A STUDY OF RELIABILITY AND VALIDITY OF THE FELDER-SOLOMAN INDEX OF LEARNING STYLES FOR BUSINESS STUDENTS

Dr. Brian H. Felkel and Dr. Ross M. Gosky
Appalachian State University
329 Walker Hall; 121 Bodenheimer Drive; Boone, NC 28608
felkelbh@yahoo.com and goskyrm@appstate.edu

Abstract

We assessed the validity and reliability of the Felder-Soloman Index of Learning Styles (ILS) instrument for a group of Business Calculus students at Appalachian State University in North Carolina, a state university with approximately 15,000 students. Students completed the ILS questionnaire twice, once at the beginning of the fall semester and again at the end of the fall semester. Our results suggest that the ILS measurement is a reasonably valid and reliable measure of learning style for Business Calculus students, who are presumed to be representative of Business students early in their college careers. Our results are consistent with published results for students from other majors as well.

Keywords: Internal validity, reliability, Felder-Soloman Index, Cronbach's Alpha, Factor Analysis, Business Education

Section 1: Introduction

Section 1.1: Learning Styles

The Felder-Soloman Index of Learning Styles (ILS) has been used among engineering educators. The ILS measure of learning style built upon a paper by Felder and Silverman, 1988, which described learning styles of engineering students. The goal of the ILS is to measure learning style preferences across four dimensions or continua (active/reflexive, sensing/intuitive, visual/verbal, and sequential/global). The ILS was introduced by Richard Felder and Barbara Soloman in 1991 and made available on the World Wide Web in 1996. A description of the different styles and learning or teaching strategies for those styles are also available online, at http://www4.ncsu.edu/unity/lockers/users/f/felder/public/ILSpage.html.

Learning styles are not very good predictors of academic success (Van Zwanenberg, Wilkinson, and Anderson, 2000). Also the dimensions are not either/or categories. Students can and do use different styles in the same dimension. A student's learning style preference can be affected by different educational experiences. Faculty who are aware of the different learning styles are able to modify their teaching style to incorporate methods that reach students with different learning styles.

Students also benefit from knowing their learning style preferences. A student who knows their learning style can be made aware of study techniques that correspond with their preferred learning style. When combined with classroom instruction sensitive to different learning styles, knowledge of one's learning style may improve student performance and satisfaction.

Section 1.2: Reliability and Validity

The main goal of this study is to assess both reliability and validity of the ILS questionnaire within a specific subpopulation of students, namely Business Calculus students at Appalachian State University in Boone, North Carolina.

The reliability of an assessment tool refers to the tool producing consistent, repeatable results. This can be assessed through test-retest reliability and internal consistency reliability. Test-retest reliability is assessed by testing the same individuals at different times and comparing the results. Internal consistency reliability refers to the consistency of results across items within a test, and is assessed by checking for correlated answers for multiple questions designed to test the same construct. This can be assessed using Cronbach's alpha, a measure of how well different items within a test designed to measure similar qualities are related. Large values of Cronbach's alpha indicate the different items are related and suggest they measure a common underlying factor. Small values of Cronbach's alpha suggest the opposite. For a more detailed discussion of these definitions and tests see http://www.socialresearchmethods.net/kb/reltypes.htm.

Validity of an assessment tool requires that the tool measure what it claims to measure (http://www.uni.edu/chfasoa/reliabilityandvalidity.htm). In particular we are interested in construct validity. The definition of construct validity is the degree to which inferences

can legitimately be made from the operationalizations in the study to the theoretical constructs on which those operationalizations were based (http://www.socialresearchmethods.net/kb/constval.htm). Construct validity assesses how well ideas or theories are translated into actual programs or measures. For the ILS questionnaire, we tested discriminant validity by checking whether the four dimensions of the ILS are truly four separate, non-overlapping dimensions of learning style. We use these results to determine whether this construct validity is found in our study.

Many studies of reliability and validity tests have been done on ILS, but all of them expressed the need for further verification (Felder and Spurlin, 2005; Lintzinger, Lee, and Wise, 2005; Zywno, 2003; Van Zwanenberg, Wilkinson, and Anderson, 2000). Almost all of the reliability/validity studies have centered on engineering majors. The work by Van Zwanenberg et al. (2000) included some Business majors with engineering majors, but the two groups were not compared to or contrasted with each other. For an excellent review of past studies and an independent validity study, we recommend the paper by Zywno (2003). Table 1 repeats information from Zywno (2003), comparing internal consistency/reliability across several studies, here. The α referred to in the table is the standard Cronbach α from Cronbach's alpha/factor analysis.

