Factor Analysis on Student's Attitude Towards Statistics (SATS©) Survey and Analysis of CSUMB Students' Responses

By

TRAVIS BARTON, JEFFREY MARTINEZ & PIERRETTE PEREDIA

CAPSTONE PROJECT

Submitted in partial satisfaction of the requirements for the degree of

BACHELOR OF SCIENCE

in

MATHEMATICS

in the

COLLEGE OF SCIENCE, MEDIA ARTS, AND TECHNOLOGY

at

CALIFORNIA STATE UNIVERSITY,

MONTEREY BAY

Approved:

(Advisor)

Michael B. Scott

Department of Mathematics and Statistics

Spring 2017

Acknowledgements

We would like to express our most sincere gratitude to the following Professors who assisted us in this work:

Prof. Michael B. Scott for his positivity and encouragement as our capstone advisor.

Prof. Steven B. Kim for his in-depth help with our R coding and technical analysis questions.

Most of all we would like to thank Prof. Alana Unfried for providing the data and an opportunity to better the Mathematics and Statistics Department through statistical analysis. She provided coaching and expertise throughout the duration of our project, and consistently made time in her schedule for us.

Abstract

Our study is centered around the Student Attitudes Towards Statistics (SATS©) survey developed by Statistics Professor Candace Shau. This survey was administered twice to 2016 STAT 100 students, once at the beginning of the semester, and again after the semester was concluded. Our project was split into two major objectives. The first objective was to perform factor analysis on the survey itself, and determine if the six attitude components the survey was designed to measure (Affect, Cognitive Competence, Difficulty, Value, Interest, and Effort) are represented, or if there are underlying, correlated factors present. This was determined through exploratory factor analysis (FCA). From our results, we determined a 5 factor structure (n = 5 eigenvalues). The first two being a combination of multiple components, which we summarized as F1: Lack of Confidence and F2: Hopefulness. The remaining three coordinated solely with the original attitude components of F3: Value, F4: Effort and F5: Difficulty.

The second objective was to analyze the paired before and after results of the survey, and to determine if STAT 100 effected a difference in student's attitudes towards statistics. This was accomplished by a two-sided matched-pairs t-test for each factor. Using a null hypothesis that the mean difference of attitude factor scores were zero to determine if there was evidence of a change in attitudes after the course. Based on our sample, using a significance level of 0.05, we failed to reject the null hypothesis for F1: "Lack of Confidence" (t=1.3998, p = 0.1636) F5: "Difficulty" (t = -1.0171, p = 0.3106) that the true mean of the difference between pre and post test scores is zero. For the remaining 3 factors, we found evidence to reject the null hypothesis for, F2: "Hopefulness" (t = -9.6029, p = 2.2e-16), F3: "Value" (t = 6.0490, p = 2.2e-16), and F4: "Effort" (t = -7.0596, p = 2.2e-16).

Contents

Introduction

At the CSUMB Mathematics and Statistics Department, we are continually looking for ways to improve our methods in garnering student interest in the fields of Mathematics and Statistics. Although our department is small compared to others on campus, most incoming freshmen must, despite their major, take Mathematics and Statistics classes as part of their general education requirements. In 2015, the Statistics Concentration was added as an option to declared Math Majors, and several Statistics courses have been added as a result.

Through this project we aim to use a known metric to assess and measure student attitudes. This metric is known as the Student Attitudes Towards Statistics SATS-36© survey, whose origin we will cover in a following section. This survey is a statistical tool, and we will explore the tool first using factor analysis, then we will impliment the tool with a two-sample t-test. The purpose of factor analysis is to measure the structure and scope of our tool, and the purpose of the t-test is to interpret the results of our tool's measurements.

The field of Statistics is and continues to be a vital and growing subject not just in academia, but in all types of industry. While the need for proficiency in Statistics is present, it is not clear whether it is interesting or appealing to the aspiring CSUMB undergraduate. Through administering the SATS-36© survey to freshmen in a required STAT 100 course, there can be clear data on whether or not students see Statistics as an engaging subject. First the students were surveyed at the beginning of the semester (what we will call our "pre-test" results), then they were given the same survey at the end of the semester (what will be called our "post-test" results).

Each question in the SATS-36© survey was designed to measure only one of the six attitude components Affect (A), Cognitive Competence (C), Difficulty (D), Value (V), Interest (I) and Effort (E). Our first main research question relates to the actual composition of the (SATS-36©) survey, and the second relates to the overall results from STAT 100 students:

- Research Question 1: Upon performing factor analysis on the CSUMB pre-test responses, does our resulting factor structure of the SATS survey stay true to the original six-component design?
- Research Question 2: Compare the student pre-test responses to the post-test responses, is there evidence of a change of student attitudes towards statistics?

The remaining sections will cover the background of Factor Analysis and previous SATS-36© survey studies, a brief explanation of a two-sided t-test, then a summary of our Factor Analysis and results of our t-test, followed by a conclusion and discussion of those results.

Background and Literature Review

Factor Analysis

After a researcher has collected his or her data, they often would like to quantify the concepts that they are trying to test with the data that they have. This is particularly prevalent in the social sciences, as something like I.Q. is not something that one can describe easily. An abundance of elements need to be considered and weighted in order to gain useful insight. Because of reasons like funding, availability of subjects, ethical concerns and more, it is most likely not possible for the researcher to run tests for all of the variables they need, and will need to settle for less than perfect data. To compensate for this, they need statistical methods that will produce meaningful results from data that they can reasonably collect. One such method is Factor Analysis.

Factor Analysis is the manipulation of one's data to create a new data set that is both condensed and weighted to give the most parsimonious and accurate model possible. This is done by taking the set of all variables, identifying and grouping together variables that have common variance. Thus revealing a smaller set of underlying, unobserved variables that are representative of the whole, called factors. This is performed through various techniques that involve principals from linear algebra, primarily the concept of eigenvectors and eigenvalues. Conditions when a researcher might use this method are diverse and include reasons such as: they may want to create a model that describes the interactions between variables; they may need to test a theory that specific unseen components describe the interactions between variables; they might want to see how a model's variables are changed by the altering of conditions which the data was taken; they may be testing a hypotheses from previous experiment with a sample of the same or different populations; and they may be testing how a result changes based on a variation of the procedure taken to get those results. (1992) This is not an exhaustive list, but it should give an idea of when a researcher might consider factor analysis as a viable approach. (Comery 1992)

