Eurasia Journal of Mathematics, Science & Technology Education, 2016, **12**(3), 549-558

Construct Validity and Reliability Measures of Scores from the Science Teachers' Pedagogical Discontentment (STPD) Scale

Murat Kahveci & Ajda Kahveci
Canakkale Onsekiz Mart University, TURKEY
King Saud University, SAUDI ARABIA
Nasser Mansour
Exeter University, UNITED KINGDOM
King Saud University, SAUDI ARABIA
Tanta University, EGYPT
Maher Mohammed Alarfaj
King Faisal University, SAUDI ARABIA
King Saud University, SAUDI ARABIA

•*Received 26 December 2014*•*Revised 21 April 2015* •*Accepted 24 April 2015*

The Science Teachers' Pedagogical Discontentment (STPD) scale has formerly been developed in the United States and used since 2006. Based on the perceptions of selected teachers, the scale is deeply rooted in the cultural and national standards. Given these limitations, the measurement integrity of its scores has not yet been conclusively established internationally, such as in the Saudi Arabia context. The items of the scale are slightly tailored to make the instrument suitable in the specific context, such as with respect to country-based regulations, reforms, and everyday practices of science teachers and their professional development initiatives. Item-based descriptive statistics, the measure's factor structure as opposed to its former validity studies, and factor-based reliability scores are investigated in the present report. Thus, this study extends the validity and reliability measures of the instrument to the international scale and further confirms its appropriateness to measure teacher attitudes towards inquiry-based science education initiatives.

Keywords: science teachers, pedagogical discontentment, construct validity, factor analysis, cross-cultural validation

INTRODUCTION

Pedagogical discontentment is viewed as teachers' assessment of contextual aspects including working conditions and other external factors such as

Correspondence: Murat Kahveci,
Faculty of Education, Çanakkale Onsekiz Mart University, E4-427, Merkez 17100, Çanakkale, Turkey.
E-mail: mkk8653@my.fsu.edu
doi: 10.12973/eurasia.2015.1417a

ISSN: 1305-8223

standardized testing, limited lesson hours to cover content, facilities and paperwork. The construct means a state of cognitive conflict that exists when an individual recognizes a mismatch between science teacher's pedagogical goals and classroom practices. For a detailed discussion about pedagogical discontentment—its connection to the conceptual change models, its difference from contextual and job (dis)satisfaction, its interaction with teacher self-efficacy, and its meaning on teachers' receptivity to reform—, please see the relatively recent work of Southerland, S. A., Sowell, S., Blanchard, M., and Granger, E. M. (2011).

The quantification of the construct pedagogical discontentment was achieved via means of developing the STPD instrument by a group of researchers (Southerland et al., 2012). For the development and evaluation of the instrument, the researchers worked with practicing science teachers from all over the U.S. The initial stage of the instrument development involved selecting a purposeful sample of 18 teachers across the country and conducting a series of semi-structured interviews. These teachers were selected to represent diverse teaching situations (i.e., grade level, science discipline) and personal characteristics (i.e., teaching experience, gender). The teaching experience of these teachers ranged from 1-5 years to 15 or more years. They taught at elementary (grades K-5), middle (grades 6-8) or high school (grades 9-12) levels, and had elementary education, secondary education or science doctorate degree. Based on the interviews the researchers constructed five categories of discontentment, which represented the teachers' common experiences. Based on the five themes the researchers created 42 Likert-scale items and sought expert opinion from 10 classroom teachers and six science educators for content validity. In addition, five science education graduate students who were also classroom science teachers provided feedback.

