Running head: SELF-EFFICACY AND CONTENT KNOWLEDGE

Self-Efficacy and Content Knowledge of students in Sequence and Series

Amit Savkar

Mathematics

196 Auditorium Road

U 3009, Storrs, CT 06269

Abstract

The purpose of this article is to describe the process of designing an new instrument to use self-efficacy and content knowledge of students on sequence and series to predict their scores on midterm exams. The instrument will have three constructs i.e, self-efficacy of students on the topic of sequence and series, content knowledge on sequence and content knowledge on series. A content validation report is generated using 8 experts in the field of Mathematics and Education. The instrument will be surveyed by approximated 300 to 350 students.The target population is first year university students taking Calculus II class in their second semester at the University of Connecticut. The response to the survey will be collected using iClicker (personal response system). The data will be analyzed using exploratory factor analysis (EFA) and the results from factor and reliability study will be discusses.

Self-Efficacy and Content Knowledge of students in Sequence and Series

Introduction

Students in the traditional Calculus II class find the topic of sequence and series more difficult than the integral calculus. Over the last five years, we have seen a steady decline in the scores of students on the midterm exam, which tests students on this topic. It has also been observed that the students have difficulty on the same material during the final exam. Although it is a common norm for professors in general to accept that the students do poorly in their exams particularly on the topic of sequence and series, there has been no effort made to examine the reasons behind such poor performance. The goal of this instrument is to look at correlation between the self-efficacy of the students on the topic of sequence and series, and their actual content knowledge. The instrument will analyze three basic constructs. The first construct will look at the student self-efficacy about sequence and series. The second construct is designed to understand the student content knowledge on sequences and the third construct is to understand the content knowledge on series. The research questions (RQ) that guide this study are:

1. Is there a correlation between the self-efficacy of students on the topic of sequence and series and their content knowledge on sequence?

2. Is there a correlation between the self-efficacy of students on the topic of sequence and series and their content knowledge on series?

3. What is the correlation between the three constructs and the students performance on the midterm exam?

The data was collected two days before their second midterm exam. This data was collected using iClicker.

Background

Social cognitive theory suggests that self-efficacy, peoples judgments of their capabilities to organize and execute course of action required to attain designated types of performance (Bandura, 1986, p.391), strongly influences the choice people make, the effort they expend and how long they persevere in the face of challenge. Bandura (1986) had hypothesized that self-efficacy beliefs mediate the effect of other determinants of performance such as gender and prior experience on subsequent performance. Bundura had argued that constructs such as self-concept, perceived usefulness, and anxiety are common mechanisms of personal agency in the sense that they, like self-efficacy beliefs, also influence the outcome. Pajeres and Miller conducted a study on the role of self-efficacy and self-concept beliefs in mathematical problem solving by using path analysis (Pajares and Miller 1994). The purpose of that study was to test Bundaras (1986) hypothesis regarding the predictive and meditational role of self-efficacy in the area of mathematics. One of the important findings of this study was that students judgments about their capability to solve math problems were more predictive of their ability to solve those problems that were other variables. Another study exploring the mathematics self-efficacy and mathematical performance correspondence (Hackett and Betz 1989) found that mathematics performance and mathematics self-efficacy were significantly and positively correlated with attitudes towards mathematics. The study also found that most of the men and minority women overestimated rather than underestimated their performance capabilities. This tendency on the part of all students towards inaccurate self-estimates and performance suggests that the student performance is highly correlated to their self-estimates of how well they would do in the mathematics course. It is also interesting to note that the student did worse than what they thought on how well they would perform. As such there is a vast literature on research studies that have looked at the affect and mathematics learning over few decades (Mcleod 1994). As such there are several factors such as self-efficacy (Betz & Hackett 1983), math attitude and math achievement (Randhawa, Beamer & Lundberg 1993), math self-concept and math performance (Marsh, Walker & Debus 1991), that have been studied over the years to understand student performance and see if these prove to be good predictors of student performance in mathematics course. Recent studies have looked into the role of self-efficacy and motivation in mathematics performance in high school students (Stevens et.al 2004). The study has found that self-efficacy is important because it not only predicts students mathematics performance but also motivational variables that influence aspects of overall mathematics achievement. The goal of the current research is to find the self-efficacy of students take Calculus II at the University of Connecticut on the topic of sequence and series. Over the last several years that the author has taught this topic, it has been observed that students perform poorly during the exam, which tests their knowledge about sequence and series. The current study has three primary constructs of interest: a) students self-efficacy about sequence and series, b) students content knowledge about sequence, and c) students content knowledge about series. Unfortunately there is no evidence in the literature citing any such study that has been conducted in the past. Literature does show, that the study of students self-efficacy can be used as good predictor for student performance in mathematics. This is one of the motivations to use self-efficacy as of the constructs in the current study. Due to lack of literature on specific instrument testing the students self-efficacy of students in the topic of sequence and series the author has designed a new instrument to study these constructs. Experts in the field of mathematics and education have validated this instrument. It is hypothesized by the author of this study that a strong correlation will exist between the students self-efficacy and their knowledge on sequence and series. For the purpose of establishing this correlation the author intends to look at the students performance on midterm 2. The details of the study are explained in the next section.