Depending on one's definitions, there are roughly five steps (Figure 1) in performing factor analysis on data. The first step in factor analysis is the same as in any statistical procedure, which is collecting/selecting the data and choosing the right variables. Data can be collected directly from the source of study or may be whatever the researcher has access to. The second step is really the first step of the actual factor analysis. It involves looking at our data's covariance matrix and then computing its generic eigenvalues. The eigenvalues with the highest values are the linear combinations of variables with the greatest variance. This is important to note, as this is how a researcher would change his or her frame of reference to see their data from a new angle, literally. This eigenvector allows the researcher to make a new matrix called the correlation matrix, which acts as a linear transformation of the covariance matrix. These values are used to create new variables called factors, which will allow the researcher to have a more parsimonious model later on in the procedure. Once they have their new data, they can begin the third step, which is to examine the factors

Diagram.png

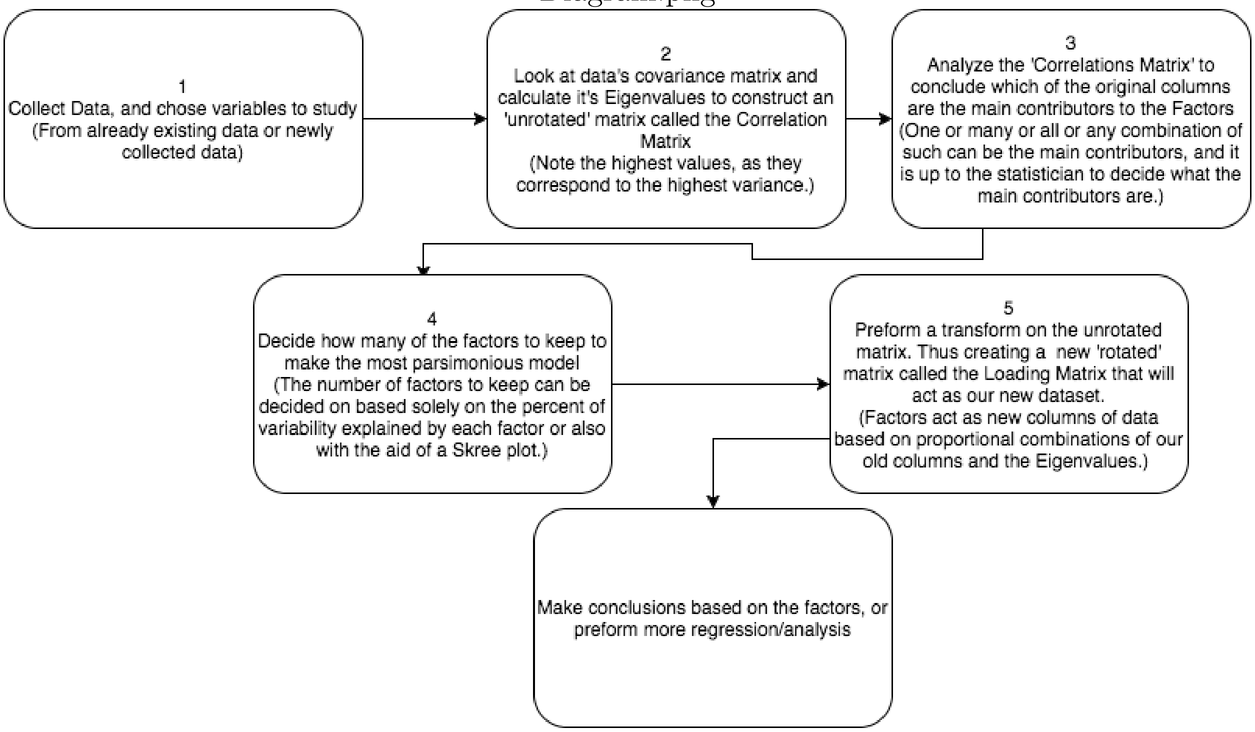

FIGURE 1. 5 Steps of Factor Analysis

and make any inferences possible before the data is rotated. A good way to visualize of this process is to picture the researcher turning the data in order to align the axis with the most variable section of data. (1992) (See Figure 2)

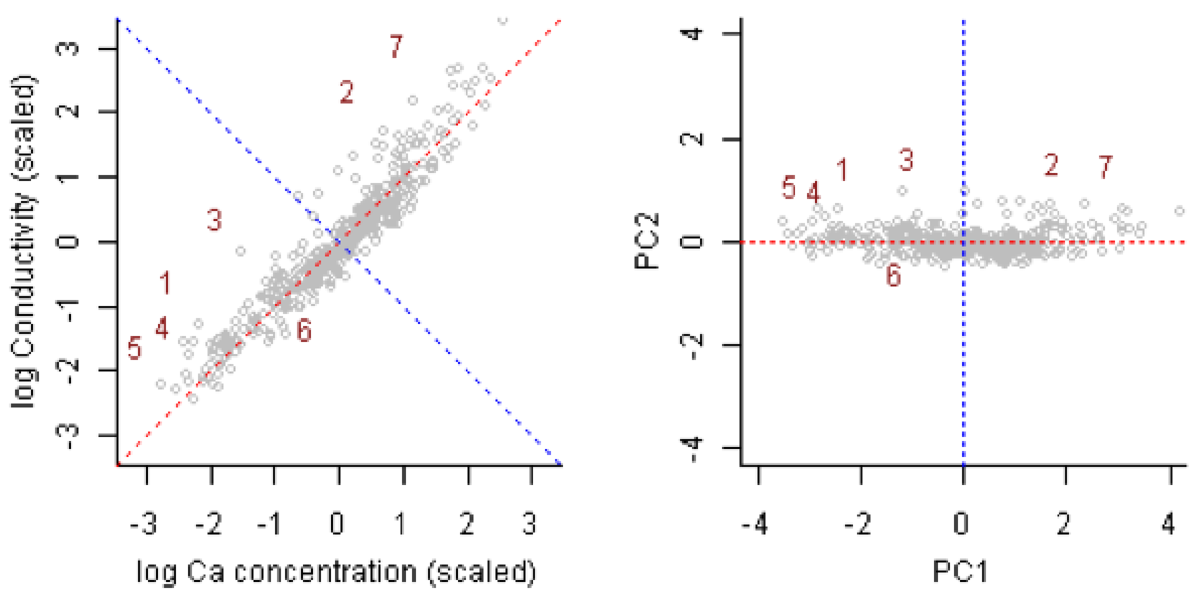

FIGURE 2. https://www3.epa.gov/caddis/da_exploratory_5.details.html

Let it be noted, that although the above figure is a result of Principal Components Analysis (PCA), our primary focus is on Factor Analysis (FA). The illustrative idea is the same, but the difference between the two methods is that PCA analyzes all sources of variance for each variable, while factor analysis only focuses on shared variance between variables (Mertler, Vanatta 2005). Also, FA assumes that there exists latent factors that account for the covariance between variables, while PCA does not.

