Technical Report

```
EXAMINE VARIABLES=Explanation_A Explanation_B
Evidence_A Evidence_B Context_A Context_B
Position_A
    Position_B Conclusion_A Conclusion_B BY
Sex Generation Class
  /PLOT BOXPLOT STEMLEAF
  /COMPARE GROUPS
  /STATISTICS DESCRIPTIVES
  /CINTERVAL 95
  /MISSING LISTWISE
  /NOTOTAL.
```

Sex

Case Processing Summary

		Cases					
		Valid		Missing		Total	
	Sex	N	Percent	N	Percent	N	Percent
Explanation_A	F	26	100.0%	0	0.0%	26	100.0%
	M	24	100.0%	0	0.0%	24	100.0%
Explanation_B	F	26	100.0%	0	0.0%	26	100.0%
	M	24	100.0%	0	0.0%	24	100.0%
Evidence_A	F	26	100.0%	0	0.0%	26	100.0%
	M	24	100.0%	0	0.0%	24	100.0%
Evidence_B	F	26	100.0%	0	0.0%	26	100.0%
	M	24	100.0%	0	0.0%	24	100.0%
Context_A	F	26	100.0%	0	0.0%	26	100.0%
	M	24	100.0%	0	0.0%	24	100.0%
Context_B	F	26	100.0%	0	0.0%	26	100.0%
	M	24	100.0%	0	0.0%	24	100.0%
Position_A	F	26	100.0%	0	0.0%	26	100.0%
	M	24	100.0%	0	0.0%	24	100.0%
Position_B	F	26	100.0%	0	0.0%	26	100.0%
	M	24	100.0%	0	0.0%	24	100.0%
Conclusion_A	F	26	100.0%	0	0.0%	26	100.0%
	M	24	100.0%	0	0.0%	24	100.0%
Conclusion_B	F	26	100.0%	0	0.0%	26	100.0%
	M	24	100.0%	0	0.0%	24	100.0%

Generation

Case Processing Summary

		Cases					
		Valid		Missing		Total	
	Generation	N	Percent	N	Percent	N	Percent
Explanation_A	0	29	100.0%	0	0.0%	29	100.0%
	1	21	100.0%	0	0.0%	21	100.0%
Explanation_B	0	29	100.0%	0	0.0%	29	100.0%
	1	21	100.0%	0	0.0%	21	100.0%
Evidence_A	0	29	100.0%	0	0.0%	29	100.0%
	1	21	100.0%	0	0.0%	21	100.0%
Evidence_B	0	29	100.0%	0	0.0%	29	100.0%
	1	21	100.0%	0	0.0%	21	100.0%
Context_A	0	29	100.0%	0	0.0%	29	100.0%
	1	21	100.0%	0	0.0%	21	100.0%
Context_B	0	29	100.0%	0	0.0%	29	100.0%
	1	21	100.0%	0	0.0%	21	100.0%
Position_A	0	29	100.0%	0	0.0%	29	100.0%
	1	21	100.0%	0	0.0%	21	100.0%
Position_B	0	29	100.0%	0	0.0%	29	100.0%
	1	21	100.0%	0	0.0%	21	100.0%
Conclusion_A	0	29	100.0%	0	0.0%	29	100.0%
	1	21	100.0%	0	0.0%	21	100.0%
Conclusion_B	0	29	100.0%	0	0.0%	29	100.0%
	1	21	100.0%	0	0.0%	21	100.0%

Class

Case Processing Summary

		Cases					
		Valid		Missing		Total	
	Class	N	Percent	N	Percent	N	Perc
Explanation_A	Complete	25	100.0%	0	0.0%	25	100.
	First	25	100.0%	0	0.0%	25	100.
Explanation_B	Complete	25	100.0%	0	0.0%	25	100.
	First	25	100.0%	0	0.0%	25	100.
Evidence_A	Complete	25	100.0%	0	0.0%	25	100.
	First	25	100.0%	0	0.0%	25	100.
Evidence_B	Complete	25	100.0%	0	0.0%	25	100.
	First	25	100.0%	0	0.0%	25	100.
Context_A	Complete	25	100.0%	0	0.0%	25	100.
	First	25	100.0%	0	0.0%	25	100.
Context_B	Complete	25	100.0%	0	0.0%	25	100.
	First	25	100.0%	0	0.0%	25	100.
Position_A	Complete	25	100.0%	0	0.0%	25	100.
	First	25	100.0%	0	0.0%	25	100.
Position_B	Complete	25	100.0%	0	0.0%	25	100.
	First	25	100.0%	0	0.0%	25	100.
Conclusion_A	Complete	25	100.0%	0	0.0%	25	100.
	First	25	100.0%	0	0.0%	25	100.
Conclusion_B	Complete	25	100.0%	0	0.0%	25	100.
	First	25	100.0%	0	0.0%	25	100.

James Madison University
Performance Assessment
Assignment
3/8/2022
John W. Lee, M. S.

Research Questions

Proposed Research Questions:

1. What is the reliability information associated with the value rubrics being utilized?

2. What differences exists between first year students and fourth year students?

3. Did fourth year students achieve a 3 or higher?

4. What are the differences between males and females, and 1st generation vs non-1st gen students on performance?

Are there differences between males and females on performance? Are there differences between 1st generation and non-1st generation student performance in critical thinking?

Approach

Statistical Software used: SPSS

Version: 27

The research questions posed above call for a mix of analytic methods in order to determine what models best fit the data.

The first question asks simply, what reliability evidence exists for the measure. Here, the Critical Thinking Value Rubric, published by the Association of American Colleges and Universities is utilized. This rubric contains five dimensions or criteria (Explanation of issues, Evidence, Influence of Context and Assumptions, Student's Position, and Conclusions and Related Outcomes. Two faculty raters provided ratings (1 through 4) on each dimension, for each of the 50 students.

The second question asks what differences exists between first year students and fourth year students.

The third question asks whether or not fourth year students achieved a score of three or above.

The fourth questions asks if there are differences on performance in males and females, and 1st generation students and non-1st generation students.

The following output and associated syntax follows the approach described above, with key elements identified with → an arrow, and annotations included in ▭ text boxes.

