SmarterMeasure Assessment Report – Fall 2014

Author: Joseph F. van Gaalen, Ph.D., Coordinator, Academic Assessment

1 INTRODUCTION

Florida SouthWestern State College's Quality Enhancement Plan goal is to provide students ways to develop strategies for effective writing and verbal communication and demonstrate independence through personal management, use of college resources, and positive relationships with peers, staff, and faculty. One measurement for the achievement of that goal is the use of the SmarterMeasure Learning Readiness Indicator. FSW has identified set criteria for defining student advancement in the Cornerstone Experience course through the published assessment tool SmarterMeasure.

The SmarterMeasure Learning Readiness Indicator was designed to help quantify the degree of student preparedness and breadth of skill set (SmarterServices, 2011). Through course completion, students are expected to statistically significantly improve in the following learning readiness indicators of the SmarterMeasure Student Readiness Test: Personal Attributes, Life Factors, Technology Knowledge, and Technology Competency. These indicators can serve as an effective tool for evaluating growth over time, allowing college assessment faculty and staff a means to identify strengths and weaknesses, and allow the college to compare FSW students with those across the country, if desired (Suskie, 2004). Results are compared by term in a longitudinal study in support of the QEP.

For additional detail or further analysis not provided in this report, please contact Dr. Joseph F. van Gaalen, Coordinator of Academic Affairs Assessment, Academic Affairs (Joseph.VanGaalen@fsw.edu; x6965).

2 STATISTICS

2.1 DESCRIPTIVE STATISTICS AND SIGNIFICANCE TESTING

During the fall 2014 semester, 2376 total tests (pre- and post-) were administered to students. Of those, between 1891 and 1930 of them were pre-/post- paired tests. Some students did not have pre-/post-counterparts for all readiness indicators. Depending on the readiness indicator, between 446 and 485 tests did not have counterparts. The differences in the means from pre-to-post test scores for each readiness indicator was tested for significance using a paired means t-test according to standard methods (Davis, 1973; McDonald, 2009; Siegel, 1956; Wilkinson, 1999) are shown in Table 1.

	Personal Attributes	Life Factors	Technology Knowledge	Technology Competency
Pre-test Mean	74.3	77.2	66.3	87.9
Pre-test standard deviation	7.3	8.3	11.7	11.9
Post-test Mean	73.6	76.1	69.3	87.8
Post-test Standard deviation	8.0	9.7	11.4	13.8
Effect size	-0.165	-0.239	0.530	-0.002
p-value	**6.84×10^{-5}**	**7.88×10^{-8}**	**3.87×10^{-27}**	0.891

*Table 1. Comparison of mean scores for Pre/post test scores. Bold denote statistically significant difference. Positive effect sizes indicate a higher mean score for Post-test scores. *Denote marginal significance as defined by Johnson (2013).*

Effect size was calculated using a method devised by Rosenthal and Rosnow (1991) for meta-analytical purposes in potential comparisons with other institutions (Lipsey and Wilson, 1993). The statistically significant results exhibit what Cohen (1988) would consider small to moderate effect sizes ranging from 0.002 to 0.53 (Table 1). In other words, non-overlap from Pre-test scores to Post-test scores range from approximately 0% to 34%.

The paired means t-test results indicate that for three of the four readiness indicators, Personal Attributes, Life Factors, and Technology Knowledge, we must reject the null hypothesis that the difference in the means of the pre- and post-test scores are equal to 0, and we can conclude this with a 95% confidence that the differences in scores are not solely due to chance. The other readiness indicator, Technology Competency, we cannot reject the null hypothesis that the slight decrease in the mean score from pre-to-post-test scores can be a result of chance.

2.2 Supporting Exploratory Data Analysis

The significance tests only provide information as to whether the two groups are unchanged. The assumption is that instruction is the main difference between pre-/post-tests, but it cannot be certain. It is necessary to explore in detail each readiness indicator as much as possible using multiple standard processes for support of significance testing.

