12-2017

Development and validation of an instrument to measure attitude of undergraduate students towards Statistics

Debrishi Bose
University of Connecticut, debrishi.bose@uconn.edu

Development and validation of an instrument to measure attitude of undergraduate students towards Statistics

Debrishi Bose

University of Connecticut

Spring 2017

Abstract

Owing to the recent increase in the number of Statistics classes across colleges, there has been a rise in fear for Statistics or Statistics' Anxiety among students, as a result of which attitude of students towards Statistics has received considerable attention from researchers. All the existing Instruments have overlooked the impact of Environmental Influence towards learning Statistics, which motivated the development of Attitude towards Statistics Instrument. An analysis of 107 responses (mostly female) revealed the hypothesized 4 factor structure of the Instrument.

Keywords: Statistics' Anxiety, Environmental Influence, Attitude towards Statistics

Development and validation of an instrument to measure attitude of undergraduate students towards statistics

Introduction

With the tremendous increase in data analysis in the past few decades, a good understanding of Statistics has become extremely crucial in most fields. (Mji, 2009, pg. 737). Consequently, at least one course in Statistics has become a core component in most undergraduate or graduate level degree programs in most universities (Hong et al., 2014, pg. 1177). In most undergraduate programs, students enroll into the program under the impression that they're supposed to study only topics related to their major without even realizing the connection or link between their majors and statistics. This becomes especially true for those students who enter undergraduate Psychology program. According to Hong.et al (2014), only 57.1 % of students were completely aware of the Statistics courses they needed to take while majoring in Psychology. Perhaps students entering such programs might not have a very strong mathematical background which in conjunction with lack of awareness create a negative perception towards Statistics courses. Furthermore, these mandatory Statistics courses are often considered high staked since they determine whether a student is qualified enough for entry into the Honors program or not, thereby spreading Statistics Anxiety among students. (Hong et al., 2014, pg. 1178).

Literature Review

Quantitative reasoning, statistical literacy, and interpretation of data have become an important components of education at school and college levels. The emphasis placed on quantitative research by funding agencies such as the Institute of Education Sciences and National Science Foundation has further contributed to an increased emphasis on statistical

training at the graduate level in the social and behavioral sciences and education. This emphasis on statistical training at the graduate and undergraduate levels in these fields has contributed to statistics anxiety among students (Hong et al., 2014, p. 1178) and has spawned research on statistics education, attitudes of students towards statistics, and statistical anxiety. Statistics Anxiety can be described "as the feelings of anxiety encountered when taking a Statistics course or doing statistical analyses" (Cruise, Cash and Bolton, 1985, pg. 92). From that definition, it seems pretty evident that such type of anxiety can be faced by both teachers and students while teaching a class or attending, studying for an exam or analyzing data.

Most of the previous studies conducted by researchers in this area has conceptualized Statistical Anxiety to be a six dimensional construct, namely - Worth of Statistics, Interpretation Anxiety, Test and Class Anxiety, Computational Self-Concept, Fear of asking for help and Fear of Statistics teachers. (Hsiao, 2010, pg. 977). According to Cruise et. al (1985), these six dimensions can further be subdivided into two major types- first factor comprising Interpretation Anxiety, Test and class Anxiety and Fear of asking for help, while the second factor comprising Worth of Statistics, Computation Self Concept and Fear of Statistics teachers. The latter factor has been referred to as the Attitudes of students towards Statistics. "Attitude towards Statistics can be described as a disposition to respond favorably or unfavorably to objects, situations or people related to statistics learning" (Arumugam, 2014). The empirical studies conducted by researchers has shown that the actual questionnaire developed for Statistics Anxiety has two subscales and that the two factor model provides better fit. According to Papousek et al. (2012), "The differentiation between anxiety and attitudes is suggested further by the two-part format of the questionnaire." The study conducted by them also confirmed the two-factor structure of the STARS-Anxiety scale (Papousek et al., 2012, pg. 86).

To the best of my knowledge, the instruments or the scales that has been developed so far or used in studies looking at “Attitude of students towards Statistics” have overlooked their background, which I believe plays a significant role in this context. Agliata et al. (2008, pg. 967) highlights the fact that "many college goers not only consider their parents to be the authoritative figures who possesses the right to take decisions for them, but also try to meet their expectations and feel obligated to do so". Accordingly to Doren et al (2012, pg. 7),”numerous studies have found parental expectations not only influencing their children's skills or abilities, but also eventually impacts their educational or occupational choices”. Empirical studies have shown these factors to be influencing the school engagement and achievements as well. Furthermore, such influences have an impact on proximal factors like- beliefs, attitudes or behaviors (Doren et al., 2012, pg. 8). Besides parental influence, college students are often influenced by their peers while taking decisions. Cohen (1983) refers to several studies where peer influence effects are well documented. According to Cohen (1983), “High school students' close friends influence their decisions as to whether or not to attend college (e.g., Campbell and Alexander, 1965; Duncan et al., 1968; Kandel and Lesser, 1969; Hauser, 1972; Sewell and Hauser, 1972; Alexander and Eckland, 1975)". Additionally, “The size of the peer influence effect has been represented by the coefficient of the direct path from best friend's college plans to respondent's college plans; coefficients for this path have often exceeded .2, suggesting a fairly substantial effect (e.g., Hauser, 1972; Alexander and Eckland, 1975; Sewell and Hauser, 1975; Alexander et al., 1975; Hauser et al., 1976; and Alwin and Otto, 1977)." Cohen (1983). While Baloglu (2004) throws some light on how the environmental factors could contribute to Statistics Anxiety among students in general, Perepiczka et al. (2012, pg. 101) discusses the importance of social support and it’s “buffering effect” on graduate students’ Statistics Anxiety and Attitude

towards Statistics. Thus, it is quite evident that parental or peer influences play a significant part in decision making for many college students. Nolan et al. (2012, pg. 105-106) gives a table of the available instruments to study “Attitude towards Statistics”. However, none of the instruments (including the most recent one SASTSc) have any information on Environmental factors. According to Papousek et al. (2012, pg. 82), Statistics anxiety seems to be a common issue among students (more specifically among Psychology novices) across the globe, so we need to have an instrument which would also incorporate the environmental factors while measuring the attitude of students towards Statistics. Thus, I propose to develop an instrument which would have 3 factors – Usefulness, Self-Concept, Statistics Instructor and an additional factor “Environmental Influence”.