While ensuring the construct validity, the researchers recruited 171 teachers from 12 states around the U.S. Similar to the initial sample, the teachers in the second sample represented a variety of teaching and personal characteristics in terms of gender, grades taught, age and teaching experience (Southerland et al., 2012). Based on the data collected, the construct validity of the scale was evaluated via factor analysis. The final STPD instrument consisted of 21 items in six subscales. According to these subscales pedagogical discontentment is classified into six categories: implementing inquiry instruction (IB) (four items), ability to teach all students science (AL) (four items), science content knowledge (SC) (four items), balance depth versus breadth of instruction (DB) (three items), teaching nature of science (TN) (three items), and assessing science learning (AP) (three items). During the evaluation process the Cronbach's alpha reliability coefficient was computed as .93 for the entire instrument. The reliability coefficients computed for each of the subscales are given in Table 1. For each of the items in the subscales respondents

State of the literature

- Pedagogical discontentment means a state of cognitive conflict that teachers realize a mismatch between pedagogical goals and actual classroom practices.
- Pedagogical discontentment is attributed to teachers' receptivity to reform, while it differs from contextual and job satisfaction.
- It is psychometrically possible to measure pedagogical discontentment however any attempt is contaminated with cultural, social, and political boundaries. Further purification for another context such as Saudi Arabia should be taken care of.

Contribution of this paper to the literature

- This study provides psychometric evidence for the STPD instrument to be used in the context of Saudi Arabia. Thus, Arabic speaking countries and similar educational contexts around that region will be able to utilize this instrument to reliably identify their teachers' openness to reform in science education—if any reform act is planned ahead.
- Measurement integrity of the STPD instrument actually turns out to be much different structure than in the US. Thus, policy-makers in the Saudi Arabia context need to understand the differences.
- This kind of instrumentation is often needed in science education in such cases as more in-depth demographics information is needed—especially when the focus is on educational reforms in science teaching.

Table 1. Reliability coefficients of the subscales in the original STPD

Subscales	Number of Items	Cronbach's Alpha
Inquiry Instruction (IB)	4	.87
Ability to Teach All Students Science (AL)	4	.82
Science Content Knowledge (SC)	4	.77
Balance Depth versus Breadth of Instruction (DB)	3	.89
Teaching Nature of Science (TN)	3	.85
Assessing Science Learning (AP)	3	.80

are required to select between options of 'no / slight / moderate / significant / very high discontentment'.

However, as pointed out by the developers, the STPD scale has a strength, on the other hand this strength holds a limitation about its use in all contexts, those of which may include international use.

> A strength of our instrument is the use of the qualitative data to derive the subscales and wording for the items. Based on the interviews and focus groups, the items in the STPD scale were constructed in ways that are familiar and consistent with classroom science teachers' experiences. However, as we pointed out, affective perceptions can be highly personalized; therefore, our instrument may not attend to all aspects of discontentment for all teachers. (Southerland et al., 2012, p. 491)

Therefore, it becomes evidently clearer that the construct validity of the STPD scale must be checked before its use in the Saudi Arabian context at large. On the other hand, Qablan et al. (2010) used an earlier version of the STPD scale in Jordan after translating it to the Arabic language, however in this study they did not report the construct validity of the instrument.

Conducting a factor analysis of the observed scores on a given instrument, one can determine if the test is measuring the variables it purports to, which is the definition of construct validation (Stapleton, 1997). Factor analysis is known as the heart of the measurement of psychological constructs (Nunnally, 1978).

The purpose of the present study is to explore the measurement integrity of the scores on the Science Teachers' Pedagogical Discontentment (STPD) scale in the Saudi Arabia context. Specifically, there were two research questions to address. First, what structure underlies responses to the measure—that is, does the (Arabic) structure correspond to that observed by Southerland et al. (2012) (in the US context)? Second, are scores on the STPD scale as a whole and as per each factor category reliable well enough to be used for future research?

METHOD

This is a survey research designed to investigate science teachers' pedagogical discontentment in Saudi Arabia (Jaeger, 1988). Specifically, this paper deals with only the construct validity of the Arabic STPD instrument.

