The Rosenberg Self-Esteem Scale:

Its Dimensionality, Reliability, and Item Analyses

Lu Wang

1) Introduction

According to William James (1890), self-esteem is a construct capturing the positive level of a person towards themselves. More recently, it was defined as the affective-cognitive component of a person's self-system (Shapurian, Hojat, & Nayerahmadi, 1987). In order to test self-esteem, Rosenberg, M (1965) published The Rosenberg Self-Esteem Scale. It became the most popular scale in self-esteem evaluation. However, Researchers argue about the dimensionality of the scale whether it is unidimensional or bifactorial. This study will try to answer the following research questions:

Research question 1: how is the reliability of the Rosenberg Self-Esteem Scale (RSES), and how is he 10 items internal consistent?

Research question 2: Is the Rosenberg Self-Esteem Scale (RSES) unidimensional or Bifactorial? How many laten variables tested in RSES?

Research question 3: Are all the 10 items work well in the RSES? Do they discriminate people with positive and negative evaluations one themselves well? Is there any extreme items in the scale?

2) Method

Subjects:

The sample consisted of 45,943 participants 6 to 100 years old (M = 26.3, SD = 7.0). The data was collected through the survey from a public online interactive version of the Rosenberg Self-Esteem Scale. Most of the respondents were female (Female = 61.2%, Male = 37.3%, unidentified or others = 1.5%). The participants were informed that their data would be saved at the beginning of the survey. At the end of the survey, 75.7% of the participants consent that they had given accurate responses and the data could be used for research. The dataset only contains the those consented.

Instrument:

Rosenberg Self-esteem Scale (RSES)

The Rosenberg Self-esteem Scale (RSES) (Rosenberg, 1989) was designed to measure the level of self-esteem of high school students, and later on, it expended to groups of all ages. It is a 10 items Guttman scale with 5 positive wording items and 5 negative wording items. Participants rate their level of self-esteem by choosing from 5 attributes: 1=strongly disagree, 2=disagree, 3=agree, and 4=strongly agree. Choosing 3 or 4 for a positive worded item, like item 2 (I feel that I have a number of good qualities) shows a higher level of self-esteem of the participants. For a negative worded item like item 3 (All in all, I am inclined to feel that I am a failure), participants choose 1 or 2 to show their higher level of self-esteem. Item 1, 2, 4, 6, and 7 are positively worded, while item

3, 5, 8, 9, and 10 are negatively worded. The total score of the 10 items indicates a participant's general self-esteem level after reverse-scoring the negatively worded items. The total score range is from 10 to 40, with a lower score meaning less positive evaluation of oneself. A Cronbach's α of .91 was observed in this sample

Procedure:

The reliability of the RSES scale was estimated through Cronbach's alpha. The Cronbach's alpha uses the variances of the component scores and variance of the total test to estimate the reliability of the test. The item responses of this study are polytomous. The confidence interval based on the sample data estimates the possible range of alpha at the population level. Third, the researcher rotated the initial loadings to obtain a set of factor loadings that are interpretable. Finally, they interpreted the item-factor associations.

The dimension of the RSES scale was tested through Exploratory Factor Analysis (EFA) since the researcher was not sure if the scale is unidimensional or not. The researcher used EFA in the early phases while conducting basic exploratory analyses of a set of items to determine the number and nature of the underlying factors that the measure have in common. EFA compromises of four steps.

First, the researcher cleaned the data and reveres-coded the negative phrased items so that large value is related to the higher level of construct. After generating the inter item

correlation matrix, they applied the Factor Extraction. It is an iterative process wherein the first factor is extracted to account for the maximum amount of variance, and the subsequent factors are extracted from the residual correlation matrix, which factors out the influence of the previously extracted factors.

The Item Analysis in CTT was implemented in this study to evaluate it the items on the test are useful, or if there are any dysfunctional items. By testing the difficulty level and discrimination level, the researcher evaluated the 10 items in the RSES scale.

3) Result

3.1 Descriptive statistics

Total scores on the RSES ranged from 10 to 40, with a mean of 26.3 and the median is 26.0 (S.D.=7.0). The frequency distribution is almost normally distributed, and the shape of the bar chart (Graph 1) is like a bell. The males and females total scores is 27.2 and 25.7 respectively. Using t-test with gender (male and female) as a grouping variable showed significant differences, $t = 23.6$, $p < 0.05$, and the true difference in means is not equal to 0, while 95 percent confidence interval is from 1.45 to 1.72. The further test in this study will combine the two genders data together. More test on both genders will be discussed in the future research.