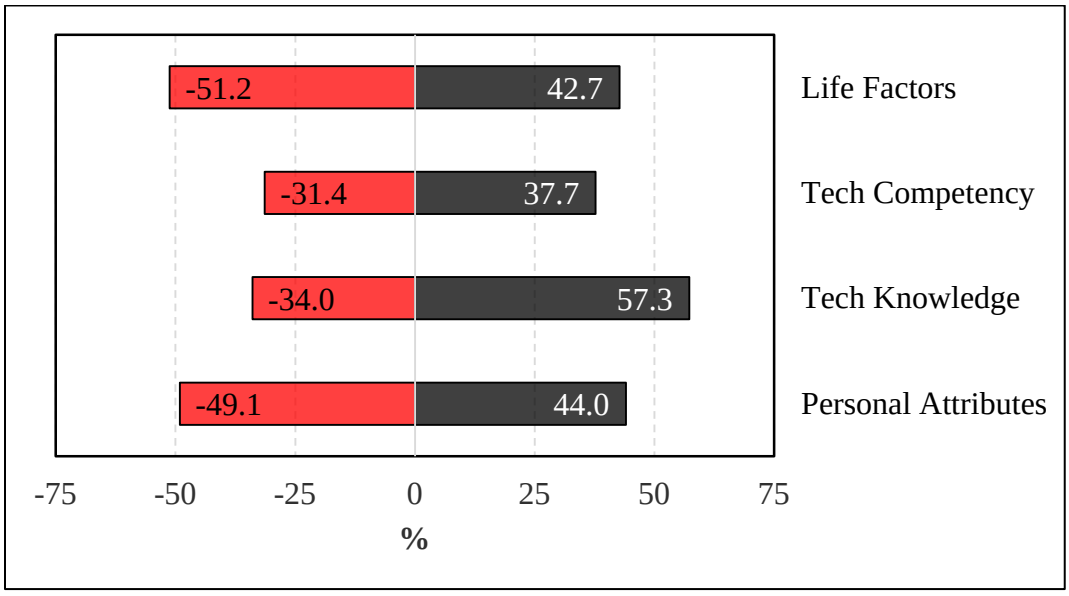

Figure 1. Percentage of students improved vs. declined by readiness indicator. Difference of the sum of increase and decline from 100 is the percentage of test takers that exhibited no change.

Figure 1 highlights the percentage of student test scores that improved and declined. The Technology Knowledge readiness indicator exhibits the greatest improvement percentage at 57.3%. The Technology Competency indicator exhibits the least improvement; however it also exhibits the least decline as 30.1% of students scored no change from pre-to-post test, the highest of any readiness indicator. The Life Factors indicator exhibits the greatest decline at 51.2%.

An empirical distribution (histogram) of pre- and post-test for each readiness indicator is reported in Figures 2 through 5. Figure 2 depicts data distribution of the Personal Attributes readiness indicator

which exhibits a subtle decrease in kurtosis and no discernable shift in data distribution or modality (central peak location). Decreases in the 70-79 scoring bin accompany increases in both the 50-59 and 60-69 scoring bins along with a slight increase in the ≥90 scoring bin.

Figure 3 depicts data distribution of the Technology Knowledge readiness indicator. A shift in modality can be seen with peak centered on 70-79 scoring bin in post-tests compared with 60-69 in the pre-test. Decreases in scoring bin 60-69 and below accompany increases in 70-79 scoring bin and above.