Cross-cultural validation

The study addresses the procedures by following rigorous steps suggested by Sperber, Devellis, and Boehlecke (1994) for the process of cross-cultural validation. The researchers emphasize the challenge of adapting an instrument in a culturally relevant and sensible form and yet maintaining the original meaning. Thus, they argue that cross-cultural validation should be planned meticulously in advance. Attending to Sperber et al.'s (1994) cautions and suggestions, the original instruments were first translated into Arabic by a bilingual science educator fluent in both English and Arabic languages (Table 2). Then, the Arabic versions of the

Table 2. Translation and adaptation procedures of the instrument (Sperber et al., 1994)

Procedures	Notes for Clarity
Back translation (two bilingual science educators).	- Translate the original instruments from English to Arabic (person 1) - Translate the translated instruments from Arabic to English (person 2) (person 2 should not see the original English questionnaires)
Comparison of the two English versions (original and back-translated) (four native English speakers preferably in science education).	- Comparability of language: the formal similarity of words, phrases, and sentences - Similarity of interpretability: the degree to which the two versions would engender the same attitude response even if the wording were not the same Likert scales used
Consulting the translators for re-translation in case there are items with moderately or less similar interpretability and/or comparability of language.	- Re-comparison by native English speakers
Revision of some items to respond to the context and reforms in Saudi Arabia.	
Evaluation of the instruments by in-service science teachers in terms of language, clarity, meaning and suitability to the Saudi science teachers' population (seven in-service science teachers).	

instruments were back-translated to English by independent science educators in Saudi Arabia also fluent in English. The Saudi science educators did not see the original English instruments.

Following the translations, the original and the back-translated versions of the instruments were compared and evaluated in terms of form (language) and meaning. Comparability of language was regarded as the formal similarity of words, phrases, sentences, and similarity of interpretability was considered as the extent to which the two versions would invoke the same attitude response even if the wording were slightly different. For this purpose, an evaluation form (see Appendix) on both dimensions (language and meaning) with a 7-point Likert-type scale with options ranging from 'extremely comparable' to 'not at all comparable' was prepared for the original and back-translated item pairs. If the items were judged as extremely comparable they were scored 1.

The comparability of the items in terms of language and meaning was performed by a total of four native English speakers holding doctoral degrees in educational sciences. Scores given by the four experts were averaged and a threshold of 3 was used in deciding whether or not retranslation was needed. Items with an average score of 3 or above for both language and meaning were the ones that needed a careful reconsideration. Items that had a score of 3 or higher for only language were slightly modified to improve their language similarity while keeping the meaning intact. Based on the comparability evaluation, in the STPD scale four items were retranslated for both meaning and language, and four items were revised for only language.

In addition to the translation, back-translation and comparability work, some of the items were revised and clarified to respond to the Saudi Arabia context. Saudi Arabia has its unique culture and there are many current reform movements concerning science education. As these reforms and teachers' responses to them would be different than any other culture, it was important that the questionnaire reflects the context of Saudi Arabia and the nature of the reforms. An example of such a revision is changing the STPD item "balancing personal science teaching goals with those of state and national standards" to "balancing personal science teaching goals with those requirements, standards, goals set out by the Ministry of Education" as the only entity to set educational standards in Saudi Arabia was the Ministry of Education. One item (balancing personal science teaching goals with state/national testing requirements) was omitted from the STPD scale because the Ministry of Education has not released any testing benchmarks or standards.

In the final stage of development, the STPD instrument that eventually included 20 items was evaluated by seven Saudi science teachers specializing in physics, chemistry, biology, and geology, in terms of suitability for the local population of teachers. Their feedback was sought to identify any ambiguities or difficulties.

Subjects

The study was conducted in Saudi Arabia context with a total of 994 science teachers (656 females and 338 males) who were teaching physics (N=90, 9.1%), chemistry (N=93, 9.4%), biology (N=56, 5.6%) at secondary school level, and general science (N=682, 68.6%) at elementary and middle school levels. 923 schools (92.9%) were located in urban area while the rest of the schools were in suburban or rural areas. In regards to teaching experience, 36.6% participants fall into their first 5 years, 26.2% up to 10 years, 15.5% up to 15 years, 13.4% up to 20 years, and 9.3% 25 or more years. Majority of teachers (57.3%) have participated more than one professıonal development programs while the rest has no professional development experience at all.

The teachers were presented with the STPD instrument and asked to choose one of the Likert-scale options for each item. An informed consent letter was attached, outlining the nature of the research and ascertaining the confidentiality of individual responses.

RESULTS

Following data collection, the ratings were entered into SPSS (2012) for the analyses of descriptive and inferential statistics.