Table 1:
Internal Consistency Reliability Comparisons from Zywno (2003)

Study	**N**	**Active Scale α**	**Sensing Scale α**	**Visual Scale α**	**Sequent. Scale α**
Van Zwanenburg et al.	279	0.51	0.65	0.56	0.41
Livesay et al.	255	0.56	0.72	0.60	0.54
Felder & Spurlin	584	0.70	0.76	0.69	0.55
Ryerson, Canada	557	0.60	0.70	0.63	0.53
Ryerson, Canada*	406	0.60	0.69	0.61	0.50

*Test-Retest Data and 1999 Sample Excluded

Among social science literature, an alpha of 0.7 or higher is preferred in order for a set of questions to be considered a scale. Lower minimum values, however, are not unheard of. In fact Tuckman (1999) asserts that the alpha for attitude tests should be above 0.5.

Section 2: Methodology

Section 2.1: Description of Participants

The study took place at Appalachian State University (ASU), in Boone, North Carolina, during the fall 2007 semester. Our goal was to assess the ILS among Business majors and compare the results to studies involving engineering majors. At ASU successful completion of Business Calculus is required before admission into the College of Business. Thus all of the Business Calculus students can be assumed to be potential Business majors, and our participants were Business Calculus students. Participation was voluntary and all students were required to sign an informed consent form. Some instructors did offer extra credit for participation, but no rewards were guaranteed to participants.

Section 2.2: Study Design

Eighty-three students, all enrolled in Business Calculus, participated in the study. In order to maximize the number of retests, we decided to test at the beginning and end of the fall 2007 semester. These 83 students all took the ILS questionnaire during the first two weeks of the fall semester. Sixty-two of the students who took the test at the beginning of the semester also took the retest during the last week of classes. Between tests there was an 11-week time period. We received a total of 145 completed questionnaires across the two testing periods.

Each of the forty-five questions of the ILS test had two choices of the best answer, and these were labeled as answers (a) and (b). For analyzing the responses we took the same approach as Zywno (2003), and coded each answer of (a) as a +1 quantitative response, and each answer of (b) as a 0 quantitative response.

Section 3: Study Results

Section 3.1: Test-Retest Reliability Results

To assess the reliability of the ILS instrument, we used the results of the 62 students who completed the ILS questionnaire at both the beginning and end of the fall 2007 semester. A set of statistical analyses regarding these paired questionnaires was performed. These measures are similar to those in Zywno (2003) for the most part, and our goal in performing them was to assess whether the conclusions raised in her 2003 article were applicable to Business majors as well.

Table 2 lists Pearson's correlation coefficients for the student's scores on the two ILS questionnaires in each of the dimensions of the ILS instrument. Of the 62 students who completed the ILS questionnaire twice, three students omitted one question. Omission of these students' scores from the correlation analysis resulted in 61 observations for each of the Active, Sensing, and Visual dimensions in Table 2, and 60 observations for the Sequential dimension. All correlations in Table 2 were statistically significant, with p-values for testing the null hypothesis that the population correlation $\rho = 0$ versus $\rho \neq 0$ less than 0.0001 for all four dimensions. Table 2 correlations were highest for the Active and Sensing scores and somewhat lower for the Visual and Sequential Scores. Comparing these correlations to those from Zywno (2003) shows that the two sets of Active, Visual, and Sequential correlations are statistically equal, but our Sensing Correlation value in Table 2 is statistically larger than that in Zywno (2003). These results suggest that, for Business students, the reliability of the ILS instrument is comparable to (and higher than, in the Sensing scale) that of engineering students.

Table 2:
Pearson's Correlation of Test-Retest Scores for the ILS

Dimension	Pearson Correlation
Active Scores	0.71 *
Sensing Scores	0.82 *
Visual Scores	0.65 *
Sequential Scores	0.57 *

* Significant at $\alpha = 0.0001$ level

Table 3 details results of matched-pairs tests for the null hypothesis that the mean difference between each student's scores on their two ILS questionnaires is zero versus a two-sided alternative hypothesis for each of the four test dimensions. We see that none of these differences are statistically significant, which is consistent with reasonable reliability in student responses across the two testing periods.

Table 3:
Matched Pairs Test Results for Test-Retest Scores for the ILS

Variable	Sample Size	Mean	Std. Dev.	T	P-Value (2 tailed)
Active scale	61	-0.26	1.62	-1.25	0.22
Sensing Scale	61	-0.23	1.73	-1.04	0.30
Visual Scale	61	0.12	1.95	0.46	0.63
Sequential Scale	60	-0.13	2.06	-0.50	0.63

Figure 1 gives histograms of the number of identically answered questions on the two ILS questionnaires for each student, for each of the four dimensions, and Table 4 provides summary statistics for these histograms.