Descriptive Statistics Of Raw Data
SPSS Syntax:

```
DESCRIPTIVES
VARIABLES=Explanation_A
Explanation_B Evidence_A
Evidence_B Context_A Context_B
    Position_A Position_B
Conclusion_A Conclusion_B
  /STATISTICS=MEAN STDDEV MIN MAX.
```

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Explanation_A	50	1.00	4.00	2.2000	1.01015
Explanation_B	50	1.00	4.00	2.2000	1.10657
Evidence_A	50	1.00	4.00	2.3400	1.08063
Evidence_B	50	1.00	4.00	2.4200	1.05153
Context_A	50	1.00	4.00	2.3800	1.04764
Context_B	50	1.00	4.00	2.5600	1.07210
Position_A	50	1.00	4.00	2.1600	.97646
Position_B	50	1.00	4.00	2.0800	.98644
Conclusion_A	50	1.00	4.00	2.2000	.98974
Conclusion_B	50	.00	4.00	2.4000	1.10657
Valid N (listwise)	50				

Frequencies
SPSS Syntax:

```
FREQUENCIES VARIABLES=Sex Generation
Class
  /ORDER=ANALYSIS.
```

Sex

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	F	26	52.0	52.0	52.0
	M	24	48.0	48.0	100.0
	Total	50	100.0	100.0	

Generation

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	29	58.0	58.0	58.0
	1	21	42.0	42.0	100.0
	Total	50	100.0	100.0	

Class

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Complete	25	50.0	50.0	50.0
	First	25	50.0	50.0	100.0
	Total	50	100.0	100.0	

Transformation
SPSS Syntax:

Computing new variables to obtain average score for each student in each one of the five dimensions. Also computing total score for each student based on averaged scores between raters.

```
DATASET ACTIVATE DataSet1.
COMPUTE Exp_avg=(Explanation_A + Explanation_B) / 2.
COMPUTE Evi_avg=(Evidence_A + Evidence_B)/2.
COMPUTE Cont_avg=(Context_A + Context_B)/2.
COMPUTE Pos_avg=(Position_A + Position_B)/2.
COMPUTE Conc_avg=(Conclusion_A + Conclusion_B)/2.
COMPUTE Totscore_avg = (Exp_avg + Evi_avg + Cont_avg +
Pos_avg + Conc_avg)/5.
EXECUTE.
```

Descriptive statistics of Transformed Data.

```
DESCRIPTIVES VARIABLES=Exp_avg
Evi_avg Cont_avg Pos_avg
Conc_avg Totscore_avg
  /STATISTICS=MEAN STDDEV MIN
MAX KURTOSIS SKEWNESS.
```

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Exp_avg	50	1.00	4.00	2.2000	.96890	.316	.337	-1.175	.662
Evi_avg	50	1.00	4.00	2.3800	.97185	.179	.337	-.957	.662
Cont_avg	50	1.00	4.00	2.4700	.99699	.070	.337	-1.102	.662
Pos_avg	50	1.00	4.00	2.1200	.93438	.613	.337	-.458	.662
Conc_avg	50	.50	4.00	2.3000	.94221	.019	.337	-1.021	.662
Totscore_avg	50	1.30	3.40	2.2940	.58324	.082	.337	-1.141	.662
Valid N (listwise)	50								

```
CORRELATIONS
  /VARIABLES=Exp_avg Evi_avg
Cont_avg Pos_avg Conc_avg
  /PRINT=TWOTAIL NOSIG LOWER
  /MISSING=PAIRWISE.
```

Correlations

		Exp_avg	Evi_avg	Cont_avg	Pos_avg	Conc_avg
Exp_avg	Pearson Correlation	--				
	N	50				
Evi_avg	Pearson Correlation	.362**	--			
	Sig. (2-tailed)	.010				
	N	50	50			
Cont_avg	Pearson Correlation	.329*	.065			
	Sig. (2-tailed)	.020	.655			
	N	50	50	50		
Pos_avg	Pearson Correlation	.187	.112	.070	--	
	Sig. (2-tailed)	.193	.440	.631		
	N	50	50	50	50	
Conc_avg	Pearson Correlation	.296*	.308*	.298*	.051	--
	Sig. (2-tailed)	.037	.030	.036	.725	
	N	50	50	50	50	50

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

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

Here we can see that a number of the computed criterion variables have a weak correlation with each other. This could foreshadow a reduction in the internal consistency of the measure.

Question #1: Reliability
SPSS Syntax:
Internal Consistency

```
RELIABILITY
  /VARIABLES=Exp_avg Evi_avg Cont_avg
Pos_avg Conc_avg
  /SCALE('ALL VARIABLES') ALL
  /MODEL=ALPHA
  /STATISTICS=DESCRIPTIVE SCALE CORR COV
ANOVA
  /SUMMARY=MEANS VARIANCE COV CORR.
```

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.568	.567	5

Internal Consistency is reflected by a coefficient alpha of .568. This is in line with the correlations alluded to above, and also indicates a moderately low level of internal consistency among the items on the measure of Critical Thinking.

Inter-Item Correlation Matrix

	Exp_avg	Evi_avg	Cont_avg	Pos_avg	Conc_avg
Exp_avg	1.000	.362	.329	.187	.296
Evi_avg	.362	1.000	.065	.112	.308
Cont_avg	.329	.065	1.000	.070	.298
Pos_avg	.187	.112	.070	1.000	.051
Conc_avg	.296	.308	.298	.051	1.000

Inter-Item Covariance Matrix

	Exp_avg	Evi_avg	Cont_avg	Pos_avg	Conc_avg
Exp_avg	.939	.341	.317	.169	.270
Evi_avg	.341	.944	.063	.101	.282
Cont_avg	.317	.063	.994	.065	.280
Pos_avg	.169	.101	.065	.873	.045
Conc_avg	.270	.282	.280	.045	.888

Summary Item Statistics

	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Item Means	2.294	2.120	2.470	.350	1.165	.019	5
Item Variances	.928	.873	.994	.121	1.138	.002	5
Inter-Item Covariances	.193	.045	.341	.296	7.591	.013	5
Inter-Item Correlations	.208	.051	.362	.311	7.097	.015	5

Item variances are high, which contributes to an increase in coefficient alpha. Inter-item correlations are low, which can reduce to strength of coefficient alpha. This could indicate that the different criterion being summed into the total score (Explanation, Evidence, Context, Position, Conclusion) are not strongly related to the construct being measured. This was foreshadowed in our correlation analysis on page 5.