Descriptive results

As given in Table 3, all of the items were evaluated in terms of their goodness of fit for the normality condition. Scores were calculated in SPSS's list-wise selection, resulting in 982 valid cases. Thus, 12 cases were omitted from further analyses at this stage allowing one to employ advanced parametric test. The main indicator in this stage was the identification of cases, those standard scores of which exceed +/-3 boundary condition. To interpret the mean values, please note that 1 denotes "very high discontentment" while 5 denotes "no discontentment."

Factor analytic results

As the STPD instrument is highly sensitive to personal perceptions and depends on educational systems and reforms (implying cultural differences), both Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) are considered to validate the underlying contracts (Kline, 2011; Pett, Lackey, & Sullivan, 2003; Swisher, Beckstead, & Bebeau, 2004; Thompson, 2004).

Step 1. Exploratory factor analysis (Split-Half)

The main data was randomly split in haft, yielding new set of data with a sample size N= 471 for EFA. The KMO index yielding .974 > .500 suggests that the correlation matrix is not an identity matrix and sampling is suitable to run the factor analysis (see Table 4). In addition, the Bartlett's Test of Sphericity (p=.000<.050) confirms that there is a high correlation among the items, and thus, factor analysis should reliably discern the patterns over the dataset (Kaiser, 1974).

Table 5 shows the factor structure of the data. The model suggests that there is only one factor component satisfying Kaiser's criteria (eigenvalue > 1) (Kaiser, 1960) as explained by the scores of STPD. The variance of the model yields 71%.

This is an acceptable result for the explained variance in the humanities (Pett et al., 2003).

Principle component analysis (PCA) with varimax rotation produced Table 6, indicating the items associated with every factors component. As there exists only one component, no rotation could be employed. PCA suggested for establishing preliminary solutions in EFA (Pett et al., 2003). To retain an item in a component,

Table 3. STPD item-based descriptive statistics

	N	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Item 1: Teaching science to students of a lower academic achievement.	982	2.42	.639	.639	.078	-.753	.156
Item 2: Teaching science to students of a higher academic achievement	982	2.04	1.151	1.151	.078	-.307	.156
Item 3: Balancing personal science teaching goals with those requirements, standards, goals set out by the Ministry of Education.	982	2.21	.888	.888	.078	-.515	.156
Item 4: Monitoring student understanding through alternative forms of assessment such as quizzes, oral questions, presentations,..	982	2.20	.815	.815	.078	-.537	.156
Item 5: Including all ability-levels during inquiry-based teaching and learning (manifest cognitive and experimental operations to reach a scientific conclusion)	982	2.02	1.158	1.158	.078	-.260	.156
Item 6: Orchestrating a balance between the needs of both high and low ability level students	982	2.17	.909	.909	.078	-.412	.156
Item 7: Preparing students to assume new roles as learners within inquiry-based learning	982	2.18	.927	.927	.078	-.439	.156
Item 8: Using inquiry-based teaching within all content areas	982	2.20	.868	.868	.078	-.504	.156
Item 9: Assessing students' understandings from inquiry-based learning	982	2.19	1.024	1.024	.078	-.187	.156
Item 10: Assessing students' nature of science understandings	982	2.10	1.072	1.072	.078	-.211	.156
Item 11: Teaching science to students from economically disadvantaged backgrounds	982	2.09	1.019	1.019	.078	-.449	.156
Item 12: Planning and using alternative methods of assessment	982	2.10	1.036	1.036	.078	-.355	.156
Item 13: Having sufficient science content knowledge to generate lessons	982	2.23	.877	.877	.078	-.432	.156
Item 14: Teaching science subject matter that is unfamiliar to me	982	2.36	.769	.769	.078	-.444	.156
Item 15: Integrating nature of science throughout the curriculum	982	2.05	1.081	1.081	.078	-.150	.156
Item 16: Having sufficient science content knowledge to facilitate classroom discussions	982	2.22	.832	.832	.078	-.538	.156
Item 17: Using assessment practices to modify science teaching methods.	982	2.17	.876	.876	.078	-.539	.156
Item 18: Developing strategies to teach nature of science	982	2.16	.868	.868	.078	-.637	.156
Item 19: Ability to plan successful inquiry-based activities/learning	982	2.28	.753	.753	.078	-.589	.156
Item 20: Balancing the depth versus breadth of science content being taught	982	2.26	.780	.780	.078	-.758	.156
Valid N (listwise)	982						

