research in North America* (Vol. 1, pp. 649-676). Rotterdam, The Netherlands: Sense Publishers.

Lunetta V. N., Hofstein A., & Clough M. (2007). Learning and teaching in the school science laboratory: an analysis of research, theory, and practice, In N, Lederman, & S. Abel (Eds,), *Handbook of research on science education* (pp. 393-441), Mahwah, NJ: Lawrence Erlbaum.

National Center for Educational Research and Development. (1997). Manahej al-ta'alim al-a'am wa ahdafaha [Public educational curricula and goals]. Beirut, Lebanon.

National Research Council. (1996). National science education standards. Washington, DC: National Academy Press.

Osborne, J., & Dillon, J. (2008). *Science education in Europe: Critical reflections*: Nuffield Foundation.

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education, 25* (9), 1049-1079.

ISSN 1648–3898

Shaw, M. E., & Wright, L. R. (1967). *Scale for the measurement of attitudes*. New York: McGraw-Hill.
Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*, 4-14.
Smolleck, L. D., Zembal-Saul, C., & Yoder, E. P. (2006). The development and validation of an Instrument to measure preservice teachers' self-efficacy in regard to the teaching of science as inquiry. *Journal of Science Teacher Education, 17* (2), 137-163.
Windschitl, M. (2002). Inquiry projects in science teacher education: What can investigative experience reveal about teacher thinking and eventual classroom practice? *Science Teacher Education, 87*, 112-143.

Received: *December 23, 2011* Accepted: *February 15, 2012*

Sebastian Szyjka	PhD., Assistant Professor of Science Education, Western Illinois University, 1 University Circle, Horrabin Hall 69, Macomb, Illinois 61455, USA. E-mail: sp-szyjka@wiu.edu Website: http://www.wiu.edu/coehs/curriculum_and_instruction/faculty_and_staff/szyjka.php
Frackson Mumba	PhD., Associate Professor of Science Education, Southern Illinois University Carbondale, 625 Wham, Mail code 4610, Carbondale, Illinois 62901, USA. E-mail: frackson@siu.edu Website: http://ci.siu.edu/faculty/Frackson_Mumba/