Question #1: Reliability

SPSS Syntax:

Interrater Reliability

Rater A and B on Explanation

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.804	.806	2

```
RELIABILITY
  /VARIABLES=Explanation_A Explanation_B
  /SCALE('ALL VARIABLES') ALL
  /MODEL=ALPHA
  /STATISTICS=DESCRIPTIVE SCALE CORR COV
  /SUMMARY=MEANS VARIANCE COV CORR
  /ICC=MODEL(MIXED) TYPE(CONSISTENCY) CIN=95
TESTVAL=0.
```

Rater A and B on Evidence

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.796	.797	2

```
RELIABILITY
  /VARIABLES=Evidence_A Evidence_B
  /SCALE('ALL VARIABLES') ALL
  /MODEL=ALPHA
  /STATISTICS=DESCRIPTIVE SCALE CORR COV
  /SUMMARY=MEANS VARIANCE COV CORR
  /ICC=MODEL(MIXED) TYPE(CONSISTENCY) CIN=95
TESTVAL=0.
```

Rater A and B on Context

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.870	.870	2

```
RELIABILITY
  /VARIABLES=Context_A Context_B
  /SCALE('ALL VARIABLES') ALL
  /MODEL=ALPHA
  /STATISTICS=DESCRIPTIVE SCALE CORR COV
  /SUMMARY=MEANS VARIANCE COV CORR
  /ICC=MODEL(MIXED) TYPE(CONSISTENCY) CIN=95
TESTVAL=0.
```

Rater A and B on Position

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.897	.897	2

```
RELIABILITY
  /VARIABLES=Position_A Position_B
  /SCALE('ALL VARIABLES') ALL
  /MODEL=ALPHA
  /STATISTICS=DESCRIPTIVE SCALE CORR COV
  /SUMMARY=MEANS VARIANCE COV CORR
  /ICC=MODEL(MIXED) TYPE(CONSISTENCY) CIN=95
TESTVAL=0.
```

Rater A and B on Conclusion

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.759	.762	2

```
RELIABILITY
  /VARIABLES=Conclusion_A Conclusion_B
  /SCALE('ALL VARIABLES') ALL
  /MODEL=ALPHA
  /STATISTICS=DESCRIPTIVE SCALE CORR COV
  /SUMMARY=MEANS VARIANCE COV CORR
  /ICC=MODEL(MIXED) TYPE(CONSISTENCY) CIN=95
TESTVAL=0.
```

Interrater agreement as calculated above is satisfactory for each one of the criteria. This indicates that the low reliability reported on page 6 is not likely due to interrater agreement.

Question #2: Analysis of Covariance
Part 1: Descriptive Statistics
SPSS Syntax:

```
MEANS TABLES=dv cov BY Group
  /CELLS=MEAN COUNT STDDEV.
```

Report

Group		Achievement	meta cognitive
non traditional females	Mean	50.9600	100.9200
	N	25	25
	Std. Deviation	3.69098	6.24446
traditional male	Mean	48.8800	101.7200
	N	25	25
	Std. Deviation	3.34564	6.38697
traditional female	Mean	49.6000	102.8000
	N	25	25
	Std. Deviation	3.67423	6.39010
Total	Mean	49.8133	101.8133
	N	75	75
	Std. Deviation	3.63055	6.30249

These descriptive stats provide the means for each one of the levels of the categorical variable on achievement and meta cognitive ability. Notice, these values differ.

Non traditional females have the highest mean on achievement, and the lowest mean on meta cog. Traditional females have a mean between the other two on achievement, and the highest meta cog mean. Also notice here, the differences between the meta cognitive means.

Of note regarding these descriptive is the relationship between the categorical IV (student type) and the continuous IV (meta cognitive). We can see that a change in student type, does not appear to have a very large impact on the level of meta cognitive ability. This indicates that the correlation between student type and meta cognitive ability is low. This is a good property of this model

Question #2: Differences between first year and fourth year students
Part 2: Independent samples T-test
Assumptions
SPSS Syntax:

```
EXAMINE VARIABLES=Totscore_avg
BY Class
  /PLOT BOXPLOT HISTOGRAM
NPPLOT SPREADLEVEL(1)
  /COMPARE GROUPS
  /STATISTICS DESCRIPTIVES
  /CINTERVAL 95
  /MISSING LISTWISE
  /NOTOTAL.
```

Descriptives

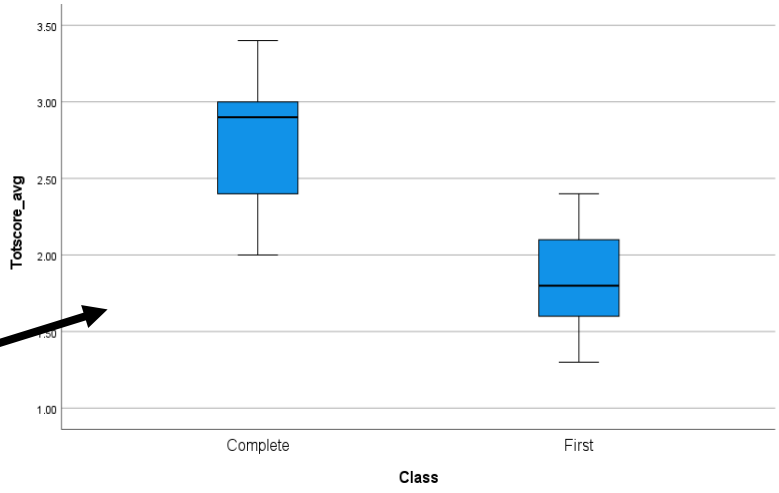
	Class			Statistic	Std. Error
Totscore_avg	Complete	Mean		2.7640	.07208
		95% Confidence Interval for Mean	Lower Bound	2.6152	
			Upper Bound	2.9128	
		5% Trimmed Mean		2.7689	
		Median		2.9000	
		Variance		.130	
		Std. Deviation		.36042	
		Minimum		2.00	
		Maximum		3.40	
		Range		1.40	
		Interquartile Range		.60	
		Skewness		-.285	.464
		Kurtosis		-.710	.902
	First	Mean		1.8240	.06462
		95% Confidence Interval for Mean	Lower Bound	1.6906	
			Upper Bound	1.9574	
		5% Trimmed Mean		1.8211	
		Median		1.8000	
		Variance		.104	
		Std. Deviation		.32311	
		Minimum		1.30	
		Maximum		2.40	
		Range		1.10	
		Interquartile Range		.50	
		Skewness		.265	.464
		Kurtosis		-1.093	.902