After the new dataset has been rotated, the fourth step is to look at the Loadings Matrix. The Loadings Matrix is a table that displays, among other things, the analysis of how much of the data's variability can be explained by each of the factors. Figure 3 is an example of how an output may look when summarizing a correlation matrix.

Total Variance Explained

	Initial Eigenvalues[b]			Extraction Sums of Squared Loadings[f]			Rotation Sums of Squared Loadings[g]		
Factor[a]	Total[c]	% of Variance[d]	Cumulative %[e]	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	6.249	52.076	52.076	5.851	48.759	48.759	2.950	24.583	24.583
2	1.229	10.246	62.322	.806	6.719	55.478	2.655	22.127	46.710
3	.719	5.992	68.313	.360	3.000	58.478	1.412	11.769	58.478
4	.613	5.109	73.423						
5	.561	4.676	78.099						
6	.503	4.192	82.291						
7	.471	3.927	86.218						
8	.389	3.240	89.458						
9	.368	3.066	92.524						
10	.328	2.735	95.259						
11	.317	2.645	97.904						
12	.252	2.096	100.000						

Extraction Method: Principal Axis Factoring.

Figure 3

Notice that the first half of the factors explain over 80% of the data's variability. Factor analysis, will not require all of the factors but only those which explain the vast majority of the variability. Cutting out some of the factors that do not contribute much to the cumulative variability allows for a more parsimonious model. There is no specific cutoff for how much of the variability needs to be explained, and it is up to the researcher on a case by case basis to decide how to best balance the reduction of the number of factors while explaining as much variability as possible. Tools that can aid a researcher in this are scree plots, parallel analysis, proportions, and more. After the factors have been created, and our data manipulated, the researcher can begin their chosen regression techniques to test their hypotheses.

While this is the general form of the results, we will be using slightly different methods of calculating the correlation matrix, estimating the mean and variance, and rotating the correlations matrix. For calculating the correlations matrix, we will be using a parallel analysis package in R called Parallel, which will calculate the correct number of factors to analyze by comparing our data to that of a randomized data set and doing the factor analysis steps on both. (Bruin 2011) To estimate the mean and specific variance of our factors, we will be using maximum likelihood estimators. Which involve finding the mean, loadings matrix, and error that maximize the parameters fit to the data that we see. (Penn State Online)

To rotate our matrix, we will use the Varimax method. It is the most common method used and relies on scaling the correlations matrix in order to see a more apparent separation of factors. Smaller loadings are diminshed while larger loadings are made expanded. (1992)

History

Although it is accepted that factor analysis is a statistical technique, the theory behind it is traced back to studies in psychometrics, the science of measuring mental capacities and processes. In the early 20th Century, psychologist Charles Spearman speculated that there was a general factor g apparent in in several human cognitive abilities. This factor g was the result of statistical analysis, "Because measures of seemingly different mental abilities consistently indicate correlations, he concluded that the prevalence of positive correlations must result from the general factor, g." (Brittanica 2008) This g eventually led to the methodology used to measure overall intelligence, what most know today as the I.Q. test. Thus, factor analysis came into practice as a method of synthesizing multiple observations, then minimizing data overlap in order to better comprehend a hidden, general underlying structure.

Application Of Factor Analysis to SATS©

For our study we apply the technique of factor analysis in the context of survey structure using the Survey of Attitudes Towards Statistics (SATS©) statistical student questionnaire developed by Candace Shau in 1992 (Shau 2003). This survey was originally administered as a 28 question survey designed to measure four attitudes of postsecondary student attitudes towards statistics: affect (A, measuring a student's feelings concerning statistics), cognitive competence (C, measuring the student's personal assessment of their knowledge concerning statistics) , difficulty (D, measuring the difficulty of statistics as a subject) and value (V, measuring the student's attitude regarding the usefulness, relevance and worth of statistics in personal and professional life) (Shau, et al. 1995). Two other attitude components: interest (I, measuring whether statistics is appealing) and effort (E, measuring the perceived personal exertion to learn statistics) were added, subsequently resulting in the more current version and focus of our study, the 36 question SATS-36©(2003). Students answer these questions on a 7-point Likert scale, with available responses spanning from strongly disagree to strongly agree. Since it is designed to measure changes in students' attitudes, the survey should be administered at least twice, usually once at the beginning of the course, and then again after or near completion.

Given that it is a 36 question survey based on 6 attitude components, it may seem fruitless to try to identify latent factors at work apart from the aforementioned 6. Through factor analysis, instead of trying to contradict or oppose the original intent of the survey, it is an endeavor of efficiency. When the survey is administered to a new population, we use factor analysis to assess whether the intended structure holds. Previous empirical studies have shown that a three-factor or two-factor construct closely fit the original four-component SATS-28© survey (Vanhoof 2011), but too little a number of factors tended to result in overgeneralization of results. These constructs combined the known components of Affect and Cognitive Competence in the three-factor construct, and the two-factor construct agglomerated Affect, Cognitive Competence and Difficulty into a single factor. These studies proved to condense the survey results, but ultimately did not offer much unique information regarding different areas of student attitudes. By this we can caution ourselves to focus not just on minimizing factors, but to accurately represent the overlap of components within the survey while retaining the overall assessment of diverse attitudes.

Other results of Factor Analysis showed the SATS-36© could be summarized by as little as four factors. In a study very similar to ours done at Winona State University, five-factor and four-factor constructs were explored (Dickey, Seifert 2014). The five-factor model saw a unique factor represented in all the Cognitive Competence questions, another unique factor was identifiable in the Effort questions, but the other three factors were a mix of the remaining components. Upon analysis using only four factors, one factor was comprised of only the Affect component, another entirely comprised of the Value component, the third factor was made of a combination of the Affect and Interest, and the last was a mix of Affect, Difficulty, Cognitive Competence and Value components (2014). We did not expect our Factor Analysis results to deviate drastically from these previous studies. The differing element is that our data set is composed of CSUMB student responses, so we could not anticipate there to be identical component combinations, but similar results.

Application of Hypothesis Testing to Survey Results

The hypothesis test is a method of determining which belief to hold based on the statistical evaluation of the evidence (study data) provided. There are two options to decide between in a hypothesis test, the Null Hypothesis, a statement about our parameter, and the Alternative Hypothesis, which is the idea that is being tested. The Null and Alternative Hypothesis are always different. We construct a known probability distribution based on the assumption of the null hypothesis, transform our observed data into a test statistic, then calculate how probable it is (p-value) to observe our test-statistic under the null hypothesis distribution. If our observed test statistic has a p-value less than our pre-selected significance level α, then we have statistical evidence to reject our null hypothesis in support of our alternative hypothesis (Kim 2015).