3.2 Reliability of the Scale

The Estimated Cronbach's α is 0.898. It means the internal consistency is perfect and the test has a high reliability. The confidence interval is (0.893, 0.902). We have 95% confidence that the Estimated Cronbach's α is in this range.

3.3 Factor analyses of the Scale

Following the Guttman-Kaiser criterion, eigenvalue greater than one, the number of PCs to retain is 5. The eigen value of 5 PCs are greater than 1 according to the Eigenvalues in Table 1. The proportion of variance accounted for (VAF) by re-running the PCA with 5 extracted PCs is 0.835. That means 83.5% of the variance in observed

data can the 5 extracted PCs explain.

The numeric values in the loading matrix in table 3 indicates correlation between the item and extracted component. The component loading of Question 4 (I am able to do things as well as most other people) on PC1 is 0.818 so that means the simple correlation between the Question 4 and the first component equals 0.818. Also, that means 66.9% of the variance in Question 4 is explained by PC1 so PC1 is a significant factor for Question 4.

The item communality is the sum of squared component loadings across all PCs for a given item. The communality of Question 8 (I wish I could have more respect for myself) is 0.998 meaning we can estimate that 99.8% of variance in Question 8 can be explained by the 5 PCs we selected, which is good.

From the scree plot in Graph 2, we concluded that only one factor retained after running the factor analysis the Maximum Likelihood (ML) estimation. The RSES scale is unidimensional, which means the scale only test the self-esteem.

3.4 Item analysis of the Scale

The difficulty level analysis in this study can be interpreted as the analysis of the positiveness a person is on his level of self-esteem since the higher the score is the more positive a participant is. The mean of the total test score for each individual is 26.295.

The minimum and maximum possible scores are 10.00 and 40.00 with a standard deviation of 6.985.

In general, the overall the participants showed little more positive on their self-esteem since the mean of the total scores (26.295) is slightly over the middle of the total score ((40-10)/2=25). In fact, more information about the test situation is needed to decide the positiveness of test since different difficulty level is required in different test situations.

Since the scale is not binary, the researcher got the p-levels by divided the scores by the total score, 4. If a participant get 3 in one question, his/her p-level on this question will be 0.75. According to Table 5, the range of the positiveness level (or difficulty level) of the items (i.e., p-levels) is from 0.553 to 0.777, and the average positiveness level is 0.657 with a standard deviation of 0.304. It means on average there is 65.7% possibility that test takers show rate their self-esteem level as positive.

The most positive item is Q2 (I feel that I have a number of good qualities) with p-level of 0.777 meaning 77.7% of the participants answer this item positively and the most negative item is Q9 (I certainly feel useless at times) with a p-level of 0.553 meaning only 55.3% of the participants feel useless at times. The average positive level is a little bit higher since it is slightly above 0.5 (0.643>0.5) meaning the scale is a little positive.

None of the items is problematic since the alphaIfDeleted values in Table 7 are between 0.9 and 0.91, which are smaller than or equal to Cronbach's alpha, 0.91. All the ten items contribute to the internal consistency.

We use biserial score (Table 8) to represent the discrimination level since this is not a binary scale. A high biserial reflects the fact that the item is doing a good job of discriminating positive participants from the negative ones. The range of the observed discrimination levels is from 0.583 to 0.805, and the average of the observed discrimination levels is 0.739 with a standard deviation of 0.073. The least discriminating item is Q8 (I wish I could have more respect for myself) with a biserial score of 0.583. All the questions work well to discriminate between the negative and positive participants.

In fact, an item with low or negative discrimination score will be a problematic item. If the score is low, it is really weak to tell the high-performing students and low-performing students apart. If the score is negative, it means basically only the low performing students can get this item right while the high-performing students will answer it wrong.

4) Conclusion or Findings.

Our aim has been to better understand the dimensionality, reliability and item quality of the Rosenberg Self-Esteem Scale (RSES). Before evaluating our conclusion, we must address several limitations in our data. Frist, the age of participants in this study ranges from 6 to 100. Although all-age evaluation is necessary for this general study of the scale, studies on different age groups are still needed. People's level of self-esteem may vary across age groups. Second, this study included both of the two genders. We know more female participated in this study than male (Female = 61.2%, Male = 37.3%, unidentified or others = 1.5%) and the average total scores for male and female were significantly different (male = 27.2, female = 25.7, $t = 23.6$, $p < 0.05$) across the gender levels. The further studies on both genders will be discussed in the future research. For example, if the item quality keeps the same across both genders. Third, RSES tests global self-esteem. We do not know if this scale has any contributions to other types of specific self-esteems, like academic self-esteem or sex self-esteem.