Table 4. KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.974
Bartlett's Test of Sphericity	Approx. Chi-Square	10752.549
	df	190
	Sig.	.000

Table 5. Factor analysis (total variance explained)

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	14.134	70.671	70.671	14.134	70.671	70.671
2	.873	4.364	75.035			
3	.695	3.474	78.510			
4	.558	2.791	81.301			
5	.474	2.370	83.671			
6	.427	2.135	85.806			
7	.361	1.803	87.608			
8	.301	1.505	89.113			
9	.269	1.344	90.457			
10	.245	1.225	91.683			
11	.219	1.095	92.777			
12	.214	1.071	93.848			
13	.203	1.014	94.863			
14	.191	.953	95.816			
15	.180	.902	96.718			
16	.171	.857	97.575			
17	.147	.737	98.312			
18	.131	.653	98.965			
19	.106	.532	99.498			
20	.100	.502	100.000			

Extraction Method: Principal Component Analysis.

Table 6. Component matrix

	Component 1
Item 1: Teaching science to students of a lower academic achievement.	.575
Item 2: Teaching science to students of a higher academic achievement	.865
Item 3: Balancing personal science teaching goals with those requirements, standards, goals set out by the Ministry of Education.	.789
Item 4: Monitoring student understanding through alternative forms of assessment such as quizzes, oral questions, presentations,..	.874
Item 5: Including all ability-levels during inquiry-based teaching and learning (manifest cognitive and experimental operations to reach a scientific conclusion)	.905
Item 6: Orchestrating a balance between the needs of both high and low ability level students	.880
Item 7: Preparing students to assume new roles as learners within inquiry-based learning	.882
Item 8: Using inquiry-based teaching within all content areas	.833
Item 9: Assessing students' understandings from inquiry-based learning	.869
Item 10: Assessing students' nature of science understandings	.916
Item 11: Teaching science to students from economically disadvantaged backgrounds	.769
Item 12: Planning and using alternative methods of assessment	.910
Item 13: Having sufficient science content knowledge to generate lessons	.848
Item 14: Teaching science subject matter that is unfamiliar to me	.684
Item 15: Integrating nature of science throughout the curriculum	.884
Item 16: Having sufficient science content knowledge to facilitate classroom discussions	.875
Item 17: Using assessment practices to modify science teaching methods.	.863
Item 18: Developing strategies to teach nature of science	.846
Item 19: Ability to plan successful inquiry-based activities/learning	.850
Item 20: Balancing the depth versus breadth of science content being taught	.822

Extraction Method: Principal Component Analysis.

the following conditions were preferred: (1) the factor loadings must be higher than .400 and (2) a component contains at least three items (variables) loaded (Henson & Roberts, 2006) —assuming the condition (1) is met (doublets are omitted in this process). In this case all 20 items loaded higher than .575, which implies a well-defined one-factor structure for the data overall, named STPD, in short. Cronbach alpha reliability statistics (.978) shows that the Arabic STPD is highly reliable (Table 7).

Table 7. Component theme and its Cronbach's alpha reliability

Component	Theme	Number of Items	Cronbach's Alpha
1	Science teachers' pedagogical discontentment (STPD)	20	.978

Table 8. CFA Model fit indices (Split-half)

Measure	Threshold*	Model Fit Indices	Implication
Chi-square/df (cmin/df)	<3 good; <5 permissible	3.198	Satisfactory
p-value for the model	> .05	.000	Not satisfactory (High sample size)
CFI	> .95 great; .90 traditional; > .80 permissible	.981	Satisfactory
GFI	> .95	.942	Satisfactory
AGFI	> .80	.908	Satisfactory
RMSEA	< .05 good; .05 - .10 moderate; >.10 bad	.066	Moderate satisfactory
PCLOSE	> .05	.003	Not satisfactory

** The threshold values are based on Hu and Bentler (1999)*

Step 2. Confirmatory factor analysis (Split-Half)