Content Validation

Generating item pool

The instrument is designed to understand the self-efficacy and content knowledge of the students on the topic of sequence and series in Calculus II. Several ideas that are conceptual and require mathematical rigor were considered. Effort was made to separate and consolidate questions that were intended to analyze the self-efficacy of the students. In this particular subject area, it is extremely challenging to distinguish between a self-efficacy (perception) question and a question that tests the knowledge of students. It is important to note that the perception of students is strongly influenced by the knowledge they have about the content. It is also important to note that the self-efficacy questions did not distinguish between self-efficacy of students on sequence and series separately. The questions designed for content knowledge were separated based on knowledge of sequence and series. A total of forty two questions were created in the pool of items. Face validity was used to check if these items fell in their described construct. A Likert scale for agreement was used for items on self-efficacy. The items on content knowledge were scaled based on 'True' 'False' and 'Don't know' options.

Content Validity

As with design of any new instrument, one must validate the instrument before using it to gather data. The process of content validation was initiated by carefully choosing questions which this author thought to be best represent each of the three factors in consideration. A total of 42 questions were prepared. The self-efficacy questions were scaled 1 through 5 using Likert scale, while the content knowledge questions were scaled as True, False and Don't know. The experts were provided a formal definition of each of the categories. They were asked to rate these items based on three things i.e., to Indicate the category that each of the item best fitted by writing a numeral (I through III and letter 'O' for others), to indicate the certainty for their placement of the item in a category by writing a number (1 = Not very sure, 2 = pretty sure and 3 =very sure) and lastly the experts were asked to indicate how relevant they felt each item was to be in the category by rating it (L = Low/No relevance, M = Moderate relevant, H = Highly relevant). The content validation was sent to the 15 experts via email. A table 1 shows the names and conceptual definition of the hypothesized factors is .

Out of the 15 possible responses the author received 8 reponses. A summary report on the content validation is attached as appendix C. Each item was scrutinized first on its placement in the correct factor, then, on how certain the experts were on its placement and last criteria for accepting or rejecting the item was its relevance to the construct. The denominators for certainty and the relevance were adjusted based on the correct placement of each items in the appropriate factors. The summary report has corrected values in terms of the percentages by adjusting the denominator. Based on this analysis, items, 3,4,8, 14-16, 18-21, 23, 29, 31-36, 41, 42 were deleted from the final survey. A copy of final survey is attached as appendix D and the order in which the survey questions were asked to the target population is attached in appendix E.

Data collection and data mining

The final survey with questions was presented to the target population as questions through a power point presentation. The responses to each of these questions were collected using clickers (personal response system). The survey was conducted in large lectures of Calculus II. Two sections were presented with the questions and a total of 319 individual responses to the questions were recorded. The scale used for collecting response on content knowledge questions was dichotomous, the recorded response to those questions had to be recoded. Data on the self-efficacy questions was analyzed separately from content knowledge to determine the number of factors.

Factor Analysis

Self-Efficacy

Principal Axis Factoring. The data was subjected to PCA using Eigen values greater than 1 as the criteria for extracting factors. There were a total of 7 questions which were associated with this factor. The PCA predicted that there were two factors. Looking at the correlations matrix question 3 and question 8 were negatively correlated and their loading on the two factors was also low. On further investigation of the original questions it was found that the questions were actually "trick" questions. This made response by the students to these question uninterpretable. These questions were deleted from further analysis.