Usefulness of Statistics can be referred to "as an individual's perception of the relevance of Statistics". (Hong et al., 2014, pg. 1178). For example, if a student is barely going to use much Statistics in the future, he or she will have a high score on this construct (Williams, 2013, pg. 48). A student majoring in Philosophy or Music would probably not be doing much Statistical Analyses in future, so he or she might have a low score on this. On the other hand, individuals with high scores on this construct are those who thinks Statistics course(s) are important or highly relevant to his or her study or the things learned from the class would surely be useful. A student majoring in Business or Finance and seeing his or her future in Stock Marketing would probably have a high score.

Self-Concept can be described "as an individual's self-perception of his or her ability to understand and calculate Statistics" (Hong et al., 2014, pg. 1178). In other words, it refers to the problems or challenges faced by an individual while trying to understand and solve a statistical problem, regardless of their true ability. For example, if a person thinks that he or she doesn't

clearly understands the statistical concepts or the problems involving Statistics are tough, then he or she would have a low score on this construct (Williams, 2013, pg. 48). Likewise, a high score on this would imply that he or she understands the concepts well or doesn't really face challenges while solving Statistical problems.

The construct "Statistics Instructors" can be referred to as an individual's mental impression or way of regarding his or her statistics teacher. If a student thinks that his or her teacher speaks in a "completely different language" (Mji, 2009, pg. 738), then that individual is expected to have a low score on this construct. On the other hand, if an individual thinks that he or she is on the same page as his or her teacher during classes or meetings or regards the Instructor to be a nice person, then that person would have a high score on the construct.

The construct “Environmental Influence” in this regard can be described as how an individual perceives the environmental conditions (family or school) and reacts to it while deciding to study Statistics. Some people hold stereotypes against specific subjects so they try to influence their children by making suggestions or recommendations. A higher score on this construct would mean that higher influences by family or peers being encountered by the student, while a lower score indicates lower degrees of family or peer influence.

Numerous studies have been done in the past which aimed at capturing the relationship between Statistics Anxiety or Attitude towards Statistics and achievements in these courses. Researchers have reported adverse or negative effects of Attitude towards Statistics on course performance. (Onwuegbuzie et al., 2010; Williams, 2013; Mji, 2009). However, this study would be different from the past studies in terms of the additional dimension that we have in the instrument.

Content Validity Procedures

The content validity procedures in my case required sending the Instrument to content experts in this area and seeking their feedback. A minimum of five feedbacks was necessary. I sent out an email to about 15 Professors at the University of Connecticut and heard back from 7 Professors, who wished to serve as content validation experts of my Instrument. However, one of them was traveling outside US on a conference and wanted an extension (if possible) while the other wanted one of her graduate students to work on the content validation. Owing to time conflicts and meeting deadlines, I couldn't have the Professor (who was traveling) as my content validation expert but I agreed with the other Professor to have her student as my content validation expert. Thus, the content validation of my instrument "Attitude of undergraduates towards Statistics" was done by Dr. Aarti Bellara, Dr. D. Betsy McCoach, Dr. Hariharan Swaminathan, Dr. Jane Rogers, Dr. Eric Loken and Ms. Shu Chiang. The first five members are the faculty in the Measurement, Evaluation and Assessment Program (under the Dept. of Educational Psychology) at the University of Connecticut, while last member is a graduate student at the Springfield College in Massachusetts.

I had 4 factors in my Instrument namely- Usefulness of Statistics, Computational Self Concept, Statistics Instructor and Social Expectations. (Definitions of these factors can be found in the Content Validation form in the appendix). In addition to the 29 items the form had, it also contained three more columns (Factor/ Certainty/Relevance) corresponding to each item. At the end of the Instrument was a space provided for reviewers to comment in general about the factors and or items. Each content expert was required to fill out the three columns corresponding to each item (instructions to fill out was provided in the content validation form).

The recommendations or feedbacks I received on my instrument can be broadly classified into two categories- one involving the factor "social expectations" and other centering on the clarity of some items. Much critic about the factor comes in terms of the way it has been defined and whether the factor actually fits the situation where the survey would be conducted. Although the factor might be relevant to some cultures, however, the experts unanimously think that the factor should either be modified or dropped from the Instrument. Some of the comments involving the factor are as follows: "Is this even a major issue for most students?", "I don't think the family factor will work well in the US at least", "May be call it family expectations". Other experts believe that the items developed for this factor wouldn't make sense to most students and hence, such items would be of no use.

As far as the items in general are concerned, most of the experts think that a majority of the items are unclear or poorly worded. About 10-12 items were found to be irrelevant to any of the factors by almost all experts. For example, "Statistics require strong Math skills". Items that were marked irrelevant by most experts were dropped. Some items were not complete in some sense, for example: "Statistics is irrelevant for me". Two of the items (11&16) were identical, so one was dropped. Some of the items like "Most people think smart kids should take up Statistics" have been voted irrelevant or not an attitude question by all experts, so it was also dropped. The items were designed in such a way that few experts found it to be a very negative-type survey. In fact, after receiving the feedback, I myself think that the general tone is negative. Also, the items were not adequately representing the constructs. With these problems in items and factors, most experts think that the instrument needs considerable improvement.

The first issue which I addressed was changing the "Social Expectations" factor to

“Environmental Influence" and defined it. Since I wanted to turn the survey into a more "positive-type" tone, consequently, I had to drop all the items and rephrase them again. Once I created this set of items, I looked back again at the comments and tried to make the items clearer and specific. A revised draft of the instrument is provided at the end of this section. However, this revised draft has not gone through content validation.