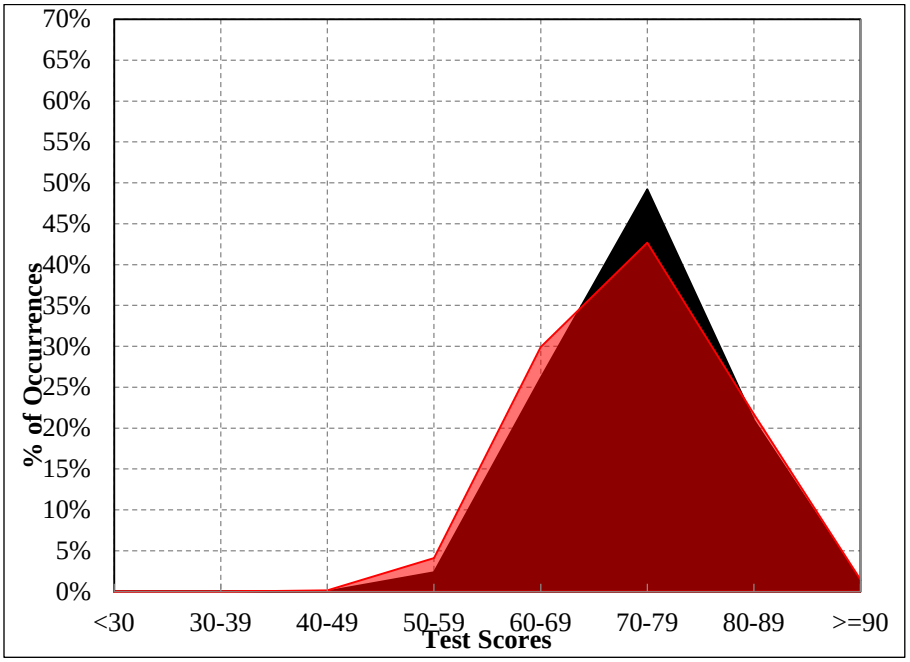

Figure 2. Personal Attributes readiness indicator distribution of test scores for Pre- (black) and Post- (red).

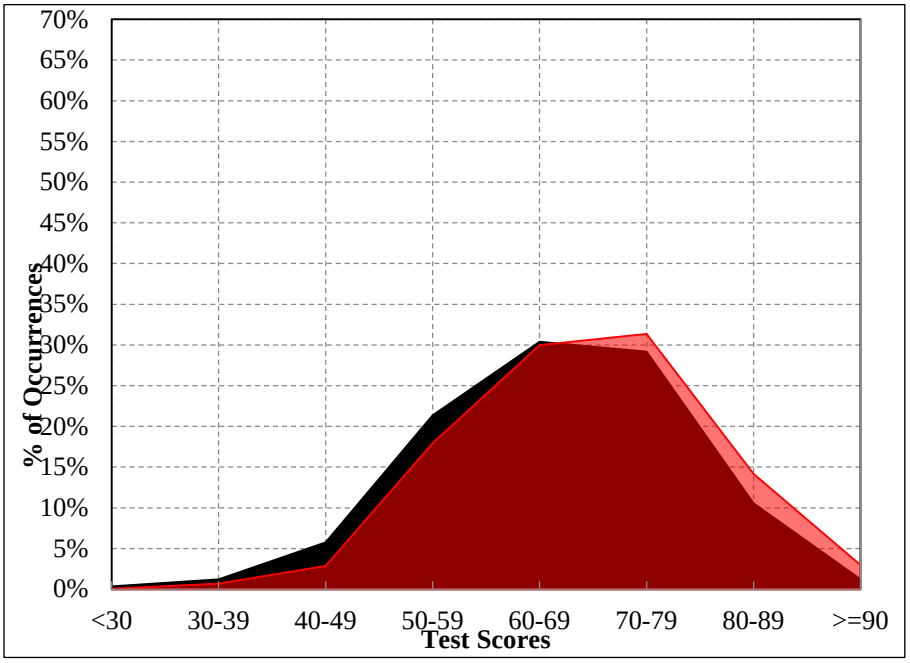

Figure 3. Technology Knowledge readiness indicator distribution of test scores for Pre- (black) and Post- (red).

Figure 4 depicts data distribution of the Technology Competency readiness indicator. Post-test scores exhibit increased kurtosis (peak height) compared with pre-test scores. Decreases in scoring bins 60-69, 70-79, and 80-89 accompany a large increase in the ≥90 scoring bin.