The particular belief that was examined in this survey (Alternative Hypothesis) was whether or not STAT 100 students that took the course changed their views of statistics. The Matched Paired t-test was used because our pre and post data was gathered from the same population. Instead of analyzing both populations as separate groups, the pre and post results are quantified into a single population by taking the mean difference between each students' pre and post results.

This method was first practiced in 1908 by William Gosset as a method of quality control for beer (2015). He noted that the sampling distribution of the differences in the population means followed a known T-distribution with varying degrees of freedom (df) based on the number of paired differences in the data. This T-distribution is very robust even for small sample sizes. For our study, we have more than enough data points (n ¿ 30) to depend on our results from this method.

Analysis and Results

Participants

The participants in our study are an aggregate of five different classes, all of which are groups of students who have taken the same survey but in different introductory statistics sections. The survey was a written questionnaire comprised of 36 questions or statements, with student answers given on a 7-point Likert scale from 'Strongly Disagree' to 'Strongly Agree'. All participants were taking the same class (STATS 100) at California State University, Monterey Bay (CSUMB). The survey was first administered to 324 students during the initial week of instruction for the fall semester, August 2016. For the remainder of our paper, we will refer to this initial survey as our *"pre-test."* In order to perform Factor Analysis, we only needed the pre-test results. We chose these results since there was a larger sample size $n_{pre} = 324$ of total student responses.

The survey was again administered on the last week of instruction, December 2016. This second version was identical to the pre-test survey, the only difference being that the tenses were changed. Due to students dropping the course, and one section of STAT 100 failing to take the second administration of the test, only $n_{post} = 191$ students took the second survey. For the remainder of our paper, these second survey results will be referred to as our *"post-test"* results.

Let it be noted that a matched-pairs t-test requires our pre-test and post-test data to be paired pre and post results. Upon exploration of our data, we determined that only 184 students took both the pre-test and post-test. Therefore, the results of our matched-pairs t-test is based on a sample of size $n_{paired} = 184$.

Statistical Analyses

We performed the following Analyses using R to test factor models for the SATS-36© survey.

Parallel Analysis. First we performed parallel analysis to extract the number of factors to retain using the parallel function in R (See Appendix A), which gave us an optimal number of eigenvalues (n=5, See scree plot in fig. 4). This plot tells us that the first 5 eigenvalues account for most of the variance in our data, so if we omit the remaining eigenvalues we are not losing valuable information.

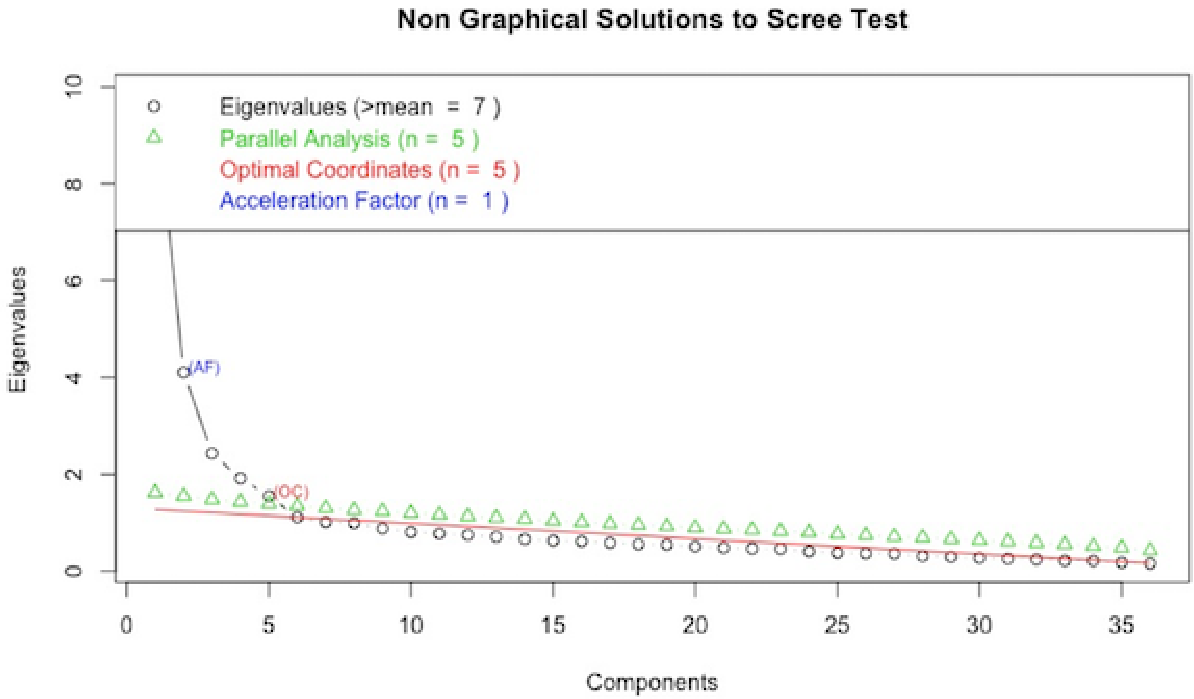

FIGURE 4. Parallel Analysis Scree Output

Factor Analysis. Using this result, we performed Factor Analysis exploring a five-factor model using the Factanal() function in R. This extracts five eigenvalues (one for each factor) from our data, generates a matrix of eigenvalue-eigenvector pairs, and rotates it based on our specified varimax method. The output results in a loadings matrix. Taking precautions to make sure our results were clear, we omit any null responses. The result gives us a loadings matrix showing the strength of which factor is represented in each statement. Loadings are given as a value between [-1,1], but we look for larger values so we omit loadings with an absolute value less than 0.40. We choose this cutoff arbitrarily, but the purpose is to see an interpretable picture of our factor structure. The resulting loadings matrix is given in Figure 5 in the results section with the questions re-ordered for clarity.