Despite the limitations, we believe that our findings are important:

4.1 Dimensionality of the Scale

We have evaluated the dimensionality of Rosenberg Self-Esteem Scale (RSES), and prove that it is unidimensional. It seems this scale measures only one laten variable and the other variable only account for only a small mount of variance of the outcome. it can be from the positive and negative wording techniques.

4.2 Internal Reliability

The internal consistency reliability was tested by Cronbach's alpha (0.91). The internal consistency is strong, meaning the 10 items in the scale agree with each other and each of them and contribute to the measurement of the construct.

4.3 Item quality

The 10 items in the Rosenberg Self-Esteem Scale (RSES) works well in differentiate or discriminate people with high level of self-esteem and low level of self-esteem. The total score of the measurement is normally distributed, meaning the scale evenly located people with different level of self-esteem around the mean. Also, all the 10 items work well to both for negative and positive people. No item is only for people have negative attitude towards themselves or positive attitude.

Appendix

Rosenberg Self-Esteem Scale (Rosenberg, M. (1965))

1=strongly disagree, 2=disagree, 3=agree, and 4=strongly

Q1. I feel that I am a person of worth, at least on an equal plane with others.
Q2. I feel that I have a number of good qualities.
Q3. All in all, I am inclined to feel that I am a failure.
Q4. I am able to do things as well as most other people.
Q5. I feel I do not have much to be proud of.
Q6. I take a positive attitude toward myself.
Q7. On the whole, I am satisfied with myself.
Q8. I wish I could have more respect for myself.
Q9. I certainly feel useless at times.
Q10. At times I think I am no good at all.