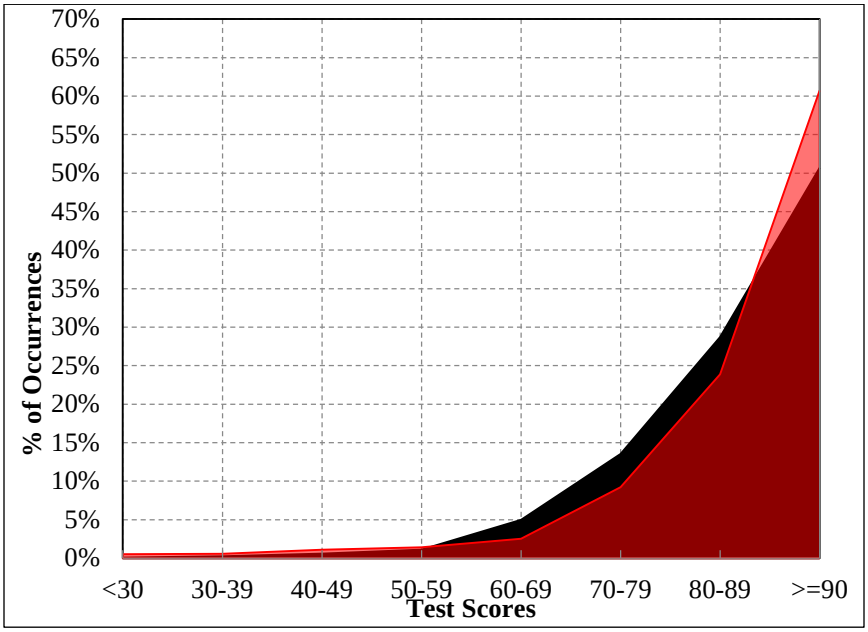

Figure 4. Technology Competency readiness indicator distribution of test scores for Pre- (black) and Post- (red).

Figure 5 depicts data distribution of the Life Factors readiness indicator. Post-test scores exhibit decreased kurtosis (peak height) compared with pre-test scores. Decreases in scoring bins 70-79 and 80-89 accompany increases in 50-59 and 60-69 scoring bins exhibiting a slight negative shift from pre-to-post test scores.

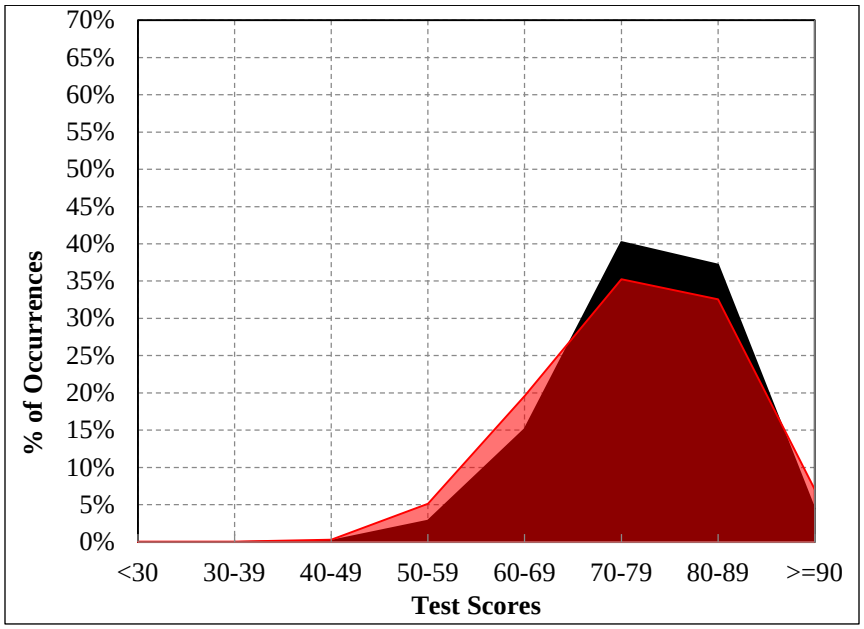

Figure 5. Life Factors readiness indicator distribution of test scores for Pre- (black) and Post- (red).

2.3 Comparison with Previous FSW Assessments

The results of paired means t-test of pre-/post- test scores for all semesters from Fall 2012 through Fall 2014 are shown in Table 2. The Technology Knowledge readiness indicator exhibits a consistent statistically significant increase in mean scores from pre-to-post test scores for the entire study period. Both the Personal Attributes indicator and the Life Factors indicator exhibit no trend in differences in the means from pre-to-post test scores.

	Fall 2012	Spring 2013	Summer 2013	Fall 2013	Spring 2014	Summer 2014	Fall 2014
Personal Attributes	-0.2	-1.0*	0.8	-0.5	-0.3	0.8	-0.7
Technology Knowledge	3.8	4.2	3.1	3.1	3.9	5.5	3.0
Technology Competency	-0.1	2.0*	3.0	-1.0	2.1	0.5	0.0
Life Factors	-0.4	0.6	2.1	-0.2	0.2	2.5	-1.1

*Table 2. SmarterMeasure Pre-/Post- test results mean difference. Comparison of significance test results for mean difference of pre-/post-test scores for Fall 2012 through Spring 2015. Shaded cells indicate statistically significant differences in the mean at the 95% confidence level. Red denotes a decrease from pre-to-post. *Denote marginal significance as defined by Johnson (2013).*

Table 3 provides additional information regarding the paired means t-test including the observed t-statistic (t_{obs}) and probability of difference due to chance (p-value) with respect to the degrees of freedom for each study. No effect size is reported since the method of calculation has not been consistent historically (Rosenthal and Rosnow, 1991; Wilson et al., 1995; Dunlop et al., 1996).