Hypothesis Testing. In order to properly analyze the results of the SATS© survey, we needed to ensure that we had data from students who took both the pre-test and the post-test. This was accomplished by merging our pre and post data sets by the unique Student ID column. The result was a larger data set showing 184 survey respondents, and included both pre-test and post-test results for those students found in both data sets. Pre-test answers were notated by capital letter then a number, i.e. A3 or C31. Post-test answers were notated by a lowercase *p*, i.e. pA3, or pC31. Since our study was based on our proposed 5-factor construct from Factor Analysis, we needed to properly interpret our results as such. Our Loadings matrix showed multiple statements as negatively loaded on certain factors, so when we constructed our values to compare sample means, we needed to transform our results. Negatively loaded answers are inverted (transformed based on a 7 point Likert scale) to reflect our factor construct. For example, statement C31 *"I can learn statistics"* loaded negatively on factor F1: "Lack of Confidence", this means the more a student agreed with this statement the less the student had a lack of confidence. Transformed or inverted answers were notated by uppercase J, i.e. if C31 = 3, then JC31 = 8 - C31 = 5. Finally, we partitioned our pre and post answers by statements loaded on F1, F2, F3, F4 and F5 respectfully. For example we combined into one data set of student answers for pre-test statements loaded on F3 (i.e. answers for statements V7, JV10, V21, V25, V16, V13 and V33 become X_{F3Pre}), and made a corresponding post-test group (i.e. answers for statements pV7, JpV10, pV21, pV25, pV16, pV13, and pV33 become Y_{F3Post}). We called these groups pre and post-test factor scores. We then took the 10 different means of the 5 pairs of groups to calculate our test statistic in our hypothesis tests.

We then set up the following hypothesis test to test whether or not the difference in pre-test and post-test factor score means for F1 was equal to zero:

- Null Hypothesis H_0: μ_{F1post} - μ_{F1pre} = 0, there is no difference of pre-test and post-test factor score means for F1.
- Alternative Hypothesis H_1: μ_{F1post} - $\mu_{F1pre} \neq 0$, there is a difference of pre-test and post-test factor score means for F1.
- We fix a significance level of $\alpha = 0.05$
- We calculated our test statistic as $T_{F1} = \frac{(\bar{Y}_{F1post} - \bar{X}_{F1pre}) \ - \ 0}{\sqrt{S^2_{F1}/184}}$ where S^2_{F1} is the estimator for the variance of $Y_{F1post} - X_{F1Pre}$.
- We use this test statistic using the known $\mathcal{T}$-distribution $T_{F1} \sim \mathcal{T}_{\mathcal{V}}$ where $\mathcal{V}$ is the degrees of freedom equal to one less than our sample size of paired F1 answers.
- We then reject or fail to reject the null hypothesis based on our calculated test statistic T. If the probability of observing our test statistic is less than our significance level ($P(T_{F1}) \leq \alpha$), then we reject the null hypothesis.

We administer this T-test using the following code in R:

```
\# input pF1 = post-test data, F1 = pre-test data
t.test(F1post-F1pre, alternative = "two-sided")
\# output of test statistic, degrees of freedom, and probability
t = 1.3998, df = 156, p-value = 0.1636

# \
^
```

The output gives us our test statistic T_{F1}, the degrees of freedom df = $\mathcal{V}$ and the probability (p-value) of observing our test statistic when μ_{F1post} - μ_{F1pre} = 0. We then repeated this process for each of our factor scores for F2, F3, F4, and F5. The results for each of the 5 factors will be shown below.

Results

Factor Analysis Results. After performing Factor Analysis on our pre-test data, the result was the loadings matrix shown in Figure 5. The survey statements have been reordered in order to better interpret our factors. We have notated each question in the first column to show the original component it was meant to represent, then we have written out the actual statement as the student would have seen it. Twelve statements were loaded on our first factor F1. These statements were originally meant to measure Cognitive Competence, Difficulty, and Affect. Seven statements heavily loaded on our second factor F2, these were designed as Value, Affect and Interest statements. Our 3rd factor F3 was loaded entirely

Original Assignment	Question/Statement	F1: Lack of Confidence	F2: Hopefulness	F3: Value	F4: Effort	F5: Difficulty
Difficulty24	Learning statistics requires a great deal of discipline					0.437
Difficulty34	Statistics in highly technical					0.657
Difficulty30	Statistics involves massive computations					0.687
Effort27	I plan to attend every statistics class session				0.638	
Effort14	I plan to study very hard for every statistics test				0.657	
Effort02	I plan to work hard in my statistics courses				0.808	
Effort01	I plan to complete all of my statistics assignments				0.812	
Value10	Statistical skill will make me more employable			-0.403		
Value21	Statistics conclusions are rarely presented in everyday life			0.455		
Value07	Statistics is worthless			0.531		
Value25	I will have no application for statistics in my profession			0.639		
Value16	Statistical thinking is not applicable in my life outside of my job			0.643		
Value13	Statistics is not useful to the typical professional			0.655		
Value33	Statistics is irrelevant in my life			0.663		
Value09	Statistics should be a required part of my professional training		0.416			
Interest12	I am interested in being able to communicate statistical information to others		0.590			
Affect3	I will like Statistics		0.597			
Interest23	I am interested in understanding statistical information		0.627			
Affect19	I will enjoy taking statistics courses		0.658			
Interest29	I am interested in learning statistics		0.713			
Interest20	I am interested in using statistics		0.765			
CogCompetence32	I will understand statistics equations	-0.522				
Difficulty06	Statistics formulas are easy to understand	-0.517				
CogCompetence31	I can learn statistics	-0.445				
CogCompetence26	I will make a lot of math errors in statistics	0.455				
Difficulty08	Statistics is a too complicated subject	0.537				
Affect15	I will get frustrated going over statistics tests in class	0.597				
CogCompetence11	I will have no idea what's going on in this statistics course	0.624				
Affect04	I will feel insecure when I have to do statistics problems	0.692				
Affect18	I will be under stress during statistics courses	0.694				
CogCompetence5	I will have trouble understanding statistics because of how I think	0.726				
CogCompetence36	I will find it difficult to understand statistical concepts	0.727				
Affect28	I am scared by statistics	0.758				
Value17	I use statistics in my everyday life					
Difficulty22	Statistics is a subject quickly learned by most people					
Difficulty36	Most people have to learn a new way of thinking to do statistics					

FIGURE 5. Loadings Matrix

with Value statements, the 4th factor F4 with Effort statements, and the 5th factor F5 with Difficulty statements. Note at the bottom of the loadings matrix, we have 3 questions that did not have loadings with magnitudes greater than 0.40.