Graph 1

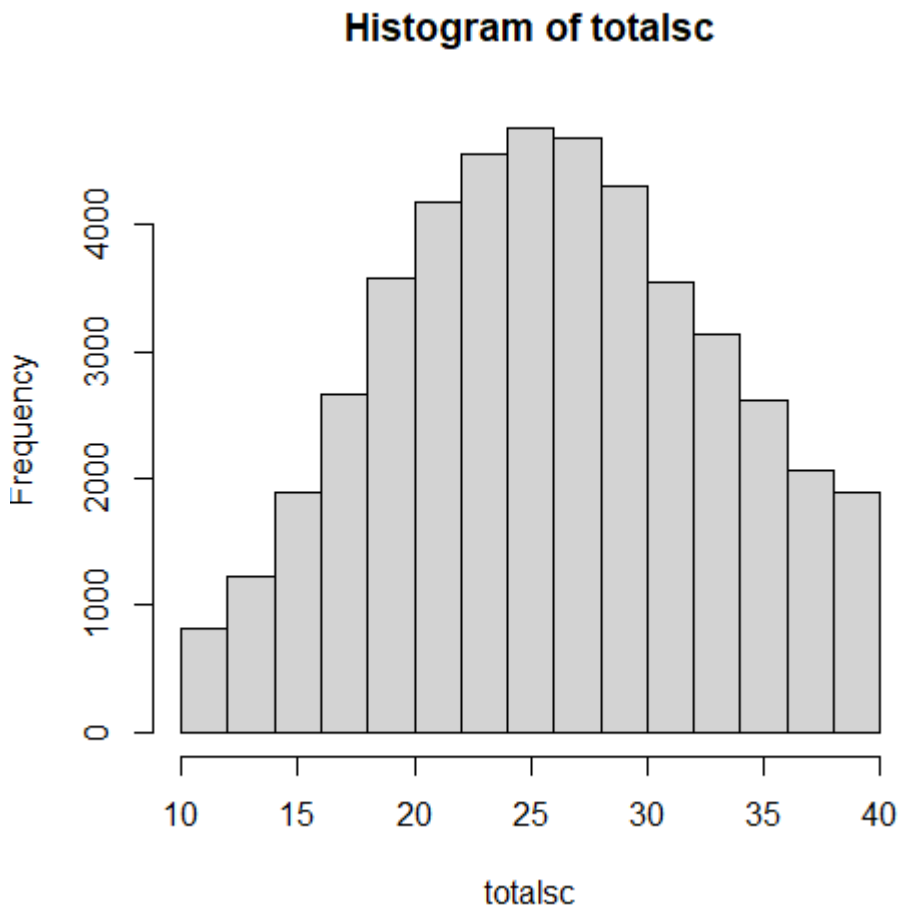

Table 1 Eigenvalues with 10 Components

	RC3	RC6	RC1	RC5	RC2	RC4	RC7	RC8	RC9	RC10
SS loadings	**1.103**	**1.092**	**1.068**	**1.037**	**1.028**	0.977	0.973	0.956	0.904	0.862
Proportion Var	0.110	0.109	0.107	0.104	0.103	0.098	0.097	0.096	0.090	0.086
Cumulative Var	0.110	0.220	0.326	0.430	0.533	0.631	0.728	0.823	0.914	1.000
Proportion Explained	0.110	0.109	0.107	0.104	0.103	0.098	0.097	0.096	0.090	0.086
Cumulative Proportion	0.110	0.220	0.326	0.430	0.533	0.631	0.728	0.823	0.914	1.000

Table 2 VAF with 5 Components

	RC2	RC4	RC1	RC5	RC3
SS loadings	2.205	1.822	1.765	1.509	1.047
Proportion Var	0.220	0.182	0.176	0.151	0.105
Cumulative Var	0.220	0.403	0.579	0.730	**0.835**
Proportion Explained	0.264	0.218	0.211	0.181	0.125
Cumulative Proportion	0.264	0.482	0.694	0.875	1.000

Table 3 Loading Matrix with 5 Components

	RC2	RC4	RC1	RC5	RC3
Q1	0.714	0.149	0.344	0.276	0.149
Q2	0.760	0.112	0.274	0.305	0.085
Q3	0.256	0.390	0.346	0.624	0.132
Q4	0.818	0.269	0.131	0.060	0.073
Q5	0.260	0.258	0.211	0.827	0.146
Q6	0.339	0.260	0.774	0.208	0.181
Q7	0.278	0.240	0.799	0.258	0.149
Q8	0.145	0.250	0.200	0.160	0.922
Q9	0.201	0.846	0.200	0.227	0.182
Q10	0.250	0.770	0.270	0.276	0.192

Table 4 Item Communality with 5 Component

Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
0.749	0.765	0.744	0.768	0.883	0.857	0.862	0.998	0.881	0.841