Initial Draft of the Instrument

Proposed Items for "Usefulness of Statistics"

Usefulness of Statistics: *This factor concerns individuals' perceptions of the relevance of Statistics.*

The following items are on a 7 point Likert Scale: 1- Strongly Disagree, 2- Disagree, 3- Slightly Disagree, 4- Neutral, 5- Slightly Agree, 6- Agree, 7- Strongly Agree

ITEMS	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Statistics is irrelevant for me							
Statistics is not necessarily needed by all fields							
Statistical skills will not improve my job prospects							
Statistics is vague							
I don't enjoy taking Statistics classes							
Statistical arguments are confusing							
I barely use Statistics							

Proposed Items for "Self-Concept"

Self-Concept: *This factor measures the anxiety experienced while solving Statistical problems or questions.*

The following items are on a 7 point Likert Scale: 1- Strongly Disagree, 2- Disagree, 3- Slightly Disagree, 4- Neutral, 5- Slightly Agree, 6- Agree, 7- Strongly Agree

ITEMS	1	2	3	4	5	6	7
Statistical questions are difficult to understand							
I don't like to do Statistics assignments							
Statistics formulas are hard to memorize							
Statistics require strong Math skills							
It is difficult to succeed in a Statistics course							
Solving Statistics assignments takes time							
It is difficult to do Statistical computations by hand							
It is common to run into problems while doing Statistical calculations							

Proposed Items for "Statistics Instructors"

Fear of Statistics teachers: *This factor measures the anxiety associated with the Statistics Instructor*

The following items are on a 7 point Likert Scale: 1- Strongly Disagree, 2- Disagree, 3- Slightly Disagree, 4- Neutral, 5- Slightly Agree, 6- Agree, 7- Strongly Agree

ITEMS	1	2	3	4	5	6	7
Statistics Instructors don't explain concepts clearly							
Statistics seems "Greek to me" during lectures							
I rarely meet my Statistics Professors during their office hours							
Statistics Instructors are intimidating							
I don't feel comfortable asking questions during a Statistics class							
Statistics Instructors rarely smile in class							

Proposed Items for "Social Expectations"

Social Expectations: *This factor measures the problems of students associated with interference of their family in school matters when they want or don't want to study Statistics.*

The following items are on a 7 point Likert Scale: 1- Strongly Disagree, 2- Disagree, 3- Slightly Disagree, 4- Neutral, 5- Slightly Agree, 6- Agree, 7- Strongly Agree

ITEMS	1	2	3	4	5	6	7
I wouldn't like if my family members force me to take up Statistics							
I can't accept the logic behind taking up Statistics, just to keep the family tradition							
I don't like the "Statistical environment" in which I grew up							

Sample Description

All undergraduate students at the University of Connecticut (across all campuses) who have taken at least one course in Statistics in their undergraduate career or currently enrolled in a Statistics course were invited to take part in the survey. Out of 265 people who started the survey, 260 of them completed it. Among the 260 completed surveys, 183 of them took a Statistics course in their undergraduate career, while the other participants (77) didn't take any Statistics course, as a result of which their responses were not included in the Analyses. 15 responses out of the 183 came from Non-undergraduates (like Doctoral or graduate students, staff members or participants who didn't disclose their year of college), so they were excluded from the analysis as well. Now among these 168 responses, 61 participants didn't answer all the questions of the survey. While some answered most of the questions, others responded to one or

two questions, as a result of which these 61 responses were dropped from further analysis. The final sample (who answered all the questions of the survey) used for the analysis consisted of responses from 107 participants, out of which 71 were females and 36 males. The mean age was 20.41 years and standard deviations was 3.603 years.

Data Collection

As a part of data collection procedure, in order to reach the target participants, the survey was advertised multiple times on UConn Daily Digest, Student Daily digest (for undergraduates), undergraduate listservs for regional campuses (via soapbox.com) and Undergraduate Tutors listserv at the Quantitative Learning Center. Besides, I've contacted Program Assistant at the Math Dept. (Uconn) to post it on their listserv. Furthermore, the Professors teaching undergraduate Math and Statistics classes in their respective Depts. were requested to share the survey with their students. Lastly, the survey was advertised and shared on Facebook.

Factor Extraction

Initially, 4 factors were hypothesized which could explain the variability in the responses to the Attitude towards Statistics survey among the participants. The phase of preliminary EFA starts off with Principal Component Analysis (PCA) to understand and have an overall yet simple idea about the factor structure of the dataset. The results of PCA conducted in SPSS yields: KMO= 0.887, p-value corresponding to Bartlett's test of Sphericity= 0.00, Kaiser Criterion= 7 factors (accounting for 72% variability), Scree plot: 5 factors, Parallel Analysis: 4 factors, MAP test: 5 factors. Also, the communalities were sufficiently high.

Looking at all these criteria, it seems reasonable to start the next phase of analysis with

5 factors. At this stage, a Principal Axis Factoring (PAF) was being conducted with Oblique Rotation, which allows for the correlation of the factors. This is the beginning of next phase of analysis. Although we have hypothesized factors, but we are still looking for the best factor structure associated with this data, so such PAF techniques were conducted repetitively until a satisfactory factor structure was attained.

With regards to five factor PAF, the Factor Structure and Pattern matrices were carefully studied and analyzed. The coefficients of the pattern matrix are like standardized Regression coefficients of the factor on the items....A careful analysis of this revealed that about 15 items load well onto one of the factors while 17 of the items (Q5-2, Q5-4,..Q9-7) loads on to multiple factors. It was also noticed that the coefficients corresponding to the 5th factor were almost negligible, except for 2-3 items, so this is an indication that perhaps the last factor is not needed. Since none of the items, except two or three load on to it, it makes sense to drop this 5th factor and consider a 4 factor model for the data.