Matched-Pairs T-Test Results. Results of our two sided T-test computed using R:

- F1: "Lack of Confidence" ($t=1.3998$, $p = 0.1636$)
- F2: "Hopefulness" ($t = -9.6029$, $p = 2.2e\text{-}16$)
- F3: "Value" ($t = 6.0490$, $p = 2.2e\text{-}16$)
- F4: "Effort" ($t = -7.0596$, $p = 2.2e\text{-}16$)
- F5: "Difficulty" ($t = -1.0171$, $p = 0.3106$)

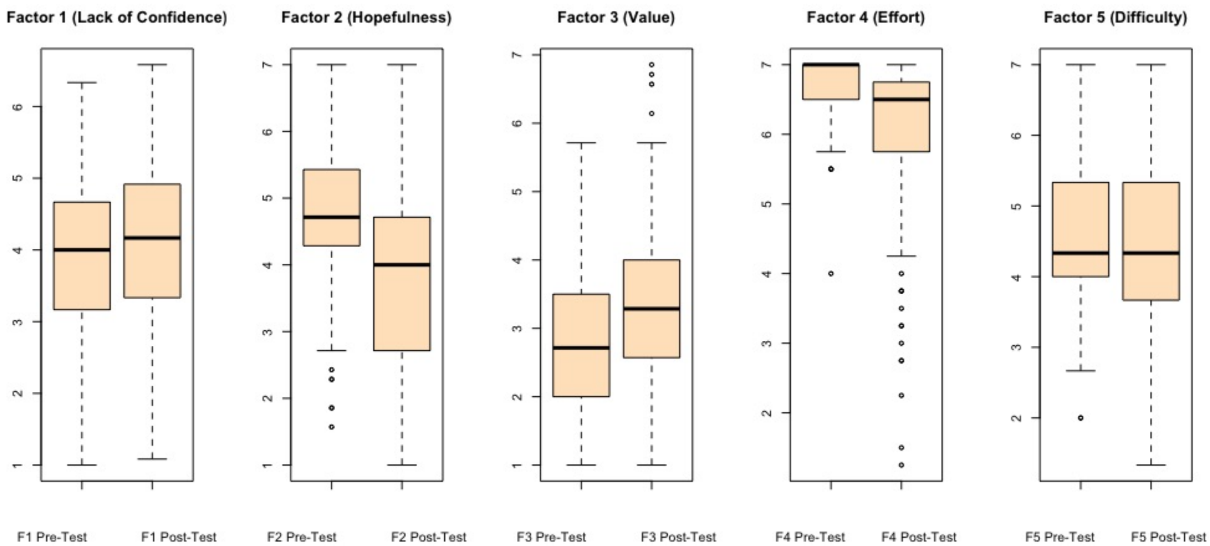

FIGURE 6. Mean Difference in Factor Scores

Conclusion

Factor Analysis Discussion. With accordance to our results, we have concluded that the number of factors needed to obtain the most parsimonious model is less than Candace Schau's original six, but more than the recommendation of some studies, four. According to our results, our data suggests that five is the best number of factors to retain in order to have the strongest possible data with the least number of explanatory variables. Therefore we answer in the negative to our first research question: "Upon performing factor analysis on the CSUMB pre-test responses, does our resulting factor structure of the SATS survey stay true to the original six-component design?" This goes against the research done by Schau for the Joint Statistical Meetings in 2008, who based her work on a collection of articles and journals from reputable mathematical and statistical sources (Schau 2008).

We have also concluded that the breakdown of our five factors is as such: Factor 1 (F1), which is comprised mostly of the original columns Affect, Cognitive Competency, and Difficulty, very much resembles the qualities one might associate with Confidence, and shall be labeled as "Lack of Confidence." Some evidence to support that conclusion lies in statements like *"I like Statistics"* being negatively associated and *"I feel insecure when I have to do a statistics problem"* being positively associated with F1. Factor 2 (F2), which is mostly comprised of the original components Affect, Value, and Interest point towards a general open mindedness and optimism towards statistics, and have thus lead us to give it the title of "Hopefulness." Evidence that supports this designation can be seen in questions like *"I like Statistics"* and *"Statistics should be a required part of my professional training"* being positively associated with F2. Factors 3 (F3), 4 (F4), and 5 (F5) all have strong associations with Value, Effort, and Difficulty respectively, and no others. Thus we have concluded that they should keep the name of their sole association. Hence F3 is to be designated Value, F4 shall be Effort, and F5 shall be Difficulty.

As stated before, we did not expect our results to vary from previous studies. Our resulting 5-factor structure does not necessarily override the original survey components, but it does shed light on the underlying commonalities intermingled within them. Perhaps the most interesting finding of our study were the three statements seen at the bottom of our loadings matrix that did not have strong loadings for any of our factors. These questions most likely would vary depending on a students field of study, or learning style. For example statements like *"Statistics is a subject quickly learned by most people"* and *"Most people have to learn a new way of thinking to do statistics"* meant to assess the difficulty attitude component does not necessarily reflect how difficult the subject seems to a student. Perhaps all college courses aren't quickly learned, and what type of thinking constitutes a *"...new way of thinking*?" Unless a student feels particularly opinionated regarding the speed or method of learning, most likely they would choose a neutral answer based the objectiveness of the wording. Perhaps these could be omitted from future revisions of the survey.

Matched Pairs T-test Discussion. We used a null hypothesis that the mean difference of attitude factor scores were zero to determine if there was evidence of a change in attitudes after the course. This was administered individually to each of the five answer sets of statements that were loaded heavily on each factor. Based on our sample, using a significance level of 0.05, we failed to reject the null hypothesis for F1: "Lack of Confidence" ($t=1.3998$, $p = 0.1636$) F5: "Difficulty" ($t = -1.0171$, $p = 0.3106$) that the true mean of the difference between pre and post test scores is zero. For the remaining 3 factors, based on our sample, using a significance level of 0.05, we found evidence to reject the null hypothesis for, F2: "Hopefulness" ($t = -9.6029$, $p = 2.2e\text{-}16$), F3: "Value" ($t = 6.0490$, $p = 2.2e\text{-}16$), and F4: "Effort" ($t = -7.0596$, $p = 2.2e\text{-}16$) and find evidence to support the alternative that the true mean of the difference between pre and post test scores for those factors is not zero.

Therefore, we have statistical evidence to support that students' "Lack of Confidence" and their personal assessment of how "Difficult" statistics was unchanged by taking STAT 100. Comparitively we found evidence to support the alternative hypothesis that there was in fact a change in students' "Hopefulness" towards statistics, a change in their perception of the overall "Value" of statistics, and a change in the perceived amount of "Effort" required to learn statistics.

Our five t-tests were two-sided, meaning we did not test for postive or negative differences in factor score means, only whether or not there was a significant difference not equal to zero. We did note that the overall trend in the three factors where a difference was apparent was a decrease in positive attitudes towards statistics. Looking at our boxplots (Fig. 6) from our results section, we see the mean difference as the change in the black lines from pre-test to post-test results. A sharp drop in student "Hopefulness" is clearly seen in Factor 2, but Factor 3 seems to show an increase in "Value." But when we consider the wording of the survey statements heavily loaded on Factor 3, we see there is positive loadings for statements like *"Statistics is irrelevant in my life"* and *"Statistics is not useful to the typical professional"*, while the statement *"Statistical skill will make me more employable"* is negativeley loaded. What this means is that our Factor 3 is in fact measuring "Value", but for the statements loaded on that factor, a higher positive score is interpreted as a decrease in value. So when we see an increase in that factor score, it is actually an unfavorable result towards statistics.

