Office of Academic Planning and Assessment

A Report of the Course Embedded

Texas Assessment of Critical Thinking Skills (TACTS)

2018-2019

Description of Introduction to Philosophy Pre- to Post-Test Assessment

Each fall and spring semester the Texas Assessment of Critical Thinking Skills (TACTS) test is administered within sections of PHIL 2303: Critical Thinking. The TACTS is a locally-developed, proprietary instrument designed to measure critical thinking, empirical, and quantitative skills. The instrument consists of 20 multiple choice questions and is administered to students enrolled in those courses at the start and end of each semester. As the instrument was developed by faculty with expertise in teaching and assessing critical thinking, it is assumed that the instrument has content related validity (Banta & Palomba, 2015). Additionally, as this test was embedded within normal sections of PHIL 2303, the student scores represent authentic student work (Banta & Palomba, 2015; Kuh et al. 2015).

The student data presented within this report reflect student performance regarding the Texas Higher Education Coordinating Board's Core Learning Objectives of Critical Thinking Skills and Empirical and Quantitative Skills (THECB, 2018). The THECB (2018) defines these concepts as follows:

- Critical Thinking Skills – creative thinking, innovation, inquiry, and analysis, evaluation, and synthesis of information
- Empirical and Quantitative Skills – manipulation and analysis of numerical data or observable facts resulting in informed conclusions

These data should therefore be used in conjunction with other data to fully understand student knowledge and ability with regards to these Core Learning Objectives.

Methodology

A total of 545 students took the pre-test and a total 362 students took the post-test for all sections of PHIL 2303: Critical Thinking for the 2018-2019 academic year; however not all student test scores were used for analysis. In order to determine whether student performance increased from pre-to-post, a dependent samples t-test was used for analysis. Student SamID's were collected along with student scores in order to identify each student's score on both the pre- and post-test. A total of 312 students provided their SamID's and took both the pre- and post-tests. All statistical analysis was therefore conducted on only those students for whom both pre- and post-test scores could be identified. In order to further disaggregate the results, the data was also analyzed separately for face-to-face and online students.

Prior to conducting inferential statistics to determine whether differences were present between the students' pre- to post-test scores, checks were conducted to determine the extent to which these data were normally distributed. All four of the standardized skewness and kurtosis coefficients were within the limits of normality of +/-3 for the face-to-face and online student populations, and three of the four coefficients were within the limits of normality for the combined populations (Onwuegbuzie & Daniel, 2002). Therefore, parametric dependent samples t-tests were used for all statistical analysis. Readers are directed to Table 1 for a breakdown of these results.

Table 1
Standardized Skewness and Kurtosis Values for Student Scores Pre- and Post-test Scores

Student Population	Standardized Skewness Coefficient	Standardized Kurtosis Coefficient
Face-to-Face Students		
Pre-Test	2.65	1.75
Post-Test	0.72	-1.04
Online Students		
Pre-Test	-1.74	-0.46
Post-Test	0.37	0.75
All Students		
Pre-Test	3.28	1.93
Post-Test	1.46	-1.69

Results

A parametric dependent samples *t*-test revealed a statistically significant difference between the pre-to-post scores for students enrolled in face-to-face sections of PHIL 2303: Critical Thinking for the 2018-2019 academic year, $t(269) = -7.27$, $p < .001$. This difference represented a small effect size (Cohen's *d*) of 0.46 (Cohen, 1988). The average student score increased from 34.91% to 40.83%, for an increase of 5.92%. This equated to an average increase of 1.19 questions answered correctly from pre-to-post. Readers are directed to Table 2 for a breakdown of these results.

Table 2

Descriptive Statistics for Student Pre- and Post-Scores on Course-Embedded Test in PHIL 2303: Critical Thinking for 2018-2019 (Face-to-Face)

Test Version	*M*	*SD*	*M %*	*SD %*
Pre-test Scores	6.98	2.44	34.91	12.21
Post-test Scores	8.17	2.72	40.83	13.60

Note. The number of students was 270.