This was different from the last step only in terms of the number of factors we wanted to extract via PAF. KMO was found to be 0.891. Again, the factor structure and pattern matrices were carefully analyzed. It now revealed that three items (Q5-4, Q7-9, and Q9-5) to be multidimensional, as the coefficients in the pattern matrix corresponding to these items were about moderate and loaded on to multiple factors, so these were eliminated along with another item Q9-8 as it did not load on to any of the 4 factors. The deletion of those 4 items and fitting a 4 factor model caused the KMO go up slightly from 0.891 to 0.892. At this stage, a careful examination of the Pattern matrix resulted in the deletion of three more multidimensional items- Q5-3, Q8-3 and Q8-4. Now, the KMO dropped slightly to 0.891. After re-examination of the

matrices, there wasn't really any need to delete any items further. Thus, based on EFA, a four factor structure of the data was obtained.

Table 1.

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure.		0.891
Bartlett's Test of Sphericity	Approx. Chi-Square	2323.351
	df	406
	Sig.	0.000

Table 2.

Communalities

	Initial	Extraction
Q5_1	0.628	0.495
Q5_2	0.594	0.415
Q5_5	0.675	0.475
Q5_6	0.701	0.641
Q5_7	0.839	0.730
Q5_8	0.728	0.655
Q5_9	0.781	0.678
Q7_1	0.658	0.546
Q7_2	0.777	0.732
Q7_3	0.535	0.518
Q7_4	0.806	0.716
Q7_5	0.586	0.369
Q7_6	0.722	0.631
Q7_7	0.475	0.347
Q7_8	0.822	0.791
Q8_1	0.848	0.769
Q8_2	0.831	0.852
Q8_5	0.834	0.726
Q8_6	0.736	0.611
Q8_7	0.682	0.618
Q8_8	0.881	0.875
Q8_9	0.579	0.400
Q9_1	0.803	0.785
Q9_2	0.375	0.414
Q9_3	0.677	0.523

Q9_4	0.716	0.637
Q9_6	0.546	0.368
Q9_7	0.516	0.371
Q9_9	0.585	0.491

Extraction Method: Principal Axis Factoring.

Table 3: Pattern Matrix

Factor	1	2	3	4
Q5_1	0.636	0.145	-0.082	-0.007
Q5_2	0.176	0.565	0.029	-0.105
Q5_5	0.698	-0.032	-0.192	0.076
Q5_6	0.874	-0.151	-0.074	-0.024
Q5_7	0.714	0.177	-0.063	0.160
Q5_8	0.299	0.055	-0.065	0.680
Q5_9	0.743	0.152	-0.225	0.046
Q7_1	0.658	-0.017	0.254	-0.009
Q7_2	0.121	0.802	0.018	-0.041
Q7_3	0.326	-0.016	0.599	-0.128
Q7_4	0.645	0.094	0.164	0.270
Q7_5	0.562	-0.076	0.051	0.212
Q7_6	0.645	0.242	0.048	-0.029
Q7_7	-0.012	0.186	0.535	-0.025
Q7_8	-0.009	0.810	-0.018	0.189
Q8_1	-0.053	0.735	0.011	0.312
Q8_2	0.076	0.130	-0.022	0.849
Q8_5	0.713	0.182	0.073	0.065
Q8_6	0.617	0.122	0.245	0.040
Q8_7	0.652	0.120	0.236	-0.070
Q8_8	-0.029	0.848	-0.041	0.228
Q8_9	-0.155	0.118	0.514	0.291
Q9_1	0.015	0.101	0.149	0.811
Q9_2	-0.015	-0.034	0.651	-0.101
Q9_3	0.566	0.250	-0.003	0.014
Q9_4	-0.037	0.823	0.026	-0.037
Q9_6	0.041	-0.117	0.589	0.140
Q9_7	0.127	0.556	-0.004	-0.054

Q9_9	0.664	0.024	0.130	-0.186

Extraction Method: Principal Axis Factoring.
Rotation Method: Oblimin with Kaiser Normalization.[a]
a. Rotation converged in 12 iterations.

Table 4: Structure Matrix

Factor	1	2	3	4		
Q5_1	0.690	0.434	0.061	0.148		
Q5_2	0.435	0.616	0.151	0.135		
Q5_5	0.660	0.298	-0.060	0.167		
Q5_6	0.783	0.248	0.059	0.065	0.815	0.568
Q5_7	0.117	0.343	0.431	0.436	0.068	
Q5_8			0.746			
Q5_9	0.783	0.487	-0.055	0.208		
Q7_1	0.696	0.342	0.373	0.124	0.503	0.848
Q7_2			0.178	0.274		
Q7_3	0.407	0.200	0.644	-0.018		
Q7_4	0.768	0.533	0.327	0.433	0.572	0.281
Q7_5	0.164	0.287	0.765	0.550	0.208	
Q7_6	0.176	0.173	0.266	0.564	0.094	
Q7_7	0.411	0.872	0.142	0.481	0.358	0.826
Q7_8	0.162	0.573	0.282	0.474	0.100	
Q8_1	0.908	0.826	0.562	0.244	0.262	
Q8_2	0.728	0.478	0.385	0.216	0.742	0.450
Q8_5	0.372	0.111	0.411	0.910	0.126	
Q8_6	0.529	0.048	0.241	0.535	0.358	
Q8_7	0.232	0.431	0.250	0.865		
Q8_8	0.072	0.037	0.632	-0.051		
Q8_9	0.688	0.527	0.147	0.203	0.358	0.797
Q9_1	0.161	0.261	0.119	0.058	0.590	
Q9_2	0.163	0.385	0.597	0.113	0.171	
Q9_3	0.668	0.299	0.240	-0.049		
Q9_4						
Q9_6						
Q9_7						
Q9_9						

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

Table 5: Factor Correlation Matrix

Factor		1	2	3	4
1	1.000	0.482	0.186	0.174	
2	0.482	1.000		0.177	0.366
3	0.186	0.177		1.000	0.099
4	0.174	0.366		0.099	1.000

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

Reliability Analyses:

After conducting EFA, the hypothesized factors have been redefined, and can be found in the following section along with Reliability Analyses. In terms of Reliability Analyses, the approach would be to consider one sub-scale with its corresponding items at a time and perform the analysis and repeat it for all other scales.