Principal axis factoring was used to determine the factors from the remaining 5 questions on self-efficacy. The PAF predicted one factor. The factor matrix in figure 1 shows the loading on the factors for each of the remaining five questions. It is clear that none of the loading are strong. Question 2 has a loading of 0.56, which may be considered good and Question 7 has a loading of 0.448 which is fair. All other loadings are poor. Looking at the correlation matrix, one can see that inter item correlations in figure 2 are at best poor. This is possible as each of the self-efficacy questions addresses different ideas with sequence and series. The total proportion in variance in the data explained by this one factor is 16.99 %. A parallel analysis using principal component and random normal data generated, suggests that the data obtained for the study follows a strong trend of being random as seen in figure 3. This suggests that the instrument has failed to truly capture the self-efficacy of students in sequence and series.

Content Knowledge

Principal Axis Factoring. As described in the data mining section, all the responses to the true or false questions were recoded to incorporate the dichotomous nature of the scale. A preliminary analysis of the data using principal axis factoring resulted in extracting 6 factors. There were only two hypothesized factors in that the instrument that should have been captured. Correlation matrix suggested removing questions 11, 12, 14, and 21. These had extremely low and negative correlations with other items. Amongst these questions, questions 11, 14 and 21 asked the students about their knowledge on partial sums. This topic within the broad topic of sequence and series is particularly difficult for students. The questions were worded to test student understanding of partial sum as related to sequence and series. As it turns out, the questions in their original form, may have been unclear to the students. After deleting these questions PAF was used to determine the factors on the remaining 10 items in the data. As seen from the factor matrix in figure 4, four factors were extracted. The factor loading of each of the items is below 0.35 with exception of one item i.e. question 10. The inter item correlations as seen in the figure 5 are extremely low. The total proportion of variance explained by these four factors together is 20.78 %. It is author's suspicion that this could be attributed to the broad nature of the knowledge on sequence and series. A parallel analysis using principal component and random normal data generated, suggests that the data obtained for the study could have two factors as seen from figure 6. The parallel analysis suggests that there could be two factors. Further investigation of the data using the oblique rotation (Direct Oblimin) show puzzling results. The pattern matrix show in figure 7 shows poor loading of each of the item on the factors. Due to the fact that the magnitude of loading of items after controlling for other extracted factors (partial standardized coefficients) is low one cannot ascertain if there are truly four factors that are being extracted. The zeroth order correlation as seen from the structure matrix in figure 8 also suggests week correlation amongst the factors and items.

Reliability Analysis

Self-efficacy

The check for internal consistency reliability which is a function of the average inter-item correlation and the number of items given by Cronbach's alpha was found to be 47.9%. Since this represents the proportion of variance in a given scale that can be attributed to a common source, one can say in the present case, more than 50% of the variation in the data is due to error. This would conclude that the instrument designed to measure self-efficacy of students in the topic of sequence and series is highly unreliable.

Content knowledge

Reliability analysis on the data for extracting the factors for content knowledge showed a Cronbach's alpha of 0.427. This means that the instrument designed to predict the content knowledge of the students on sequence and series is only 42.7% reliable.

Conclusion

Based on the factor analysis of data on self-efficacy and content knowledge of students in sequence and series, it can be concluded that the instrument in its current form does not clearly measure the constructs that the author had hypothesized. There are several issues that need to be addressed starting with narrowing the area of interest within the broad topic of sequence and series. It is also possible that the self-efficacy part of the instrument measured the self-efficacy of students on topic of sequence and series separately and not as one single topic. The inter-item correlations for this factor were extremely low indicating that each item was probably testing a different idea. It is important to consider writing more items which, are testing the same idea within each construct. A similar inference can be made based on there factor analysis of content knowledge data. None of the factor loadings were more than 0.6. Most of the item loadings were under 0.3. The reliability analysis also confirmed the need to have better inter-item correlations and a need to increase the number of items in each of the constructs. Thus the current instrument is not sufficient in predicting the scores of students in midterm exams. As such a preliminary regression analysis was performed to check if there was any correlation between individual items from the instrument to questions on the exam and it was found to have at best a weak correlation.

Future work

To design an instrument that effectively measure the self-efficacy and content knowledge of the students in sequence and series should be done in an incremental fashion. It is necessary to first concentrate one of the content ideas for example sequence. It is required that the questions which are formulated to address the self-efficacy should be clear and should have easily interpretable responses. These will allow the researcher a better insight in the measurement of self-efficacy. Same can be said about designing an instrument that measures the content knowledge of students. It is possible to separate the content knowledge questions that pertain to sequence and series even though there is conceptual overlap between the two.

References

Bandura, A. (1986). Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall.

Pajares, F., Miller, D. M. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. Journal of education psychology, 86, 2, 193-203.