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Class	Statistic	df	Sig.	Statistic	df	Sig.
Totscore_avg	Complete	.167	25	.070	.953	25	.288
	First	.156	25	.119	.946	25	.200

a. Lilliefors Significance Correction

Here we can see that the test of normality for both the first year students, and the students soon to be complete, are normally distributed. Meeting our assumption of normality to an independent samples t-test.



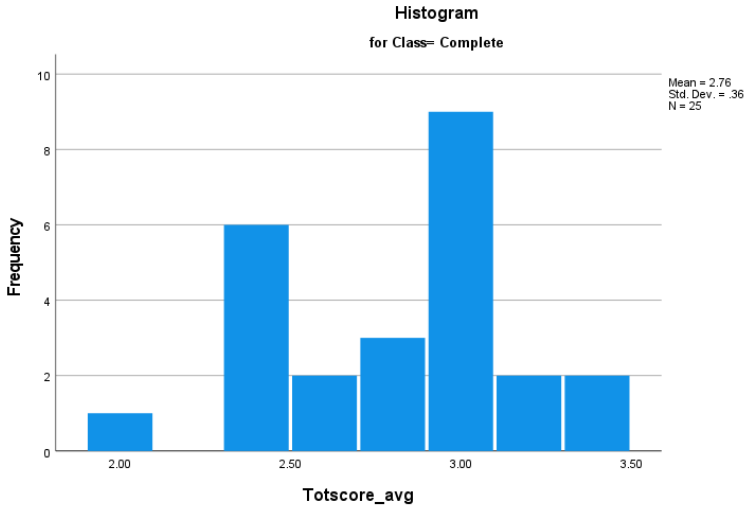
There are no outliers in the data set, and the means appear to be different based on visual inspection of box plot.

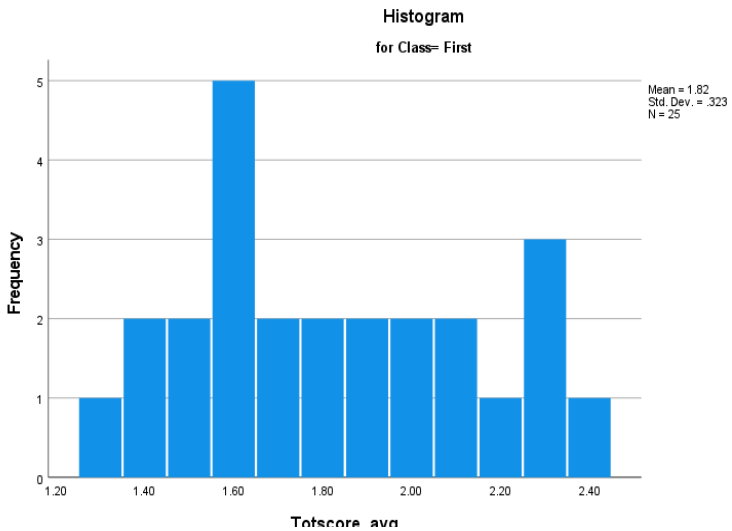
Test of Homogeneity of Variance

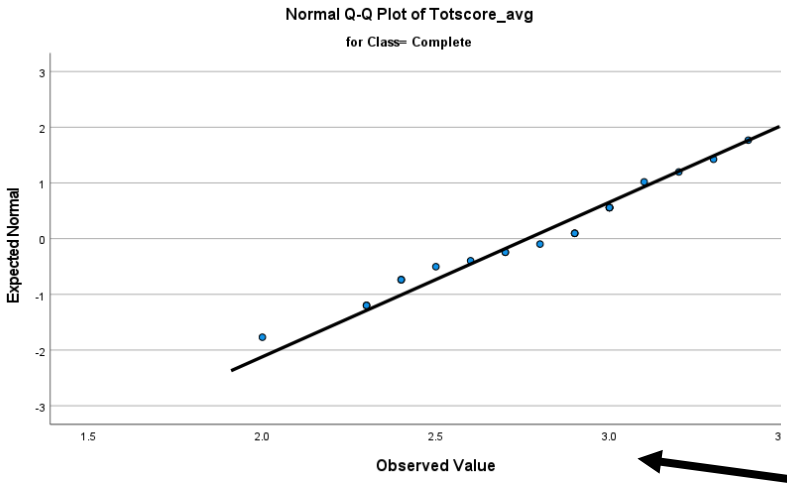
		Levene Statistic	df1	df2	Sig.
Totscore_avg	Based on Mean	.366	1	48	.548
	Based on Median	.168	1	48	.684
	Based on Median and with adjusted df	.168	1	42.785	.684
	Based on trimmed mean	.357	1	48	.553

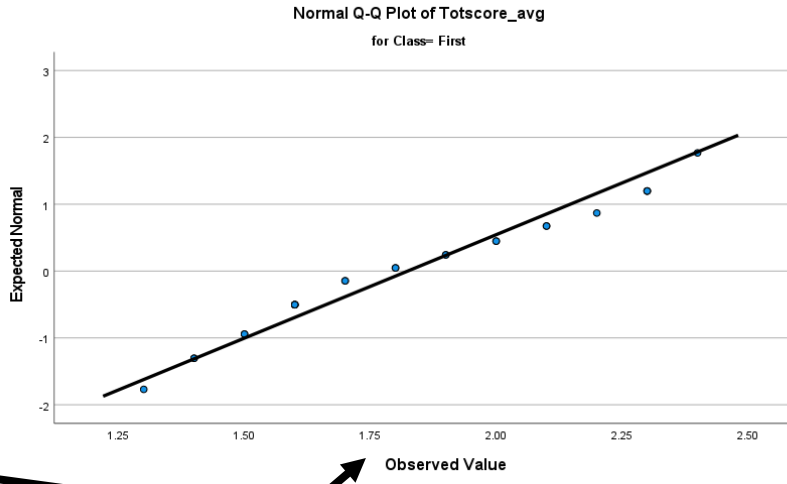
Test of homogeneity of variance is not significant, indicating the assumption is satisfied.