Statistics' Interest and Future Use:

This sub-scale consisted of 14 items and the internal consistency Reliability of this is 0.942. This can be interpreted as 95% of the total variance in this sub-scale can be attributed to a common source. In other words, the correlation coefficient between the observed scores and true scores is about 0.97. From the table given below, it is clear that deletion of any items from this sub scale would reduce the internal consistency reliability, as a result of which all the 14 items were retained.

Table 7: Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Q5_1	59.60	286.073	0.659	0.541	0.939
Q5_5	59.49	291.063	0.610	0.630	0.941
Q5_6	59.54	286.383	0.718	0.645	0.938
Q5_7	60.72	272.015	0.814	0.776	0.935
Q5_9	59.81	285.172	0.752	0.744	0.937
Q7_1	60.79	286.416	0.677	0.557	0.939
Q7_4	60.77	274.596	0.777	0.720	0.936
Q7_5	59.39	301.599	0.566	0.477	0.942
Q7_6	59.97	284.594	0.761	0.653	0.937
Q8_5	60.30	273.155	0.837	0.801	0.934
Q8_6	60.78	277.194	0.727	0.651	0.938
Q8_7	60.83	279.953	0.732	0.622	0.937
Q9_3	60.31	282.423	0.692	0.595	0.939
Q9_9	60.04	293.131	0.634	0.483	0.940

This sub-scale measures the degree to which students take interest in learning Statistics and its relevance in their future career. So a person receiving a high score on this construct means that he/she takes interest or has an inclination towards learning Statistics and considers it to be relevant in terms of their future career or job prospects. On the other hand, a low score on this construct implies that the person is reluctant to learn Statistics and considers it to be irrelevant in terms of their future career or job prospects. The mean and standard deviation of this sub-scale are 4.63 and 1.30 respectively. The mean score can be interpreted as the average response (score) over all possible items in that sub-scale across all participants.

Table 8: Inter-Item Correlations

	Q5_1	Q5_5	Q5_6	Q5_7	Q5_9	Q7_1	Q7_4	Q7_5	Q7_6	Q8_5
Q5_1	1.000	0.538	0.626	0.549	0.625	0.524	0.443	0.398	0.473	0.531
Q5_5	0.479	0.687	0.415	0.394	0.502	0.503	0.417	0.626	0.679	1.000
Q5_6	0.527	0.421	0.570	0.539	0.549	0.479	0.586	1.000	0.695	0.473
Q5_7	0.824	0.625	0.687	0.635	0.695	1.000	0.434	0.585	0.415	0.580
Q5_9	0.523	0.473	0.434	1.000	0.532	0.429	0.568	0.566	0.443	0.394
Q7_1	0.532	1.000	0.439	0.616	0.758	0.398	0.502	0.421	0.420	0.415
Q7_4	0.623	0.517	0.473	0.503	0.570	0.641	0.580	0.568	0.616	0.623
Q7_5	0.531	0.417	0.539	0.824	0.614	0.566	0.758	0.517	0.717	1.000
Q7_6										
Q8_5										
Q8_6	0.458	0.357	0.463	0.647	0.470	0.575	0.755	0.407	0.559	0.710
Q8_7	0.490	0.357	0.524	0.631	0.447	0.615	0.638	0.455	0.605	0.716
Q9_3	0.469	0.447	0.464	0.674	0.701	0.471	0.558	0.351	0.564	0.632
Q9_9	0.424	0.391	0.549	0.496	0.479	0.570	0.536	0.281	0.457	0.533

Statistical Ability

This sub-scale had 7 items to begin with and the estimate of internal consistency coefficient was found to be 0.923. This means about 92 % of the variation in this sub-scale could be attributed to a common source. The tables for internal consistency reliability coefficients for "item-deleted" cases were generated which are discussed below:

Table 9: Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Q5_2	30.28	72.638	0.584	0.410	0.921
Q7_2	30.22	67.081	0.802	0.688	0.900
Q7_8	29.90	64.546	0.845	0.767	0.895
Q8_1	29.20	65.140	0.809	0.792	0.899

Q8_8	29.40	63.771	0.886	0.857	0.890
Q9_4	30.47	67.119	0.738	0.645	0.906
Q9_7	29.07	76.655	0.562	0.345	0.922

From the table presented above, it is clear that if the item Q9-7 is deleted, then the internal consistency reliability coefficient would increase to 0.922. Hence this item was removed from the sub scale. The following table suggests that further deletion of item Q5-2 would result in increasing reliability, hence it was deleted as well.

Table10: Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Q5_2	24.60	59.356	0.584	0.410	0.932
Q7_2	24.54	54.364	0.804	0.683	0.905
Q7_8	24.21	51.906	0.856	0.765	0.897
Q8_1	23.51	52.856	0.799	0.790	0.905
Q8_8	23.72	51.562	0.880	0.853	0.894
Q9_4	24.79	54.321	0.742	0.645	0.913

Now the table given below suggests that no further deletion of items from this sub scale is necessary, so other items were retained.

Table 11: Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Q7_2	20.07	40.232	0.784	0.643	0.923
Q7_8	19.74	37.516	0.872	0.765	0.906
Q8_1	19.04	38.433	0.807	0.789	0.919
Q8_8	19.24	37.525	0.879	0.845	0.905
Q9_4	20.31	39.461	0.760	0.640	0.928

This subscale measures an individuals' ability or self-confidence in understanding and solving Statistical problems. A high score on this sub scale implies that a student has no problem understanding statistical questions and manages to solve statistical problems by himself or herself. On the other hand, a low score on this means, a person faces problems understanding statistical questions and have difficulty in solving statistical by himself/ herself. The Mean score is given by 4.92 and standard deviation is given by 1.54. The mean score (4.92) represents the average score in this subscale over the 5 items across 107 participants.