	Fall '12 t_{crit} = 1.97	Spring '13 t_{crit} = 1.97	Summer '13 t_{crit} = 1.98	Fall '13 t_{crit} = 1.97	Spring '14 t_{crit} = 1.97	Summer '14 t_{crit} = 1.97	Fall '14 t_{crit} = 1.97
Personal Attributes	t(292)=-0.44, p=0.66	t(200)=-2.01, p=0.0464*	t(166)=1.81, p=0.072	t(779)=-1.89, p=0.0594	t(430)=-0.90, p=0.371	t(191)=1.89, p=0.060	t(1925)=-3.6, $p=6.84x10^{-5}$
Technology Knowledge	t(292)=7.04, p<<0.001	t(193)=6.37, p<<0.001	t(166)=4.74, p<<0.001	t(775)=8.91, p<<0.001	t(425)=7.72, $p=9.57x10^{-14}$	t(189)=9.20, $p=7.09x10^{-17}$	t(1890)=11.6, $p=3.87x10^{-27}$
Technology Competency	t(292)=-0.09, p<<0.001	t(194)=2.29, p=0.023*	t(166)=3.17, p=0.002	t(775)=-1.84, p=0.066	t(426)=3.28, $p=8.71x10^{-4}$	t(189)=0.48, p=0.631	t(1921)=0.0, p=0.891
Life Factors	t(292)=-0.79, p<<0.001	t(200)=0.86, p=0.390	t(166)=3.57, p=0.0005	t(779)=-0.71, p=0.478	t(431)=0.63, p=0.533	t(191)=4.88, $p=2.29x10^{-6}$	t(1928)=-5.2, $p=7.88x10^{-8}$

*Table 3. Additional significance testing statistics for learning dimensions including observed t-stat (t_{obs}), probability of difference due to chance (p-value), degrees of freedom (df), and critical t-stat. In some cases, earlier reports did not include p-value when p<0.05 or in later studies, p<<0.001 and are indicated where applicable. Mean difference of pre-/post-test scores are reported in Table 3. *Denote marginal significance as defined by Johnson (2013).*

Figure 6 is a graphical representation of the difference in mean scores of readiness indicators across all semesters. The consistently strong improvement in scores from pre-to-post tests of the Technology Knowledge indicator (red bar) is clearly visible. All other readiness indicators at exhibit negative changes during at least one term. The demographic of each term is also visible. Note that each fall term exhibits a strong positive for Technology Knowledge with negative changes for all other indicators. Spring terms exhibit positive changes for all but the Personal Attributes indicator. Finally, summer terms exhibit positives for all readiness indicators.