Assumptions continuted:

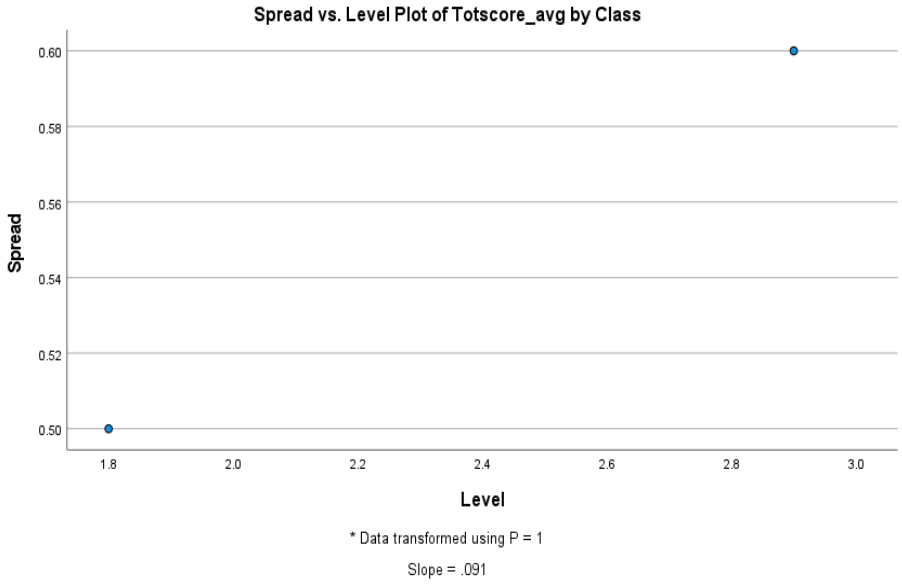








Based on the Histograms, and Q-Q plots, we can see that there is a linearity for both first year students and students soon to complete the program.



The spread vs level plot displays the mean and standard deviation. No relationship is visible between means and variances for the groups.

Part 2: Independent samples T-test

SPSS Syntax:

```
T-TEST
GROUPS=Class('First'
'Complete')
  /MISSING=ANALYSIS
  /VARIABLES=Totscore_avg
  /ES DISPLAY(TRUE)
  /CRITERIA=CI(.95).
```

Group Statistics

	Class	N	Mean	Std. Deviation	Std. Error Mean
Totscore_avg	First	25	1.8240	.32311	.06462
	Complete	25	2.7640	.36042	.07208

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Totscore_avg	Equal variances assumed	.366	.548	-9.710	48	.000	-.94000	.09681	-1.13465	-.74535
	Equal variances not assumed			-9.710	47.438	.000	-.94000	.09681	-1.13471	-.74529

Because our assumptions for Independent Samples T test are met (above analysis), we see that the test is significant at $p<.001$.

Independent Samples Effect Sizes

		Standardizer[a]	Point Estimate	95% Confidence Interval	
				Lower	Upper
Totscore_avg	Cohen's d	.34227	-2.746	-3.518	-1.959
	Hedges' correction	.34774	-2.703	-3.463	-1.929
	Glass's delta	.36042	-2.608	-3.517	-1.679

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control group.

Although the t-test is significant, Cohen's D effect size is moderately small. This means that even though the difference between the two means is statistically significant, it is a relatively small effect.

Based on the above analysis we can conclude that there is a statistically significant difference between first year students and students who are soon to complete the program in their performance on the critical thinking measure.

Question #3: Did fourth year students achieve a 3 or better.
SPSS Syntax:

```
DATASET ACTIVATE DataSet1.
SORT CASES  BY Class.
SPLIT FILE SEPARATE BY Class.
```

Descriptive Statistics[a]

	N	Minimum	Maximum	Mean	Std. Deviation
Totscore_avg	25	2.00	3.40	2.7640	.36042
Valid N (listwise)	25				

a. Class = Complete

```
DATASET ACTIVATE DataSet1.
SORT CASES  BY Class.
SPLIT FILE SEPARATE BY Class.
DESCRIPTIVES
VARIABLES=Totscore_avg
  /STATISTICS=MEAN STDDEV MIN
MAX.
```

We see here that overall, Fourth year students did not achieve a 3 or better. Their mean total score was 2.76.

Additional analysis:

```
DESCRIPTIVES VARIABLES=Exp_avg
Evi_avg Cont_avg Pos_avg Conc_avg
  /STATISTICS=MEAN STDDEV MIN MAX.
```

Descriptive Statistics[a]

	N	Minimum	Maximum	Mean	Std. Deviation
Exp_avg	25	1.00	4.00	2.7400	.87939
Evi_avg	25	1.00	4.00	2.8600	[illegible]
Cont_avg	25	1.00	4.00	2.9600	[illegible]0046
Pos_avg	25	1.00	4.00	2.5000	1.02062
Conc_avg	25	1.50	4.00	2.7600	.70887
Valid N (listwise)	25				

a. Class = Complete

We see that Context had the highest mean score (*M*=2.96), and Position had the lowest mean score (*M*=2.50) for seniors.