Table 12: Inter-Item Correlations

	Q7_2	Q7_8	Q8_1	Q8_8	Q9_4
Q7_2	1.000	0.739	0.641	0.704	0.743
Q7_8	0.739	1.000	0.770	0.829	0.743
Q8_1	0.641	0.770	1.000	0.885	0.597
Q8_8	0.704	0.829	0.885	1.000	0.679
Q9_4	0.743	0.743	0.597	0.679	1.000

College Influence

This sub scale had 5 items with an internal consistency reliability of 0.738. This means about 74% of the variation in this sub-scale could be attributed to a common source. The table

presented suggests that elimination of any item would result in reduction of internal consistency reliability, so all the items were retained.

Table 13: Item-Total Correlation

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Q7_3	17.45	21.287	0.493	0.295	0.696
Q7_7	16.19	20.078	0.513	0.309	0.689
Q8_9	15.45	22.193	0.493	0.382	0.697
Q9_2	17.38	20.559	0.483	0.262	0.700
Q9_6	15.78	21.327	0.528	0.401	0.683

This factor measures the influence of Professors or peers on an individual in studying Statistics. A high score on this sub scale means higher degrees of influence on the person by Professors or friends for studying Statistics whereas a lower score indicates lower degrees of influence by Professors or friends on that person in studying Statistics. The mean score and standard deviation are given by 4.112 and 1.112 respectively. A mean score of 4.112 implies the average score in this sub-scale over the 5 items across 107 participants.

Table 14: Inter-Item Correlations

	Q7_3	Q7_7	Q8_9	Q9_2	Q9_6
Q7_3	1.000	0.469	0.236	0.408	0.289
Q7_7	0.469	1.000	0.373	0.348	0.287
Q8_9	0.236	0.373	1.000	0.258	0.579
Q9_2	0.408	0.348	0.258	1.000	0.387
Q9_6	0.289	0.287	0.579	0.387	1.000

Statistics Instructors

This factor had three items with an internal consistency reliability of 0.896. However, deletion of item Q5-8 from this sub scale would result in an increase in the reliability coefficient to 0.909.

Table 15: Item- Total statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Q5_8	10.85	8.751	0.737	0.554	0.909
Q8_2	10.31	9.291	0.854	0.748	0.808
Q9_1	10.36	8.948	0.807	0.707	0.841

This can be referred to as an individuals' mental impression or the way of regarding his or her Statistics Instructor. A higher score on this construct indicates the person regards his or her Statistics Instructor in a positive way while a lower score is an indicative of the person regarding his or her Statistics Instructor in a negative way. The mean score and standard deviation are given by 5.425 and 1.479 respectively. The average score on this sub-scale across 107 participants on the two items is 5.425. In this case, the correlation between the two items is 0.835.

Table 16: Comparison across sub-scales

Sub-Scales	Reliability	Number of Items	Average I.I.C	STD.Dev(I.I.C)
Interest/Future Use	0.942	14	0.537	0.105
Stat Ability	0.932	5	0.733	0.08
College Influence	0.738	5	0.363	0.1
Stat Instructor	0.909	2	0.835	0

Average Inter-Item Correlations can be defined as the average correlation among all the items in a given sub-scale.

Table 17: Means and SD by Sub-scales

Sub-Scales	Scale Means	Scale Std Dev
Interest/Future Use	4.63	1.3
Stat Ability	4.92	1.54
College Influence	4.112	1.112
Stat Instructor	5.425	1.479

Table 18: Subscale Correlations

	future mean	ability_mean	col_infl_mean	stat_inst_mean
future_mean	1	.561**	.292**	.356**
		0.000	0.002	0.000
	107	107	107	107
Ability_mean	.561**	1	.227*	.562**
	0.000		0.019	0.000
	107	107	107	107
Col_infl_mean	.292**	.227*	1	.207*
	0.002	0.019		0.032
	107	107	107	107
Stat_inst_mean	.356**	.562**	.207*	1
	0.000	0.000	0.032	
	107	107	107	107

***. Correlation is significant at the 0.01 level (2-tailed).*

**. Correlation is significant at the 0.05 level (2-tailed).*

Limitations, Implications, Suggestions for future research and Conclusions

The current study is consistent with previous studies in terms of Attitude towards Statistics" being a multi-dimensional construct. However, this study differs from the previous studies in terms of the incorporation of the new factor, Environmental or College Influence. While Cronbach Alpha reliability estimates for three subscales were 0.942, 0.932 and 0.909 were

high, the reliability of the College Influence sub-scale was found to be 0.738, indicating that more items are needed for this subscale. One major drawback of this study is the low sample size of 107. More data need to be collected before a CFA study is carried out.

It is evident from prior studies that environmental factors, parental expectations and peer influences affect or shape students' attitude or beliefs of college students. It stands to reason that that positive influence from peers and/or parental or expectations, semi malleable factors, might improve the attitude of students towards Statistics. The current study suggests that undergraduate students do think that mastery of statistical skills could improve their job prospects in future and can therefore serve as a motivating factor. Statistical anxiety, on thither hand seems to be related to the instructor factor. Perhaps, during introductory lectures, statistics instructors could highlight or discuss the importance of statistics and its relevance in terms of future job prospects with their students in details so as to give them a better idea of the future uses of Statistics. To alleviate statistical anxiety and increase student's self-efficacy, instructors could show through examples that use simple calculations, the principles underlying statistical concepts. These suggestions would not only help the students taking Statistics classes, but would also create a positive impression of the statistics instructor, thereby creating a positive attitude of students towards Statistics.

References:

Agliata, A, Renk, K. (2008). College Students' Adjustment: The Role of Parent–College Student Expectation Discrepancies and Communication Reciprocity. *Journal of Youth and Adolescence*. 37(8), 967-982.

Arumugam, R.N. (2014). Students Attitude towards Introductory Statistics course at Public Universities using Partial Least Square Analyses. *Interdisciplinary Journal of Contemporary Research in Business,* Vol 4, No 4.