Hackett, G. and Betz, N. E. (1989) An exploration of mathematics self-efficacy/mathematics performance correspondence. Journal for research in mathematics education, 20, 261-273

Mcleod, D. B. (1994) Research on affect and mathematics learning in JRME: 1970 to the present. Journal for research in mathematics education, 25, 637-647.

Betz, N. E. and Hackett, G. (1983). The relationship of mathematics self-efficacy expectations to the selection of science-based college majors. Journal of Vocational behavior, 23, 329-345

Randhawa, B. S., Beamer, J. E., & Lundberg, I. (1993). Role of mathematics self-efficacy in the structural model of mathematics achievement. Journal of Education Psychology, 85, 41-48.

Marsh, H W., Walker, R., & Debus R. (1991). Subject-specific components of academic self-concept and self-efficacy. Contemporary education psychology, 16, 331-345

Stevens, T., Olivarez., A. Jr., Lan, W. Y., & Tallent-Runnels, M. K. (2004) Role of mathematics self-efficacy and motivation in mathematics performance across ethnicity. The Journal of Educational Research, 97, 208-221

Appendix A

Self-Efficacy Factor analysis and reliability analysis output

Appendix B

Content Knowledge Factor analysis and reliability analysis output

Appendix C

Content Validation Summary Report

Appendix D

Final copy of Instrument

Appendix E

Copy of question order

Author Note

Thanks to Dr. Betsy McCoach, Dr Thomas Defranco and experts who validated the instrument

Table 1
Factors and description

Categories	Conceptual Definition
I . Sequence and Series Self-Efficacy	A person's beliefs about his/her ability to tackle problems in sequence and series
II. Content knowledge of sequences	The knowledge that each student has about sequences
II. Content knowledge of series	The knowledge that each student has about series

Figure Captions

Figure 1. Factor Matrix for self-efficacy of students in sequence and series.

Figure 2. Correlation coefficients on self-efficacy of students in sequence and series.

Figure 3. Parallel Analysis for self-efficacy.

Figure 4. Factor matrix for content knowledge using Principal axis factoring.

Figure 5. Inter-item correlation coefficients for items testing content knowledge.

Figure 6. Parallel Analysis for content knowledge.

Figure 7. Pattern matrix for extracting factors on content knowledge.

Figure 8. Structure matrix for extracting factors on content knowledge.

Factor Matrix[a]

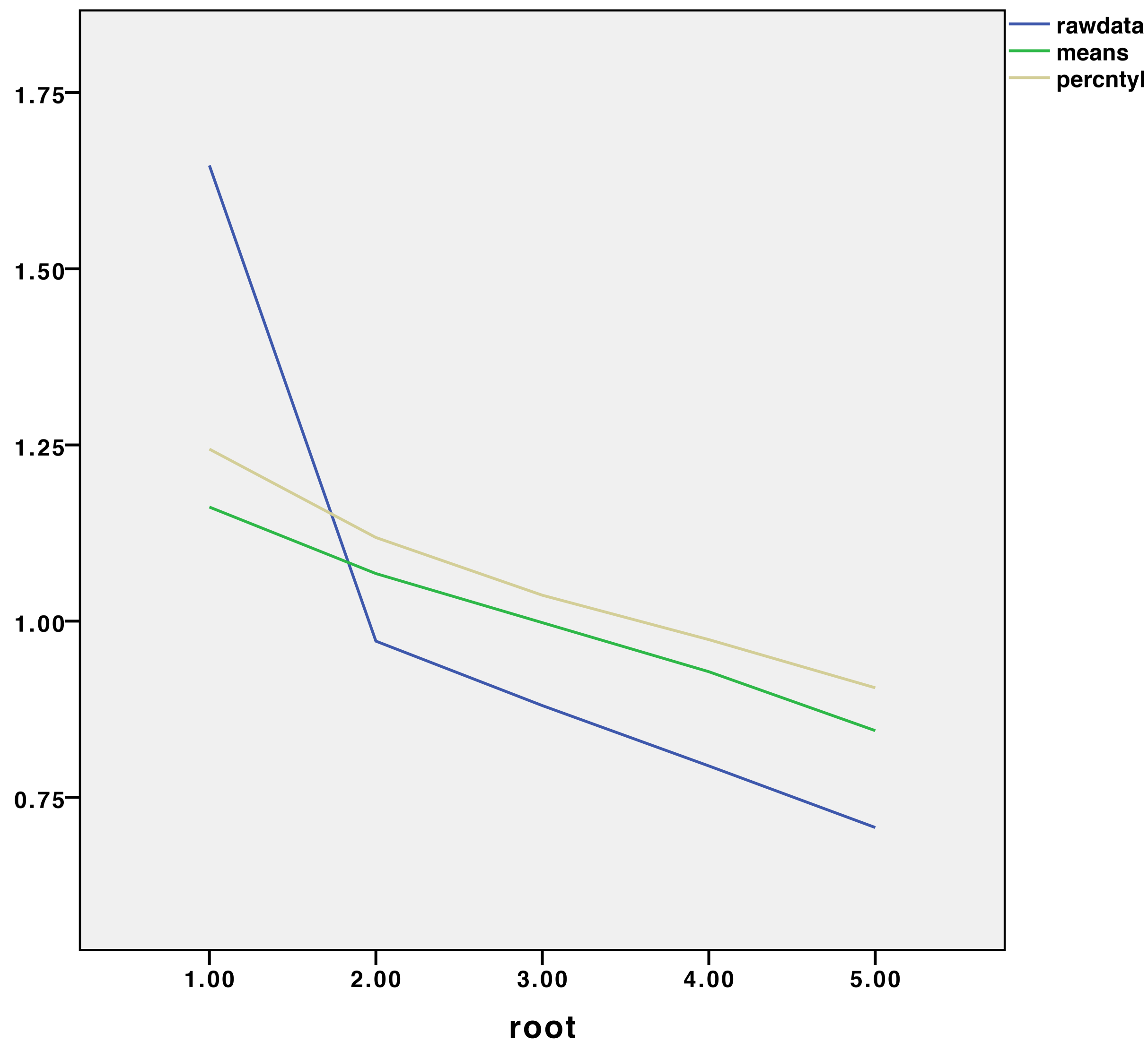
	Factor
	1
Q 1	.297
Q 2	.560
Q 4	.302
Q 7	.448
Q 18	.392