An interesting analysis here is the difference between the means for first year and fourth year students on their average scores for each criterion. An independent samples t-test was conducted to determine if the difference between the first year and fourth year students was significant.

```
SPLIT FILE OFF.
T-TEST GROUPS=Class('First'
'Complete')
  /MISSING=ANALYSIS
  /VARIABLES=Exp_avg Evi_avg
Cont_avg Pos_avg Conc_avg
  /ES DISPLAY(TRUE)
  /CRITERIA=CI(.95).
```

Group Statistics

	Class	N	Mean	Std. Deviation	Std. Error Mean
Exp_avg	First	25	1.6600	.73201	.14640
	Complete	25	2.7400	.87939	.17588
Evi_avg	First	25	1.9000	.80364	.16073
	Complete	25	2.8600	.89582	.17916
Cont_avg	First	25	1.9800	.84755	.16951
	Complete	25	2.9600	.90046	.18009
Pos_avg	First	25	1.7400	.66332	.13266
	Complete	25	2.5000	1.02062	.20412
Conc_avg	First	25	1.8400	.93229	.18646
	Complete	25	2.7600	.70887	.14177

Additional Analysis continued:

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Exp_avg	Equal variances assumed	1.690	.200	-4.720	48	.000	-1.08000	.22884	-1.54011	-.61989
	Equal variances not assumed			-4.720	46.471	.000	-1.08000	.22884	-1.54050	-.61950
Evi_avg	Equal variances assumed	.471	.496	-3.988	48	.000	-.96000	.24069	-1.44395	-.47605
	Equal variances not assumed			-3.988	47.445	.000	-.96000	.24069	-1.44409	-.47591
Cont_avg	Equal variances assumed	.220	.641	-3.962	48	.000	-.98000	.24732	-1.47727	-.48273
	Equal variances not assumed			-3.962	47.825	.000	-.98000	.24732	-1.47732	-.48268
Pos_avg	Equal variances assumed	7.521	.009	-3.122	48	.003	-.76000	.24345	-1.24948	-.27052
	Equal variances not assumed			-3.122	41.205	.003	-.76000	.24345	-1.25158	-.26842
Conc_avg	Equal variances assumed	1.621	.209	-3.928	48	.000	-.92000	.23424	-1.39096	-.44904
	Equal variances not assumed			-3.928	44.799	.000	-.92000	.23424	-1.39183	-.44817

Independent Samples Effect Sizes

		Standardizer[a]	Point Estimate	95% Confidence Interval	
				Lower	Upper
Exp_avg	Cohen's d	.80906	-1.335	-1.944	-.714
	Hedges' correction	.82199	-1.314	-1.914	-.703
	Glass's delta	.87939	-1.228	-1.872	-.566
Evi_avg	Cohen's d	.85098	-1.128	-1.721	-.525
	Hedges' correction	.86457	-1.110	-1.694	-.516
	Glass's delta	.89582	-1.072	-1.694	-.432
Cont_avg	Cohen's d	.87440	-1.121	-1.713	-.518
	Hedges' correction	.88837	-1.103	-1.687	-.510
	Glass's delta	.90046	-1.088	-1.713	-.446
Pos_avg	Cohen's d	.86072	-.883	-1.460	-.297
	Hedges' correction	.87446	-.869	-1.438	-.292
	Glass's delta	1.02062	-.745	-1.331	-.145
Conc_avg	Cohen's d	.82815	-1.111	-1.703	-.509
	Hedges' correction	.84138	-1.093	-1.676	-.501
	Glass's delta	.70887	-1.298	-1.952	-.625

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control group.

From both of the independent samples t-test of the mean scores for each criteria between the first year and fourth year students, we can see that although the fourth year students did not meet the target score of 3 or higher, there scores were significantly higher than the first year students. This is evident in viewing the changes in the overall score, as well as each criterion score. The effect size for each of these criteria was significant. This could indicate that that the weak correlations between separate items on this measure contribute to the low effect size obtained in the overall score.

Part 4: Differences between males and females, and 1st generation vs non-1st gen students on performance.

Analysis: Two way between subjects ANOVA.

Descriptive statistics and Assumptions.

```
EXAMINE VARIABLES=Totscore_avg BY Sex by Generation
  /PLOT BOXPLOT HISTOGRAM NPPLOT SPREADLEVEL(1)
  /COMPARE GROUPS
  /STATISTICS DESCRIPTIVES
  /CINTERVAL 95
  /MISSING LISTWISE
  /NOTOTAL.
```

Descriptives

	Sex	Generation			Statistic	Std. Error
Totscore_avg	F	0	Mean		2.3214	.14720
			95% Confidence Interval for Mean	Lower Bound	2.0034	
				Upper Bound	2.6394	
			5% Trimmed Mean		2.3071	
			Median		2.3000	
			Variance		.303	
			Std. Deviation		.55077	
			Minimum		1.50	
			Maximum		3.40	
			Range		1.90	
			Interquartile Range		.90	
			Skewness		.416	.597
			Kurtosis		-.340	1.154
		1	Mean		2.2917	.18359
			95% Confidence Interval for Mean	Lower Bound	1.8876	
				Upper Bound	2.6957	
			5% Trimmed Mean		2.2796	
			Median		2.3500	
			Variance		.404	
			Std. Deviation		.63598	
			Minimum		1.50	
			Maximum		3.30	
			Range		1.80	
			Interquartile Range		1.20	
			Skewness		.260	.637
			Kurtosis		-1.279	1.232
	M	0	Mean		2.2533	.14036
			95% Confidence Interval for Mean	Lower Bound	1.9523	
				Upper Bound	2.5544	
			5% Trimmed Mean		2.2648	
			Median		2.3000	
			Variance		.296	
			Std. Deviation		.54362	
			Minimum		1.30	
			Maximum		3.00	
			Range		1.70	
			Interquartile Range		1.10	
			Skewness		.042	.580
			Kurtosis		-.924	1.121
		1	Mean		2.3222	.23850
			95% Confidence Interval for Mean	Lower Bound	1.7722	
				Upper Bound	2.8722	
			5% Trimmed Mean		2.3302	
			Median		2.7000	
			Variance		.512	
			Std. Deviation		.71550	
			Minimum		1.40	
			Maximum		3.10	
			Range		1.70	
			Interquartile Range		1.45	
			Skewness		-.369	.717
			Kurtosis		-2.026	1.400

Case Processing Summary

			Cases					
			Valid		Missing		Total	
	Sex	Generation	N	Percent	N	Percent	N	Percent
Totscore_avg	F	0	14	100.0%	0	0.0%	14	100.0%
		1	12	100.0%	0	0.0%	12	100.0%
	M	0	15	100.0%	0	0.0%	15	100.0%
		1	9	100.0%	0	0.0%	9	100.0%

Assumption of Normality Is met

Tests of Normality

			Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Sex	Generation	Statistic	df	Sig.	Statistic	df	Sig.
Totscore_avg	F	0	.158	14	.200*	.960	14	.725
		1	.157	12	.200*	.921	12	.291
	M	0	.150	15	.200*	.935	15	.327
		1	.257	9	.089	.839	9	.056

*. This is a lower bound of the true significance.