A parametric dependent samples *t*-test did not reveal a statistically significant difference between the pre-to-post scores for students enrolled in online sections of PHIL 2303: Critical Thinking for the 2018-2019 academic year, $t(41) = 0.41$, $p = .685$. The average student score decreased from 23.45% to 22.86%, for a decrease of 0.59%. This equated to an average decrease of 0.12 questions answered correctly from pre-to-post. Readers are directed to Table 3 for a breakdown of these results.

Table 3

Descriptive Statistics for Student Pre- and Post-Scores on Course-Embedded Test in PHIL 2303: Critical Thinking for 2018-2019 (Online)

Test Version	M	SD	M %	SD %
Pre-test Scores	4.69	1.47	23.45	7.37
Post-test Scores	4.57	1.43	22.86	7.17

Note. The number of students was 42.

      A parametric dependent samples *t*-test revealed a statistically significant difference between the pre-to-post scores for all students enrolled in sections of PHIL 2303: Critical Thinking for the 2018-2019 academic year, $t(311) = -6.80$, $p < .001$. This difference represented a small effect size (Cohen's *d*) of 0.38 (Cohen, 1988). The average student score increased from 33.37% to 38.41%, for an increase of 5.04%. This equated to an average increase of 1.01 questions answered correctly from pre-to-post. Readers are directed to Table 4 for a breakdown of these results.

Table 4

Descriptive Statistics for Student Pre- and Post-Scores on Course-Embedded Test in PHIL 2303: Critical Thinking for 2018-2019 (All Students)

Test Version	M	SD	M %	SD %
Pre-test Scores	6.67	2.46	33.37	12.30
Post-test Scores	7.68	2.86	38.41	14.30

Note. The number of students was 312.

      Additional important information regarding student performance can also be gained through an item analysis of student pre- and post-test performance on individual test questions for each of the examined student populations. This item analysis revealed that students in face-to-face sections scored statistically significantly higher on 5 of the 20 test questions (Questions 1, 8, 9, 11, 12) from pre-to-post. Readers are directed to Table 5 for a complete breakdown of item analysis data for face-to-face students.

Table 5

Percentage of Face-to-Face Students Correctly Answering Pre- and Post-Test Questions

	Pre-Test	Post-Test	Mean Difference	Cohen's *d*
Question 1	27.8%	46.3%	18.5%***	0.39
Question 2	58.1%	54.1%	-4.0%	
Question 3	16.3%	10.7%	-5.6%	
Question 4	30.7%	34.4%	3.7%	
Question 5	71.1%	66.3%	-4.8%	
Question 6	4.4%	5.2%	0.8%	
Question 7	31.1%	31.5%	0.4%	
Question 8	15.9%	57.0%	41.1%***	0.94

Question 9	17.4%	46.3%	28.9%***	0.65
Question 10	7.8%	11.5%	3.7%	
Question 11	39.6%	48.1%	8.5%*	0.17
Question 12	22.2%	31.9%	9.7%**	0.21
Question 13	72.2%	77.0%	4.8%	
Question 14	14.1%	15.6%	2.5%	
Question 15	24.4%	29.6%	5.2%	
Question 16	66.3%	72.2%	5.9%	
Question 17	33.7%	28.1%	-5.6%	
Question 18	42.6%	48.9%	6.3%	
Question 19	29.6%	23.7%	-5.9%	
Question 20	72.6%	78.1%	5.5%	

Note. $n = 270$. * significant at $p \leq 0.05$; ** significant at $p \leq 0.01$; *** significant at $p \leq 0.001$. Cohen's *d* from 0.2 – 0.49 indicate a small effect size, 0.50-0.79 indicate a moderate effect size, and 0.80 and higher indicate a large effect size (Cohen, 1988).

An item analysis revealed that students in online sections scored statistically significantly higher on 2 of the 20 test questions (Questions 2, 9) from pre-to-post. Furthermore, students in online sections scored statistically significantly lower on 4 of the 20 test questions (Questions 3, 11, 13, 14) from pre-to-post. Readers are directed to Table 6 for a complete breakdown of item analysis data for online students.