Extraction Method: Principal Axis Factoring.

a. 1 factors extracted. 10 iterations required.

Correlation Matrix

		Q 1	Q 2	Q 4	Q 7	Q 18
Correlation	Q 1	1.000	.153	.055	.148	.145
	Q 2	.153	1.000	.155	.242	.253
	Q 4	.055	.155	1.000	.196	.096
	Q 7	.148	.242	.196	1.000	.130
	Q 18	.145	.253	.096	.130	1.000


rawdata
means
percntyl
1.75
1.50
1.25
1.00
0.75
1.00
2.00
3.00
4.00
5.00
root

Factor Matrix[a]

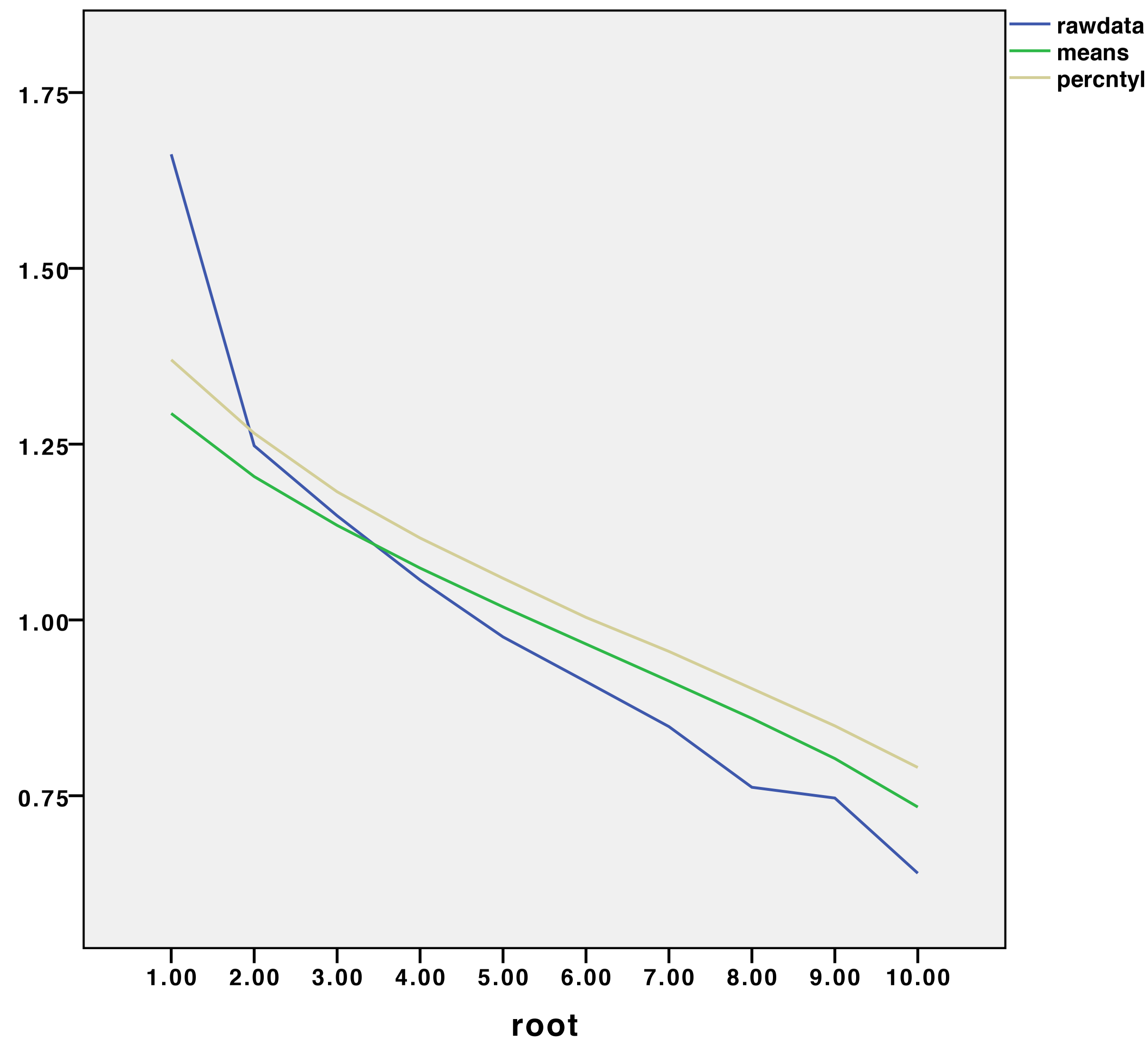
	Factor			
	1	2	3	4
Q 5	.213	.086	-.007	.127
Q 6	.333	.172	.164	-.217
Q 9	.309	-.455	.099	.128
Q 10	.474	-.313	.031	-.192
Q 15	.249	.130	-.276	.116
Q 16	.310	.283	-.120	-.217
Q 17	.254	.164	.139	.110
Q 19	.200	.120	.242	.091
Q 20	.139	.204	.217	.210
Q 13	.366	-.003	-.290	.162

Extraction Method: Principal Axis Factoring.

a. 4 factors extracted. 14 iterations required.

Inter-Item Correlation Matrix

	Q 5	Q 6	Q 9	Q 10	Q 13	Q 15	Q 16	Q 17	Q 19	Q 20
Q 5	1.000	.053	.051	.039	.081	.107	.064	.091	.099	.034
Q 6	.053	1.000	.019	.144	.037	.047	.185	.079	.142	.067
Q 9	.051	.019	1.000	.269	.113	.001	-.071	-.007	.042	.023