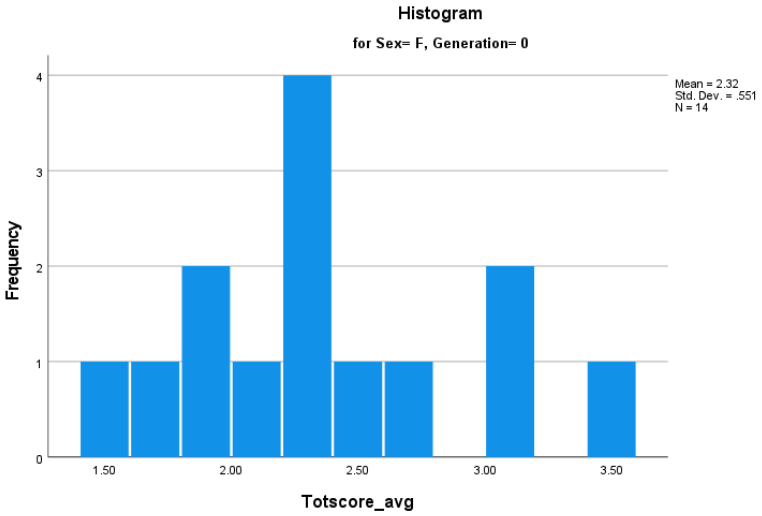
a. Lilliefors Significance Correction

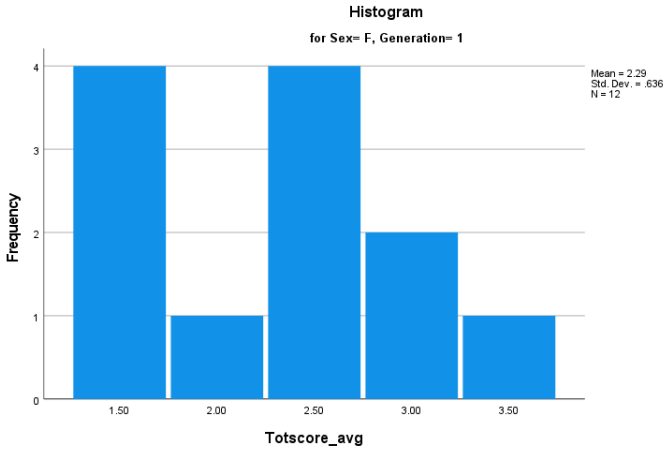
Assumption of Homogeneity of Variance Is met

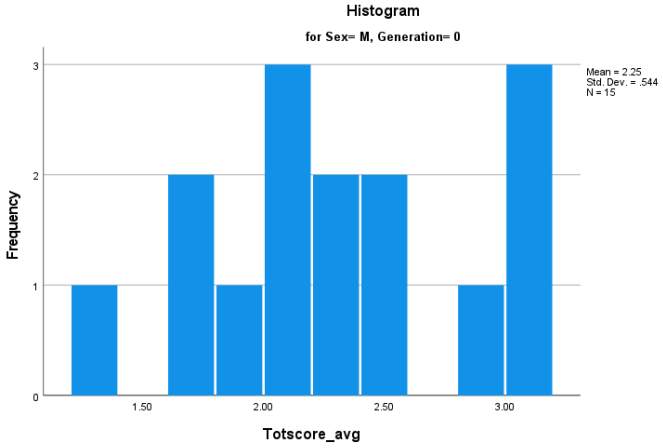
Test of Homogeneity of Variance

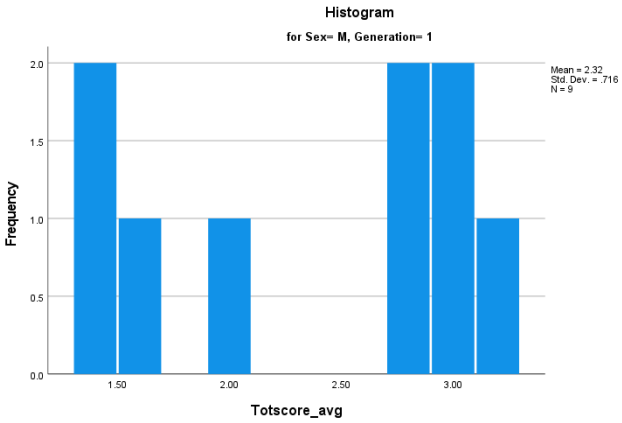
		Levene Statistic	df1	df2	Sig.
Totscore_avg	Based on Mean	1.233	3	46	.309
	Based on Median	.634	3	46	.597
	Based on Median and with adjusted df	.634	3	38.425	.598
	Based on trimmed mean	1.252	3	46	.302