Confirmatory Factor Analysis (CFA) is run over the second half of the data (N=511) derived from the Arabic STPD instrument. To employ the CFA analysis, IBM SPSS Amos 21 (Amos Development Corporation, Meadville, PA, USA) is used (Arbuckle, 2012). The CFA was run with the following analysis properties: (a) Discrepancy: Maximum likelihood, (b) Fit measures with incomplete data: Fit the saturated and independence model, and (c) Output: Standardized estimates, Residual moments, Modification indices. The threshold values on Table 8 are based on Hu and Bentler (1999), which indicates the goodness of fit indices. As chi-square per degree of freedom is in the acceptable region (<5), and similarly CFI, GFI, AGFI, and RMSEA values are under the threshold values, one would conclude that the suggested one-factor model (See Figure 1) is a "good fit." Thus, Arabic STPD instrument is a one-dimensional instrument measuring one construct, named STPD. With respect to treatments to suggested modification indices, many of the highly loaded ones are applied because there is only one latent variable and observed variables predicting common latent variables could share error variances.

This could be considered as a unique case because all 20 items are predicting one latent variable. In order to maximize the model fit, the standardized residual covariance matrix is carefully investigated (See Figure 2). Items 1, 4, 14, 17 and 18 were removed from the model as they loaded very high. Normally, concurrent occurrences of an item >.4 is considered as a source of discrepancy between the proposed and estimated models. The Cronbach alpha reliability coefficient (.975) of this model confirms highly reliable instrument.

DISCUSSION AND CONCLUSIONS

The results of EFA and CFA results are overall promising concerning the validity of the scores from the STPD scale. Having only one component, namely "Science teachers' pedagogical discontentment (STPD)" indicates that the teachers in Saudi Arabia perceive their pedagogical discontentment as one-dimensional mismatch with their pedagogical goals and teaching practice. In the original instrument (Southerland et al., 2012), there were six distinct dimensions in the measure. The potential reason for one-dimensional result could be that Arabic science teachers might have less experience in professional development towards alternative teaching methods such as inquiry-based science education as opposed to their American counterparts. This reason might reveal broader understandings about reforms and their effects in Saudi Arabia context. This evidence is alignment with the current debate in Saudi Arabia, indeed. Although there has been a shift from traditional teaching to inquiry-based science education, Almazroa and Alorini, as cited by Almazroa (2013) claim that professional development activities that are offered to teachers do not meet the demand of new curriculum.

In conclusion, the results of this study confirm that the Arabic version of the STPD scale is valid and highly reliable. It can be used in this specific context as a one-dimensional affective state to measure teachers' pedagogical discontentment towards teaching science.

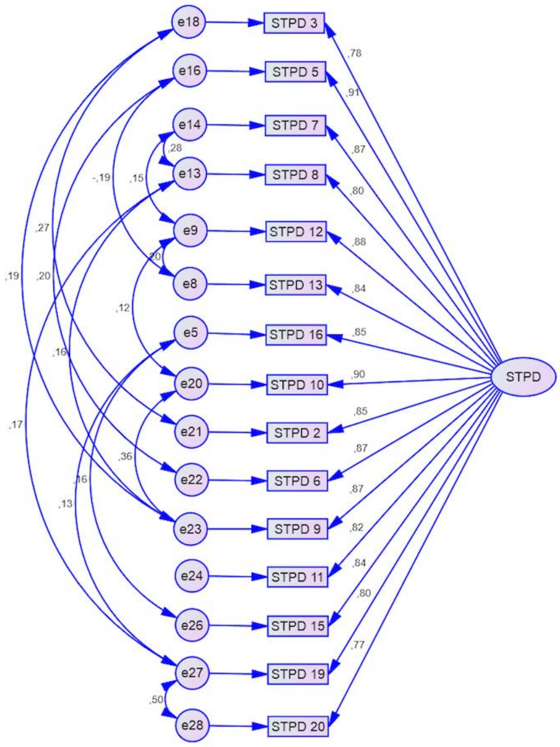

Figure 1. Path diagram for the CFA analysis of the overall Arabic STPD data. (Step 2)