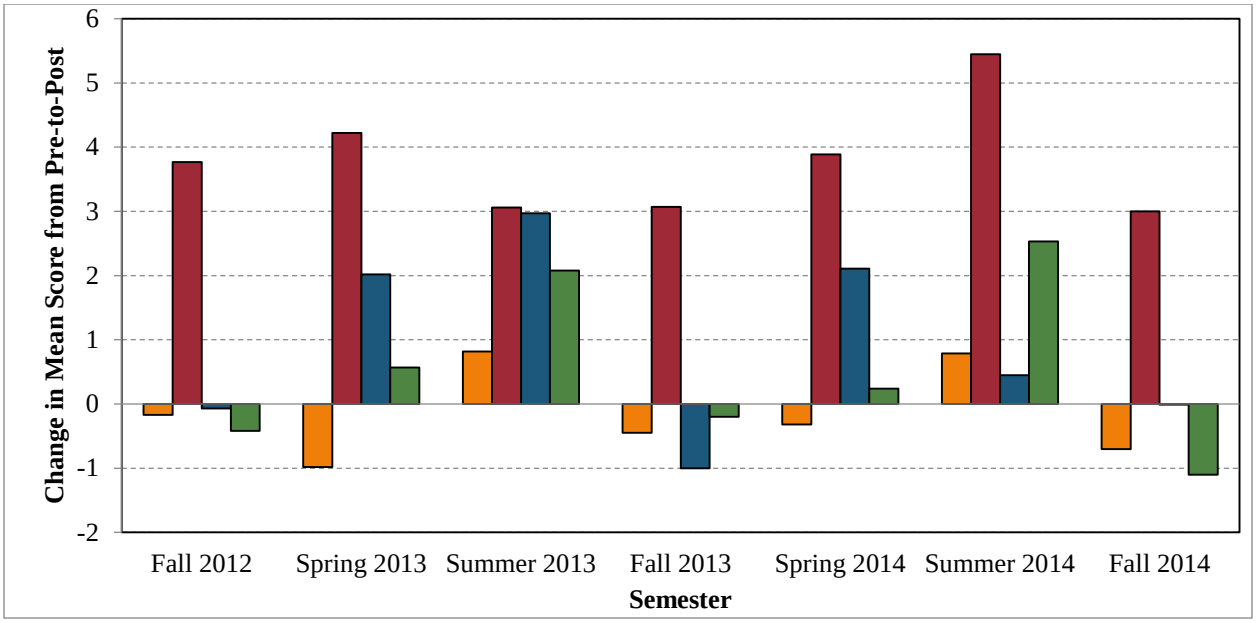

Figure 6. Comparison of the difference in mean scores from pre-to-post tests across semester by readiness indicator. Orange (Personal Attributes), red (Technology Knowledge), blue (Technology Competency), and green (Life Factors).

3 CONCLUSIONS

In FSW's QEP assessment, the goal is to measure changes in student disposition in the four readiness indicators of the SmarterMeasure test: Personal Attributes, Technology Knowledge, Technology Competency, and Life Factors. These dispositions are measured based on testing dates between the first three weeks of the semester and the last three weeks of the semester in a 16-week semester.

A drilldown of the results of the SmarterMeasure Fall 2014 assessment are as follows:

1. The means of post-test scores exhibit statistically significant improvement over means of pre-test scores in the Technology Knowledge readiness indicator only. Both the Personal Attributes and Life Factors indicators exhibit statistically significant declines.
2. The Technology Knowledge readiness indicator exhibits no statistically significant change between pre-test and post-test scores.
3. In a comparison of students who improved from pre-to-post and those that didn't, both the Technology Competency and Technology Knowledge readiness indicators exhibit a greater population of students who improved from pre-to-post test scores than declined.
4. In that same study, both the Life Factors and Personal Attributes readiness indicators exhibit a greater population of students who declined from pre-to-post test scores than improved.
5. In a longitudinal study, since comparisons began in Fall 2012, Technology Knowledge has consistently exhibited the largest improvement in scores from pre-/post-test scores. The indicator has exhibited a statistically significant increase in mean score across all semesters.
6. In the same longitudinal study, since comparisons began, semesters have consistently exhibited the same trends. All fall semesters exhibit an increase in Technology Knowledge with declines in the other three indicators. All spring semesters exhibit an increase in all indicators except Personal Attributes. And all summer semesters exhibit increases in all indicators.

4 REFERENCES

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

Davis, J.C. 1973. Statistics and Data Analysis in Geology. John Wiley & Sons, New York, New York, 564 pp.

Dunlop, W.P., Cortina, J.M., Vaslow, J.B., and Burke, M.J. 1996. Meta-analysis of experiments with matched groups or repeated measures designs. Psychological Methods, 1, 170-177.

Johnson, V. 2013. Revised Standards for Statistical Evidence. Proceedings of the National Academy of Science, 110(48), 19313-19317.

Lipsey, M.W. and Wilson, D.B. 1993. The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. American Psychologist, 48, 1181-1209.

McDonald, J.H. 2009. Handbook of Biological Statistics (2nd ed.). Sparky House Publishing, Baltimore, Maryland.

Rosenthal, R. and Rosnow, R.L. 1991. Essentials of behavioral research: Methods and data analysis. McGraw Hill, New York, 692 pp.

Siegel, S. 1956. Nonparametric statistics for the behavior sciences. McGraw-Hill, New York, New York, 312 pp.

SmarterServices. 2011. 2011 Student Readiness Report. Deatsville, Alabama: unknown.

Suskie, L. 2004. Assessing Student Learning. Anker Publishing Co., Inc., Bolton, Massachusetts, 331 pp.

Wilkinson, L. 1999. APA Task Force on Statistical Inference. Statistical Methods in Psychology Journals: Guidelines and Explanations. American Psychologist 54 (8), 594–604.