Graph 2 Scree Plot with 10 Components

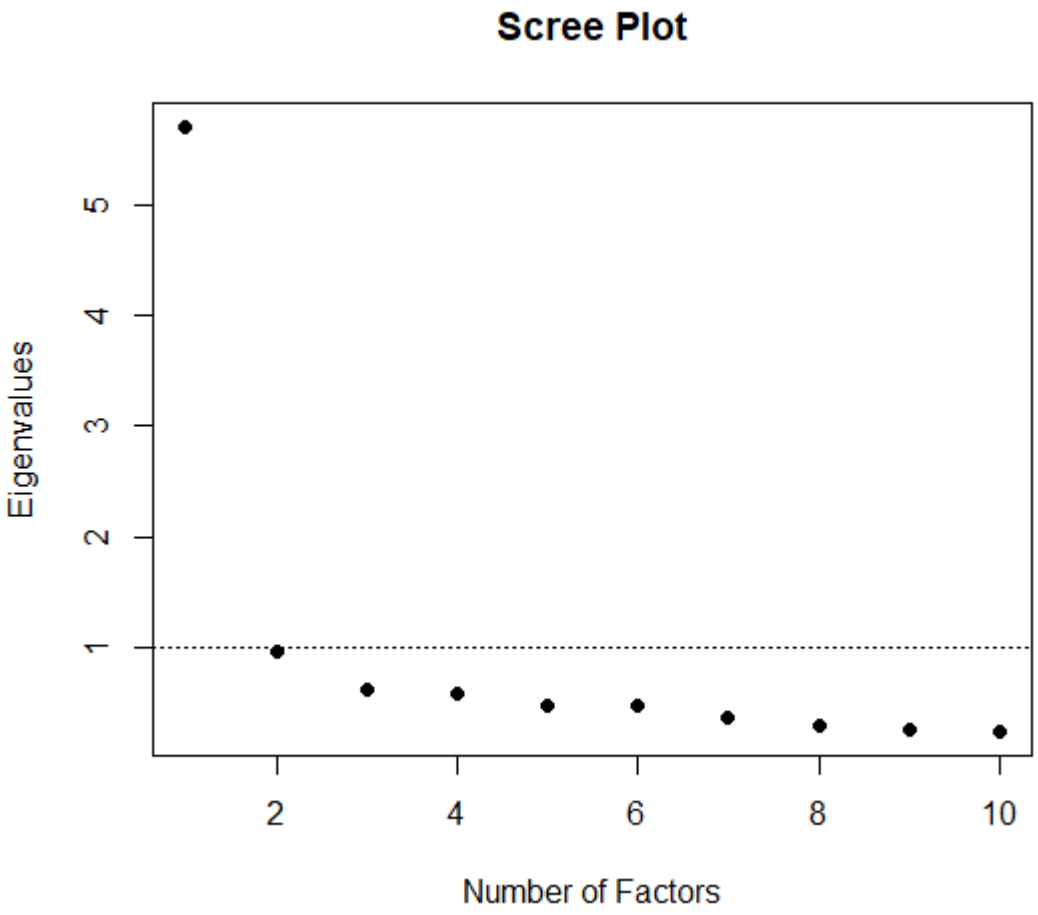

Table 5 Difficulty Level

Question	Q9	Q8	Q10	Q7	Q6	Q5	Q3	Q4	Q1	Q2
P-value	0.553	0.576	0.604	0.612	0.641	0.654	0.673	0.731	0.753	0.777

Table 7 Alpha If Delete

Question	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
alphaIfDeleted	0.905	0.907	0.902	0.911	0.905	0.901	0.902	0.914	0.905	0.902

Table 8 biserial score

Question	**8**	**4**	**2**	**5**	**9**	**1**	**7**	**3**	**10**	**6**
Bis	0.583	0.639	0.738	0.741	0.749	0.766	0.784	0.785	0.802	0.805

R Code:

```
#research question

#1. reliability (internal consistency; croback alpha)

#2. item analysis

#3. Is the sacle unidimenstional (FA)

library(psych)

library(GPArotation)

# Basic statistical analysis of data at the test level (e.g., distribution of total score


# Descriptive Statistics


dim(data_1)

# [1] observations = 45943; items = 15

attach(data_1)

mean(totalsc)

# Mean = 26.29495

median(totalsc)

# Median = 26

sd(totalsc)

# Standard Deviation = 6.984597

summary(totalsc)

#      Min. 1st Qu.    Median        Mean 3rd Qu.        Max.
```

```
# 10.00      21.00      26.00      26.29      31.00      40.00

hist(totalsc)


mean(age)

# 26.61554

sd(age)

# 12.3927

hist(age)

table(gender)

#         N         M         F         O

#       166 17160 28121       497

prop.table(table(gender))

#                   0                   1                   2                   3

#    0.003591407 0.373506301 0.612084540 0.010817752


tapply(totalsc, gender, mean)

# Mean of the male = 26.12121 the Female = 25.7


t.test(data_3$totalsc ~ data_3$gender,var.equal=TRUE)

# Using t-test with gender (male and female) as a grouping variable showed significant

differences,

# t = 23.6, p < 0.05, and the true difference in means is not equal to 0,
```

```
# while 95 percent confidence interval is from 1.45 to 1.72.


### Factor Analysis ###


cormat <- cor(data_1[,1:10])

pca_full <- principal(r = cormat, nfactors = 10)

pca_full$loadings

# SS loadings          1.103 1.092 1.068 1.037 1.028 0.977 0.973 0.956 0.904 0.862

# 5 factors are identified


write.csv(pca_full$Vaccounted, file="pca_full_loadings.csv")


pca_full$Vaccounted


pca_full$communality

pca_ide <- psych::principal(r = cormat, nfactors = 5)


pca_ide$Vaccounted

write.csv(pca_ide$Vaccounted, file="pca_ide_Vaccounted.csv")

pca_ide$loadings

write.csv(pca_ide$loadings, file="pca_ide_loadings.csv")
```

```
pca_ide$communality

write.csv(pca_ide$communality, file="pca_ide_communality.csv")



fa_init <- psych::fa(r = cormat, nfactors = 5)


fa_init_ml <- fa(r=cormat, nfactors=5, SMC=TRUE, fm="ml")

fa_init_ml$loadings



plot(seq(1,10,1), fa_init_ml$e.values, pch = 16,

     xlab="Number of Factors", ylab="Eigenvalues", main="Scree Plot")

abline(1, 0, lty="dotted")



# 1 factor


#### CTT Item Analysis ###


# Difficulty
```

```
datai <- data_1[,1:10]

colMeans(datai)

#        Q1        Q2        Q3        Q4        Q5        Q6        Q7

Q8        Q9        Q10

# 3.012167 3.109745 2.692336 2.923688 2.614871 2.563720 2.446009 2.303920

2.210652 2.417844




apply(datai,2, var )    # 2 indicates columns

#        Q1        Q2        Q3        Q4        Q5        Q6

Q7        Q8        Q9        Q10

# 0.7462061 0.6183159 0.9098466 0.6478266 0.9547586 0.8540161 0.8737065

0.9137905 0.9784795 1.1495723


apply(datai,2, sd )    # 2 indicates columns

#        Q1        Q2        Q3        Q4        Q5        Q6

Q7        Q8        Q9        Q10

# 0.8638322 0.7863306 0.9538588 0.8048768 0.9771175 0.9241299 0.9347227

0.9559239 0.9891812 1.0721811


# Item difficulty indices and the average difficulty
```

```
diff <- colMeans(datai)

mean(diff/4)

# 0.657


sd(diff)


# Identify the easies and hardest item

sort(diff/4)

# easies: Q9 = 2.200946;    hardest: Q2 = 3.088178


write.csv(sort(diff/4), file="sort_diff.csv")




result_sort_diff <- sort.int(diff, index.return = TRUE, decreasing=TRUE)

result_sort_diff$x # p-levels of items in descending order

result_sort_diff$ix # item index whose p-levels are in descending order


## barplot

barplot(diff, main="Difficulty")
```