Baloglu, M. (2003). Individual differences in Statistics Anxiety among college students. *Personality and Individual Differences*, 34(5), 855-865.

Cohen. J (1983). Peer Influence on College Aspirations with Initial Aspirations Controlled. *American Sociological Review*, 48(5), 728-734.

Doren. B, Gau. J.M, Lindstrom. L. E. (2012). The Relationship between Parent Expectations and Postschool Outcomes of Adolescents with Disabilities. *Council for Exceptional Children*, 79(1), 7-23.

Hong, K, Chew, P, Dillon, D. B. (2014). Statistics anxiety and the Big Five personality factors. *Social and Behavioral Sciences,* 112, 1177 – 1186.

Hsiao, T.Y. (2010). The Statistical Anxiety Rating Scale: Further Evidence for Multidimensionality. *Psychological Reports*, 107, 977-982.

Mji, A. (2009). Differences in University Students' Attitudes and Anxiety about Statistics. *Psychological Reports*, 104, 737-744.

Nolan. M. M, Beran. T, Hecker. K. G. (2012). Survey Assessing Students Attitudes toward

Statistics: A systematic review of Validity and Reliability. *Statistics Education Research Journal*, 11(2), 103-123.

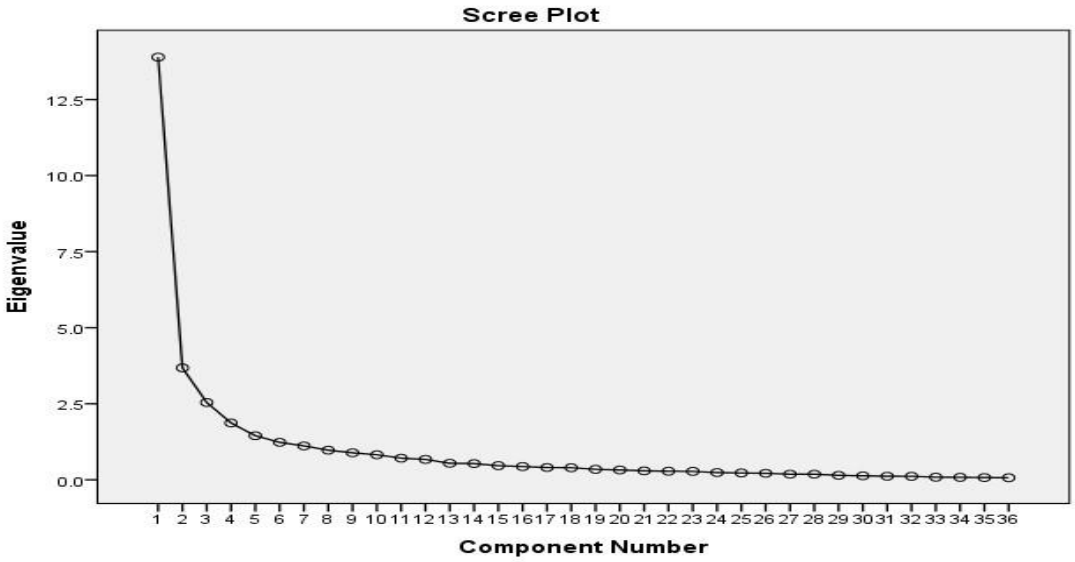
Onwuegbuzie, A. J., Leech, N. L., Murtonen, M., Tahtinen, J. (2010). Utilizing Mixed Methods in Teaching Environments to reduce Statistical anxiety. *International Journal of Multiple Research Approaches*, 4, 28-39.

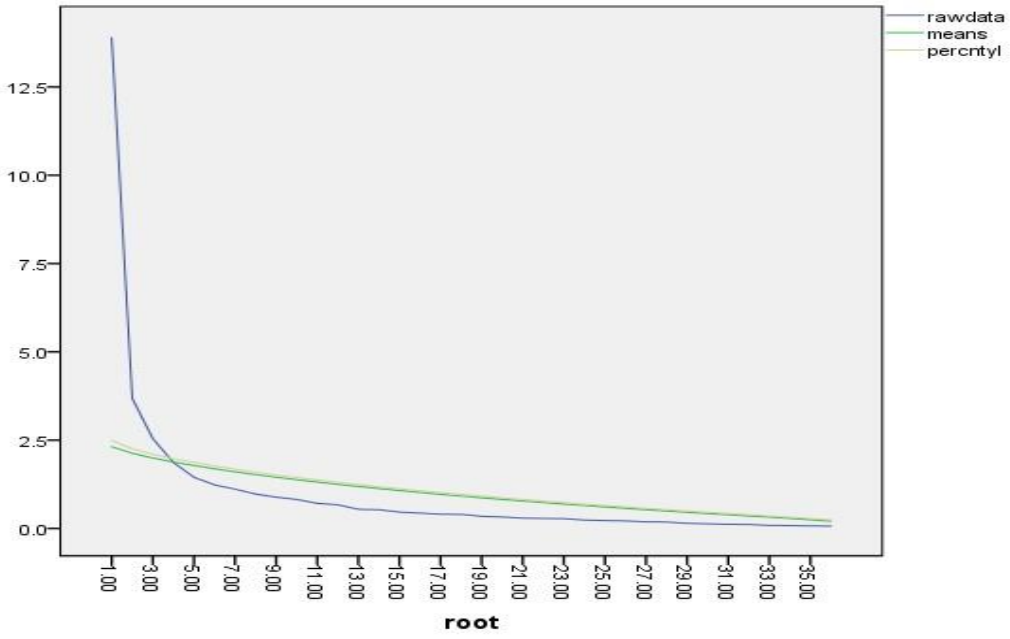
Pappusek, I, Ruggeri, K, Macher, D, Paechter, M, Heene, M, Weiss, E.M.,... Freudenthaler, H.H (2012). Psychometric Evaluation and Experimental Validationof the Statistics Anxiety Rating Scale. *Journal of Personality Assessment*, 94(1), 82-91.