Table 6

Percentage of Online Students Correctly Answering Pre- and Post-Test Questions

	Pre-Test	Post-Test	Mean Difference	Cohen's *d*
Question 1	33.3%	35.7%	2.4%	
Question 2	38.1%	69.0%	30.9%**	0.64
Question 3	31.0%	11.9%	-20.0%*	0.46
Question 4	33.3%	28.6%	-4.7%	
Question 5	57.1%	57.1%	0.0%	
Question 6	2.4%	7.1%	4.7%	
Question 7	21.4%	35.7%	14.3%	
Question 8	2.4%	14.3%	11.9%	
Question 9	23.8%	61.9%	40.5%**	0.82
Question 10	21.4%	11.9%	-9.5%	
Question 11	26.2%	9.5%	-16.7%*	0.44
Question 12	0.0%	4.8%	4.8%	
Question 13	40.5%	7.1%	-33.4%***	0.84
Question 14	28.6%	0.0%	-28.6%***	0.89
Question 15	26.2%	33.3%	7.1%	
Question 16	33.3%	33.3%	0.0%	
Question 17	0.0%	0.0%	0.0%	
Question 18	11.9%	9.5%	-2.4%	
Question 19	9.5%	11.9%	2.4%	
Question 20	28.6%	14.3%	-14.3%	

Note. $n = 42$. * significant at $p \leq 0.05$; ** significant at $p \leq 0.01$; *** significant at $p \leq 0.001$. Cohen's *d* from 0.2 – 0.49 indicate a small effect size, 0.50-0.79 indicate a moderate effect size, and 0.80 and higher indicate a large effect size (Cohen, 1988).

Finally, an item analysis revealed that all students combined scored statistically significantly higher on 4 of the 20 test questions (Questions 1, 8, 9, 12) from pre-to-post. Furthermore, all students combined scored statistically significantly lower on 1 of the 20 test questions (Question 3) from pre-to-post. Readers are directed to Table 7 for a complete breakdown of item analysis data for face-to-face students.

Table 7

Percentage of All Students Correctly Answering Pre- and Post-Test Questions

	Pre-Test	Post-Test	Mean Difference	Cohen's *d*
Question 1	28.5%	44.9%	16.4%***	0.34
Question 2	55.4%	56.1%	0.7%	
Question 3	18.3%	10.9%	-7.4%**	0.21
Question 4	31.1%	33.7%	2.6%	
Question 5	69.2%	65.1%	-4.1%	
Question 6	4.2%	5.4%	1.2%	
Question 7	29.8%	32.1%	2.3%	
Question 8	14.1%	51.3%	37.2%***	0.86
Question 9	18.3%	48.4%	30.1%***	0.67
Question 10	9.6%	11.5%	1.9%	
Question 11	37.8%	42.9%	5.1%	
Question 12	19.2%	28.2%	9%**	0.21
Question 13	67.9%	67.6%	-0.3%	
Question 14	16.0%	13.5%	-2.5%	
Question 15	24.7%	30.1%	5.4%	
Question 16	61.9%	67.0%	5.1%	
Question 17	29.2%	24.4%	-4.8%	
Question 18	38.5%	43.6%	5.1%	
Question 19	26.9%	22.1%	-4.8%	
Question 20	66.7%	69.6%	2.9%	

Note. $n = 312$. * significant at $p \leq 0.05$; ** significant at $p \leq 0.01$; *** significant at $p \leq 0.001$. Cohen's *d* from 0.2 – 0.49 indicate a small effect size, 0.50-0.79 indicate a moderate effect size, and 0.80 and higher indicate a large effect size (Cohen, 1988).

References

Banta, T. W., & Palomba, C. A. (2015). *Assessment essentials: Planning implementing, and improving assessment in higher education* (2nd ed.). San Francisco, CA: Jossey-Bass.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.)*.* Hillsdale, NJ: Lawrence Erlbaum.

Kuh, G. D., Ikenberry, S. O., Jankowski, N. A., Cain, T. R., Ewell, P. T8., Hutchings, P., Kinzie, J. (2015). *Using evidence of student learning to improve higher education*. San Francisco, CA: Jossey-Bass.

Onwuegbuzie, A. J., & Daniel, L. G. (2002). Uses and misuses of the correlation coefficient. *Research in the Schools, 9*(1)*,* 73-90.