In Factor 4 from our boxplots we see a drop in factor scores related to "Effort." This could be interpreted as a good sign, as a positive attitude towards statistics, that students through taking a course in statistics see it as less effort as they first thought. Once again, the wording of the four statements gives us insight to explain this change. For each of the four statements, they all have a common wording: *"I plan to..."*. We would argue that answers to this statement would not necessarily give a clear reading into the students' perceived amount of "Effort" required to do statistics. Instead it is more insightful into whether or not the student plans to take additional statistics courses beyond STAT 100.

Final Discussion and Future Studies. As we analyze the results of our study, we can make alterations to the course curriculum to achieve our desired results. For our study, since we saw a drop in student "Hopefulness", perhaps more resources could be made for STAT 100 students, such as tutors and study groups to boost student optimism towards tackling the material. Since we saw a negative result for our "Value" factor, perhaps throughout the course there could be brief examples of statistics being used in current events or in modern technology. Or even pathways and salary information of statistical careers (Sports

Statisticians, Database Analysts, Data Scientists, Systems Engineers... etc) could be given to pique student curiosity.

These methods are given as examples, and whether or not they work could be assessed by repeating the study. It would be advantageous to practice administering the SATS©survey to STATS 100 courses in the future, so we can determine if the same trends in opinion hold for various study groups. At the CSUMB Mathematics and Statistics department, the aim is not to just instruct students in the field of Statistics or to gain more declared majors, instead the goal is to cultivate a basic foundation and a passion for the science. Moving forward we could use this study and analysis as a system to improve the presentation and content of the material, in hopes to increase appeal. It is unrealistic to believe that every student would suddenly fall in love with the subject, but perhaps they would have a clearer understanding of the overall worth and relevance of Statistics.

Bibliography

[1] Bruin, J *newtest: command to compute new test* 2011: UCLA

[2] Clark, Kristi Lynn *Undergraduate students‘ attitudes toward statistics in an introductory statistics class* 2013: UGA

[3] Comrey, Andrew *A First Course in Factor Analysis* 1992: Lawrence Erlbaum.

[4] Kim, Tae Kyun. *"T test as a parametric statistic." Korean Journal of Anesthesiology.* The Korean Society of Anesthesiologists, 25 Nov. 2015. Web. 08 May 2017.

[5] Mertler, Craig A and Vannatta, Rachel A *Advanced and Multivariate Statistical Methods* 2005: Pyrczak Publishing

[6] *Multivariate Explore* 2006: US EPA

[7] Penn State *12.7 - Maximum Likelihood Estimation Method.* Penn State Online

[8] Sanet Coetzee *Psychology Students' Attitudes Towards Statistics* 2010: Journal of Industrial Psychology

[9] Schau, Candace *Common Issues in STATS© reaserch* 2008: JSM

[10] Schau, Candace *Students' attitudes: The other important outcome in statistics education* 2003

[11] Schau, Candace, et al. *"The development and validation of the survey of attitudes toward statistics."* Educational and psychological measurement 55.5 (1995): 868-875.

[12] Spearman, Charles *The Proof and Measurement of Association Between Two Things* 1904: JSTOR

[13] Spearman, Charles *The Abilities of Man* 1927: Macmillan and Co., Limited.

[14] Vanhoof, Stijn and Kuppens, Sofie and Sotos, AE Castro and Verschaffel, Lieven and Onghena, Patrick *Measuring statistics attitudes: Structure of the survey of attitudes toward statistics (SATS-36)* 2011: International Association for Statistical Education

Appendix A: R-code and Output

Factor Analysis

```
\# combining 5 original STAT 100 sections
au1pre.survey = cbind(AU_1_Pre_Post_anon[,4], AU_1_Pre_Post_anon[,6:41])
au2pre.survey = cbind(AU_2_Pre_Post_anon[,4], AU_2_Pre_Post_anon[,6:41])
au3pre.survey = cbind(AU_3_Pre_Post_anon[,4], AU_3_Pre_Post_anon[,6:41])
au4pre.survey = cbind(AU_4_Pre_Post_anon[,4], AU_4_Pre_Post_anon[,6:41])
au5pre.survey = cbind(AU_5_Pre_Post_anon[,4], AU_5_Pre_Post_anon[,6:41])
\# Combining our data to include only survey results and ROSA ID
aupre = rbind(au1pre.survey, au2pre.survey, au3pre.survey, au4pre.survey, au5pre.survey)

View(aupre) \# view all pre-test student responses

\# Determine Number of Factors to Extract
library(nFactors)     \# installing nFactor package to extract factors
library(parallel)     \# installing parallel to extract factors
ev <- eigen(cor(na.omit(aupre)))  \# get eigenvalues
attach(ev)       \# attaching the values to eignvalues
ev               \# printed the output to see what was going on
ev$values      \# printing the output only the values

ap <- parallel(subject=nrow(aupre), var=ncol(aupre), rep=100, cent=.05)
\#  nScree returns info about number of components
\#  factors, kaiser and parallel analysis
nS <- nScree(ev$values, aparallel=ap$eigen$qevpea)
nS
\# plotting parallel Analysis
plotnScree(nS) \# Gives us the output of figure 4 in text

\# using factor analysis with 5 factors on the pretest
facs <- factanal(na.omit(aupre), factors = 5, length(x(x > 0.4)), rotation = "varimax")
\# prints the data
facs
\# shows loadings greater than 0.4
print(facs, cutoff = 0.4)
Loadings:
    Factor1 Factor2 Factor3 Factor4 Factor5
E1                            0.812
E2                            0.808
A3  -0.482   0.597
```

```
A4   0.692
C5   0.726
D6  -0.517
V7                  0.531
D8   0.537
V9          0.416
V10                -0.403
C11  0.624
I12         0.590
V13                 0.655
E14                         0.657
A15  0.597
V16                 0.643
V17
A18  0.694
A19 -0.444  0.658
I20         0.765
V21                 0.455
D22
I23         0.627
D24                                 0.437
V25                 0.639
C26  0.455
E27                         0.638
A28  0.758
I29         0.713
D30                                 0.687
C31 -0.445
C32 -0.522
V33                 0.663
D34                                 0.657
C35  0.727
D36

               Factor1 Factor2 Factor3 Factor4 Factor5
SS loadings      5.483   3.929   3.592   2.608   1.711
Proportion Var   0.152   0.109   0.100   0.072   0.048
Cumulative Var   0.152   0.261   0.361   0.434   0.481

Test of the hypothesis that 5 factors are sufficient.
The chi square statistic is 808.48 on 460 degrees of freedom.
The p-value is 1.56e-21

  \# ^
```