Perepiczka. M, Chandler. N, Becerra. M. (2011). Relationship between Graduate Students' Statistics Self-Efficacy, Statistics Anxiety, Attitude towards Statistics, and Social Support. *The Professional Counselor*, 1(2), 99-108.

Williams, A. S. (2013). Worry, Intolerance of Uncertainty, and Statistics Anxiety. *Statistics Education Research Journal, 12, 48-59.*

APPENDIX: I


Scree Plot
Eigenvalue
12.5
10.0
7.5
5.0
2.5
0.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36
Component Number


rawdata
means
percntyl
12.5
10.0
7.5
5.0
2.5
0.0
1.00
3.00
5.00
7.00
9.00
11.00
13.00
15.00
17.00
19.00
21.00
23.00
25.00
27.00
29.00
31.00
33.00
35.00
root

Appendix J

Revised Instrument (after Content Validation)

Factors	Conceptual definitions
Usefulness of Statistics	This measures an individuals' perception of the role or relevance of Statistics in his/her life
Computational Selfconcept	This can be described as an individual's perception of his/her ability to understand and solve statistical problems regardless of actual ability
Statistics Instructor	This can be referred to as an individual's mental impression or way of regarding his/her Statistics Instructor
Environmental Influence	This factor can be described as how an individuals' perceives the environmental conditions while deciding to study Statistics

All the questions of the survey are on a 7 point Likert Scale (Strongly Disagree to Strongly Agree), where:

1- Strongly Disagree, 2- Disagree, 3- Slightly Disagree, 4- Neutral, 5- Slightly Agree, 6- Agree, 7- Strongly Agree.

Please write a number from 1-7 (in the 3rd column) against each of the questions given below.

Index	Questions of the survey	Indicate your response (1-7)
1	I think Statistics is relevant for my learning	
2	Statistical questions are not difficult to understand	

3	I feel I learn a lot from Statistics lectures.	
4	I would be more likely to study Statistics if a family member suggests	
5	I think Statistics is used in all fields.	
6	I think Statistical skills will improve my job prospects	
7	I like to do Statistics assignments	
8	I feel comfortable approaching my Statistics teacher	
9	I feel everyone should learn Statistics	
10	I use a lot of Statistics in everyday life	
11	I think it is easy to solve Statistics assignments	
12	I was motivated by my friends to study Statistics	
13	I like going to my Statistics classes	
14	I think Statistical decisions are valid and trustworthy	
15	I use various Statistical techniques to analyze data	
16	I feel I'm more likely to study Statistics if my adviser thinks so	
17	I find it easy to succeed in a Statistics course	
18	I think Statistical explanations are reasonable	
19	I feel I can earn a good grade in a Statistics class like my other friends	
20	I find my Statistics Instructor to be nice	

21	I use Statistics for my research projects	
22	My adviser thinks Statistics would help me in other classes that involves data analysis	
23	I find it interesting solving Statistics problems	
24	I like to participate in my Statistics classes	
25	The increase in data analysis in recent years motivated me to study Statistics	
26	I think I can do well on the Statistics exams	
27	I think Statistics classes are easier if you have a friend to study with	
28	I think my Statistics teacher supports our learning	
29	I enrolled in the Statistics class with my partner so that I could discuss with him/her	
30	I feel Statistics should be made mandatory in colleges	
31	Solving Statistics assignments doesn't take much time for me	
32	I feel my Statistics teacher explains the concepts clearly	
33	I am more likely to succeed in a Statistics class if I took it with my friends	
34	I think I was smart enough to enroll in a Statistics course	
35	I like when my Statistics teacher assigns in-class problems	

36	My parents think Statistics would help me in future career.	

ADDITIONAL QUESTIONS: Please respond to the 4 questions below

Questions	Your Response
What is your major?	
What is your age?	
Which year of college are you in (like freshman, junior etc.)?	
What is your Gender?	

Appendix K

Final Instrument (After EFA and Reliability Analyses)

All the questions of the survey are on a 7 point Likert Scale (Strongly Disagree to Strongly Agree), where:

1- Strongly Disagree, 2- Disagree, 3- Slightly Disagree, 4- Neutral, 5- Slightly Agree, 6- Agree, 7- Strongly Agree.

Please write a number from 1-7 (in the 3rd column) against each of the questions given below.

Index	Questions of the survey	Indicate your response (1-7)
1	I think Statistics is relevant for my learning	
2	I think Statistics is used in all fields.	
3	I think Statistical skills will improve my job prospects	
4	I like to do Statistics assignments	
5	I feel everyone should learn Statistics	
6	I use a lot of Statistics in everyday life	
7	I think it is easy to solve Statistics assignments	
8	I was motivated by my friends to study Statistics	
9	I like going to my Statistics classes	
10	I think Statistical decisions are valid and trustworthy	
11	I use various Statistical techniques to analyze data	

12	I feel I'm more likely to study Statistics if my adviser thinks so	
13	I find it easy to succeed in a Statistics course	
14	I feel I can earn a good grade in a Statistics class like my other friends	
15	I find my Statistics Instructor to be nice	
16	I find it interesting solving Statistics problems	
17	I like to participate in my Statistics classes	
18	The increase in data analysis in recent years motivated me to study Statistics	
19	I think I can do well on the Statistics exams	
20	I think Statistics classes are easier if you have a friend to study with	
21	I think my Statistics teacher supports our learning	
22	I enrolled in the Statistics class with my partner so that I could discuss with him/her	
23	I feel Statistics should be made mandatory in colleges	
24	Solving Statistics assignments doesn't take much time for me	
25	I feel my Statistics teacher explains the concepts clearly	
26	My parents think Statistics would help me in future career.	

ADDITIONAL QUESTIONS: Please respond to the 4 questions below

Questions	Your Response
What is your major?	
What is your age?	
Which year of college are you in (like freshman, junior etc.)?	
What is your Gender?	