	X20	X19	X15	X11	X9	X6	X2	X10	X3	X5	X7	X8	X12	X13	X16
X20	,000														
X19	,155	,149													
X15	,601	,641	,000												
X11	-,699	-,561	,072	,000											
X9	-,519	,231	-,346	,863	-,012										
X6	,032	-,168	-,531	-,244	-,277	,000									
X2	,334	-,278	-,105	,530	-,371	,607	,000								
X10	-,445	,014	,259	,851	,054	-,446	-,320	,034							
X3	-,006	,205	-,323	,450	-,066	,383	-,058	,152	-,033						
X5	-,365	-,562	,089	-,174	-,285	,026	,542	-,289	,376	,008					
X7	,279	,106	-,660	-,600	,022	,663	,062	-,144	-,146	,568	-,026				
X8	,792	,481	-,602	-,532	-,085	,679	-,350	-,286	-,529	,612	-,036	,035			
X12	,135	-,004	,290	,184	,480	,155	-,432	,198	-,410	-,199	-,150	-,384	-,003		
X13	,430	,521	,383	-,285	,354	-,155	-,758	,324	-,586	-,086	-,323	-,363	,064	,028	
X16	,207	,218	,062	-,371	,095	-,153	-,109	,152	-,043	-,248	-,169	,281	,004	,532	,017

Figure 2. Standardized residual covariances. (Step 2)

ACKNOWLEDGEMENT

This research was conducted as part of the professional development for in-service science and mathematics teachers research group with support of the Excellence Research Centre of Science and Mathematics Education - King Saud University, Kingdom of Saudi Arabia.

REFERENCES

Almazroa, H. (2013). Professional development: A vision for Saudi science teachers. *Esera.org*. Retrieved April 9, 2015, from http://www.esera.org/media/esera2013/Hiya_Almazroa_13%20Feb2014.pdf

Arbuckle, J. L. (2012). *IBM SPSS Amos 21 user's guide* (pp. 1–680). AMOS Development Corporation.

Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological Measurement*, *66*(3), 393–416.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.

Jaeger, R. M. (1988). Survey research methods in education. In R. M. Jaeger, *Compementary methods for research in education* (pp. 303–330). Washington, DC: American Educational Research Association.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*, 141–151.

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

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

Nunnally, J. (1978). *Psychometric teory* (2nd ed.). New York: McGraw-Hill.

Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. California: Sage Publications Inc.

Qablan, A. M., Khasawneh, S. A., & Al-Omari, A. A. (2010). Measuring the level of pedagogical discontentment among Jordanian science and vocational student teachers. *The Educational Journal*, *2*(95), 1–24.

Southerland, S. A., Nadelson, L., Sowell, S., Saka, Y., Kahveci, M., & Granger, E. M. (2012). Measuring One Aspect of Teachers' Affective States: Development of the Science Teachers' Pedagogical Discontentment Scale. *School Science and Mathematics*, *112*(8), 483–494. doi:10.1111/j.1949-8594.2012.00168.x

Southerland, S. A., Sowell, S., Blanchard, M., & Granger, E. M. (2011). Exploring the construct of pedagogical discontentment: A tool to understand science teachers' openness to reform. *Research in Science Education*, *41*(3), 299–317.

Sperber, A. D., Devellis, R. F., & Boehlecke, B. (1994). Cross-cultural translation methodology and validation. *Journal of Cross-Cultural Psychology*, *25*(4), 501–524. doi:10.1177/0022022194254006

SPSS. (2012). SPSS 19 software. Chicago, IL: SPSS Inc.

Stapleton, C. D. (1997). Basic concepts in exploratory factor analysis (EFA) as a tool to evaluate score validity: A right-brained approach (pp. 1–8). Presented at the Southwest Educational Research Association, Austin. Retrieved from http://ericae.net/ft/tamu/Efa.htm

Swisher, L. L., Beckstead, J. W., & Bebeau, M. J. (2004). Factor analysis as a tool for survey analysis using a professional role orientation inventory as an example. *Physical Therapy*, *84*(9), 784–799.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: American Psychological Association

◈◇◈