Q 10	.039	.144	.269	1.000	.127	.050	.097	.092	.035	-.045
Q 13	.081	.037	.113	.127	1.000	.193	.113	.105	.016	.003
Q 15	.107	.047	.001	.050	.193	1.000	.117	.004	.018	.045
Q 16	.064	.185	-.071	.097	.113	.117	1.000	.089	.021	.045
Q 17	.091	.079	-.007	.092	.105	.004	.089	1.000	.096	.158
Q 19	.099	.142	.042	.035	.016	.018	.021	.096	1.000	.106
Q 20	.034	.067	.023	-.045	.003	.045	.045	.158	.106	1.000



rawdata
means
percntyl
1.75
1.50
1.25
1.00
0.75
1.00
2.00
3.00
4.00
5.00
6.00
7.00
8.00
9.00
10.00
root

Pattern Matrix[a]

	Factor			
	1	2	3	4
Q 5	.165		-.165	
Q 6	.148			-.403
Q 9		-.558		.165
Q 10		-.528		-.243
Q 13		-.119	-.468	
Q 15			-.406	
Q 16			-.169	-.426
Q 17	.301			
Q 19	.330			
Q 20	.401			

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

a. Rotation converged in 10 iterations.

Structure Matrix

	Factor			
	1	2	3	4
Q 5	.202		-.202	
Q 6	.249	-.123		-.431
Q 9		-.552		
Q 10		-.548	-.155	-.284
Q 13		-.185	-.481	-.115
Q 15			-.407	-.130
Q 16	.119		-.259	-.454
Q 17	.334		-.143	-.168
Q 19	.338			-.131
Q 20	.381			

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