CTT Item Analysis - Item Discrimisniation

```
dataif <- as.data.frame(datai)

# data input must be data.frame; if not transform using "as.data.frame"

result_ctt <- CTT::itemAnalysis(dataif, itemReport=TRUE)

str(result_ctt)

result_ctt$itemReport$alphaIfDeleted

write.csv((result_ctt$itemReport$alphaIfDeleted), file="alphaIfDeleted.csv")



# nItem: The number of items

# nPerson: The sample size used in calculating the values

# alpha: Cronbach's alpha

# scaleMean: Average of total score

# scaleSD: Standard deviation of total sum score

# alphaIfDeleted: Cronbach's alpha if the corresponding item were deleted

# pBis: The item total correlation, with the item's contribution removed from the total

(CORRECTED) (for true dichotomy)

# bis: The item total biserial correlation, with the item's contribution removed

# from the total (CORRECTED)

# itemMean: Average of each item
```

```
bis <- result_ctt$itemReport$bis # item biserial correlations

mean(bis)

tmp <- sort.int(bis, index.return = TRUE)

tmp$x

tmp$ix

sd(bis)


write.csv(tmp, file="bis.csv")



# mean bis = 0.7268325

#  0.5784572  0.6268226  0.7041272  0.7324152  0.7350403  0.7609913  0.7697587

0.7729052 0.7920146 0.7957924

# 8   4   2   5   9   1   7   3 10   6

# 0.578 - 0.796

barplot(bis)

names(bis) <- c("Item 1", "Item 2", "Item 3", "Item 4", "Item 5","Item 6", "Item 7",

"Item 8", "Item 9", "Item 10")

barplot(bis, main="Biserial")
```

```
ItemStats <- cbind(diff,bis)

colnames(ItemStats)<-c("Difficulty", "Corrected rbis")

round(ItemStats, 2)

write.csv(ItemStats, file="CTT_ItemAnalysis.csv")


###reliablility - Internal Consistency Measures - Cronbach's Alpha ###


rel_alpha <- psychometric::alpha(datai)

# 0.9109343


### Confidence interval for alpha

psychometric::alpha.CI(alpha = rel_alpha, k = 10, N = 47336, level = 0.95)
```

Citations:

Shapurian, R., Hojat, M., & Nayerahmadi, H. (1987). Psychometric characteristics and dimensionality of a Persian version of Rosenberg Self-esteem Scale. *Perceptual and Motor Skills*, 65(1), 27–34. Retrieved from https://doi.org/10.2466/pms.1987.65.1.27

James, W. (1890). The principles of psychology, by William James ... Henry Holt and company.

Brown, J. (1998). The Self (1st ed.). Psychology Press. https://doi-org.ezproxy.lib.utexas.edu/10.4324/9780203837641

Rosenberg, M. (1965). Society and the adolescent self-image. Princeton University Press.