Matched Pairs T-test

```
################### Hypothesis Tests of pre-test vs. post-test Factor means

\# Condensing our post-test data to ROSA ID & survey results only
au1post.survey = cbind(AU1Post[,4], AU1Post[,6:41])
au2post.survey = cbind(Au2Post[,3], Au2Post[,5:40])
au3post.survey = cbind(Au3Post[,3], Au3Post[,5:40])
au4post.survey = cbind(Au4Post[,3], Au4Post[,5:40])

aupost = rbind(au1post.survey, au2post.survey, au3post.survey, au4post.survey)
\ #combining all of our post-test results into one master set "aupost"

\# combining results by student ID in order to make pre/post results matched pairs
prepost<-merge(aupre, aupost, by = "ROSA ID")

\# Pre-test answers are notated by capital letter then a number, i.e. A3 or C31
\# Post-test answers are notated by a lowercase p, i.e. pA3, or pC31. Our
\# Loadings matrix showed multiple statements as negatively loaded on certain
\# factors So when we construct our values to compare sample means, we need to
\# transform our results. Negatively loaded answers are inverted (transformed
\# based on a 7 point Likert scale) to reflect our factor construct. Transformed
\# answers are notated by uppercase J, i.e. if A3 = 3, then JA3 = 8 - A3 = 5

\# Transforming the answer of negatively loaded statements
JD6 = (8-D6)   \# Q D6: "Statistics Formulas are easy to understand"
JpD6 = (8-pD6) \# (Negatively Loaded on F1)

JC31 = (8-C31)
JpC31 = (8-pC31) \# Q C31: "I can learn statistics" (Negatively Loaded on F1)

JC32 = (8-C32)  \# Q C32: "I will understand statistics equations"
JpC32 = (8-C32) \# (Negatively Loaded on F1)

JV10 = (8-V10)   \# Q V10: "Statistical skills will make me more employable"
JpV10 = (8-pV10) \# (Negatively Loaded on F3)

attach(prepost) \# attaching matched pairs data set for t.test
F1Pre <- cbind(A4, C5, JD6, D8, C11, A15, A18, C26, A28, JC31, JC32, C35)
\# Making a dataset of only pre-test Factor 1 (F1) Responses
F2Pre <- cbind(A3, V9, I12, A19, I20, I23, I29)
\# Making a dataset of only pre-test Factor 2 (F2) Responses
F3Pre <- cbind(V7, JV10, V13, V16, V21, V25, V33)
\# Making a dataset of only pre-test Factor 3 (F3) Responses
F4Pre <- cbind(E1, E2, E14, E27)
\# Making a dataset of only pre-test Factor 4 (F4) Responses
F5Pre <- cbind(D24, D30, D34)
\# Making a dataset of only pre-test Factor 5 (F5) Responses

F1Post <- cbind(pA4, pC5, JpD6, pD8, pC11, pA15, pA18, pC26, pA28, JpC31, JpC32, pC35)
```

```
\# Making a dataset of only post-test Factor 1 (F1) Responses
F2Post <- cbind(pA3, pV9, pI12, pA19, pI20, pI23, pI29)
\# Making a dataset of only post-test Factor 2 (F2) Responses
F3Post <- cbind(pV7, JpV10, pV13, pV16, pV21, pV25, pV33)
\# Making a dataset of only post-test Factor 3 (F3) Responses
F4Post <- cbind(pE1, pE2, pE14, pE27)
\# Making a dataset of only post-test Factor 4 (F4) Responses
F5Post <- cbind(pD24, pD30, pD34)
\# Making a dataset of only post-test Factor 5 (F5) Responses

F1 <- (rowMeans(F1Pre)) \# Pre-test mean of answers of statements loaded on Factor 1
F2 <- (rowMeans(F2Pre)) \# Pre-test mean of answers of statements loaded on Factor 2
F3 <- (rowMeans(F3Pre)) \# Pre-test mean of answers of statements loaded on Factor 3
F4 <- (rowMeans(F4Pre)) \# Pre-test mean of answers of statements loaded on Factor 4
F5 <- (rowMeans(F5Pre)) \# Pre-test mean of answers of statements loaded on Factor 5

pF1 <- (rowMeans(F1Post)) \# Post-test mean of answers of statements loaded on Factor 1
pF2 <- (rowMeans(F2Post)) \# Post-test mean of answers of statements loaded on Factor 2
pF3 <- (rowMeans(F3Post)) \# Post-test mean of answers of statements loaded on Factor 3
pF4 <- (rowMeans(F4Post)) \# Post-test mean of answers of statements loaded on Factor 4
pF5 <- (rowMeans(F5Post)) \# Post-test mean of answers of statements loaded on Factor 5

## Factor 1 Hypothesis Test for Equal Means
> t.test(pF1 - F1, alternative = "two.sided")

One Sample t-test
data:  pF1 - F1
t = 1.3998, df = 156, p-value = 0.1636
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.05542443  0.32506349
sample estimates:
mean of x
0.1348195

## Factor 2 Hypothesis Test for Equal Means
> t.test(pF2 - F2, alternative = "two.sided")

One Sample t-test
data:  pF2 - F2
t = -9.6029, df = 157, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -1.2285789 -0.8093957
sample estimates:
mean of x
-1.018987
```

```
## Factor 3 Hypothesis Test for Equal Means
> t.test(pF3 - F3, alternative = "two.sided")

One Sample t-test
data:  pF3 - F3
t = 6.049, df = 165, p-value = 9.457e-09
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.3309987 0.6517896
sample estimates:
mean of x
0.4913941

## Factor 4 Hypothesis Test for Equal Means

> t.test(pF4 - F4, alternative = "two.sided")

One Sample t-test
data:  pF4 - F4
t = -7.0596, df = 170, p-value = 4.068e-11
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.7632819 -0.4297005
sample estimates:
 mean of x
-0.5964912

## Factor 5 Hypothesis Test for Equal Means
> t.test(pF5 - F5, alternative = "two.sided")

One Sample t-test
data:  pF5 - F5
t = -1.0171, df = 166, p-value = 0.3106
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.27005268  0.08641994
sample estimates:
  mean of x
-0.09181637

# \
^
```