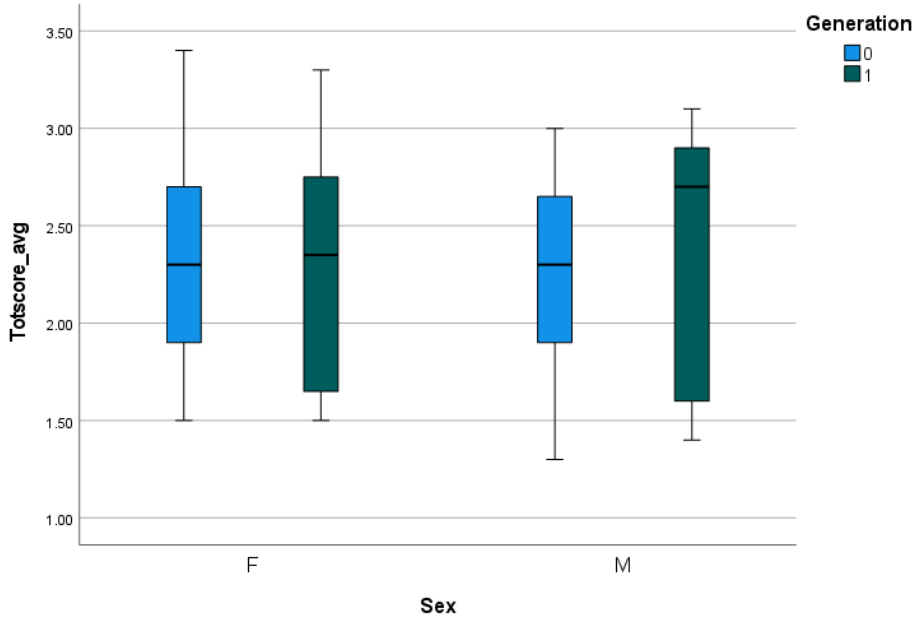
Assumptions Continued











Part 4: Differences between males and females, and 1st generation vs non-1st gen students on performance.

Analysis: Two way between subjects ANOVA.

```
UNIANOVA Totscore_avg BY Sex Generation
  /METHOD=SSTYPE(3)
  /INTERCEPT=INCLUDE
  /PLOT=PROFILE(Sex*Generation
Generation*Sex) TYPE=LINE ERRORBAR=NO
MEANREFERENCE=NO YAXIS=AUTO
  /EMMEANS=TABLES(Sex)
  /EMMEANS=TABLES(Generation)
  /EMMEANS=TABLES(Sex*Generation)
  /PRINT ETASQ DESCRIPTIVE
  /CRITERIA=ALPHA(.05)
  /DESIGN=Sex Generation Sex*Generation.
```

Descriptive Statistics

Dependent Variable: Totscore_avg

Sex	Generation	Mean	Std. Deviation	N
F	Non first-generation	2.3214	.55077	14
	first generation	2.2917	.63598	12
	Total	2.3077	.57960	26
M	Non first-generation	2.2533	.54362	15
	first generation	2.3222	.71550	9
	Total	2.2792	.59926	24
Total	Non first-generation	2.2862	.53833	29
	first generation	2.3048	.65382	21
	Total	2.2940	.58324	50

Tests of Between-Subjects Effects

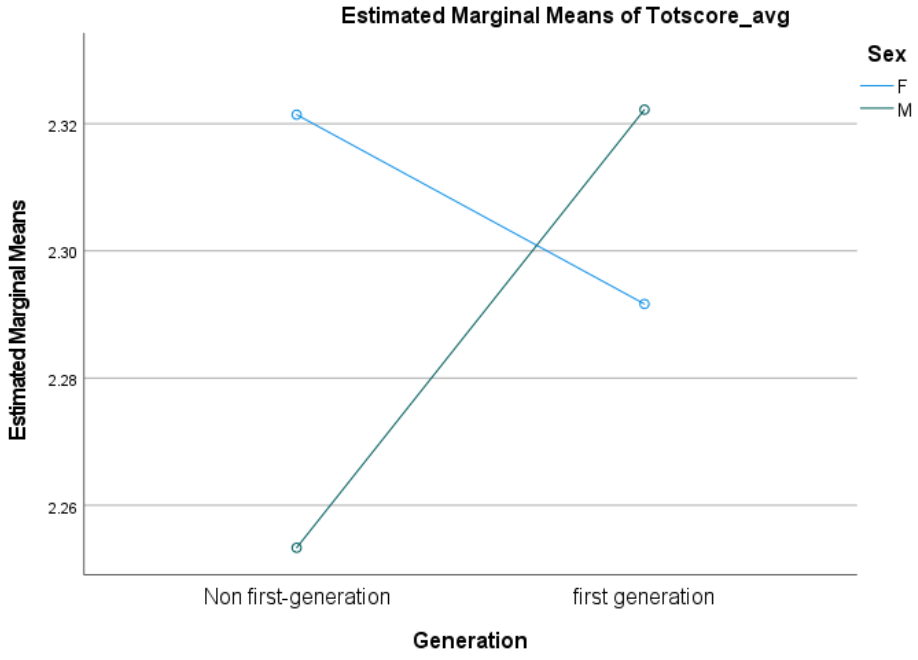
Dependent Variable: Totscore_avg

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	.043[a]	3	.014	.039	.989	.003
Intercept	253.898	1	253.898	702.489	.000	.939
Sex	.004	1	.004	.012	.914	.000
Generation	.005	1	.005	.013	.911	.000
Sex * Generation	.029	1	.029	.081	.777	.002
Error	16.626	46	.361			
Total	279.790	50				
Corrected Total	16.668	49				

a. R Squared = .003 (Adjusted R Squared = -.062)

The interaction between Sex and Generation is not significant.

The main effect of sex, and the main effect of generation are both non-significant on performance.



The intercept is displayed visually. However, we have unequal samples sizes and know from above that our interaction and main affects were non-significant. Future analysis may be needed with larger samples sizes in order to examine the relationship between Sex and Gender on performance on the critical thinking measure.

Summary

In summary, the above output indicates that the internal consistency of the critical thinking VALUE rubric is moderate to low (page 6). The inter rater agreement for scoring of the 50 students included in the study was high (page 7). First year students score significantly lower on the critical thinking VALUE rubric than do fourth year students (page 11, 13). Additionally, fourth year students are not currently meeting a score of three or above on the Critical Thinking VALUE Rubric (page 12). There are no differences between sex and generation on performance (page 16).

The critical thinking VALUE rubric contains 5 criteria. Many of these criteria have a weak correlation with each other, which could contribute to low internal consistency in reliability analysis. Further investigation into the reliability and validity of the critical thinking VALUE rubric is warranted. The established desired score of three should be evaluated for relevancy based on these analyses.

This study did not achieve balanced samples sizes for males and females, or generation. Future iterations may benefit from increased samples sizes.

Fourth year students are performing significantly better than first year students on the critical thinking VALUE Rubric.