Figure 1:
Histograms of Identical Responses for ILS Test-Retest Data

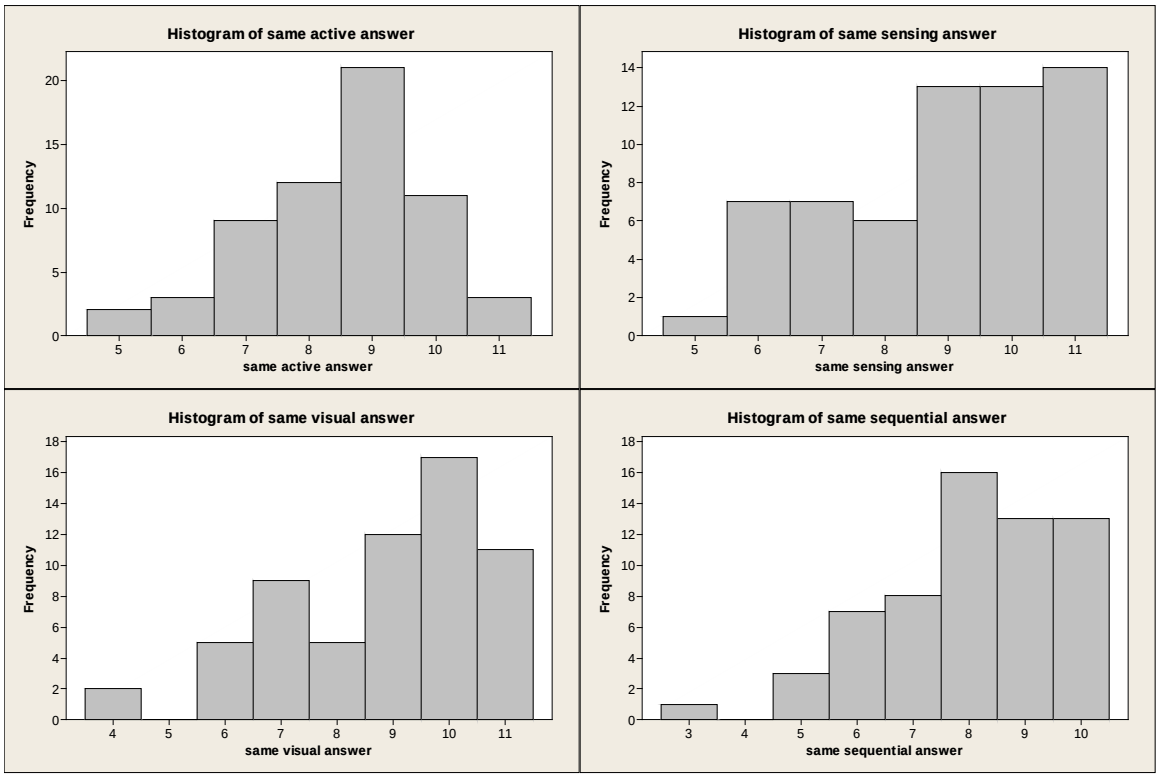


Table 4:
Statistics for the Distributions of Identically Answered Questions on Test-Retest

Statistic	Overall 44 Items	Active Scale 11 Items	Sensing Scale 11 Items	Visual Scale 11 Items	Sequential Scale 11 Items
Mean	34.34	8.51	8.93	8.85	8.05
Std. Dev.	3.28	1.39	1.74	1.81	1.60
Skewness	-0.22	-0.52	-0.50	-0.79	-0.74
Kurtosis	-0.21	0.04	-0.88	-0.02	0.37

The students in our study have higher number, on average, of identically answered test-retest questions than the Engineering students in Zywno (2003). These differences are statistically significant, using a series of Independent Samples T tests. However, the magnitudes of these differences are small, and may be due to the different time spans between re-tests (eight months for the Engineering students versus approximately three months in our study).

Overall this analysis suggests that the ILS instrument is reliable for the measured population. To an extent, these reliability measures are a function of the proximity of the retest to the initial test, and a longer-term study verifying these results over a longer period would be useful.

Section 3.2: Internal Consistency

We used the Cronbach's alpha statistic to measure the consistency of the ILS instrument across the four dimensions. Results of this analysis are given in Table 5. Comparing our results with those in Table 1, the Cronbach's alpha statistics in the last column of Table 5 are consistent with other studies on the Visual-Verbal and Sequential-Global scales. The Cronbach's alpha measure for the Active-Reflective scale in our study is lower than the results of other studies in Table 1. However, the alpha statistic for the Sensing-Intuitive scale in our study is notably higher than those of the other studies listed in Table 1.

Table 5:
Internal Consistency Reliability for the ILS – Cronbach's alpha

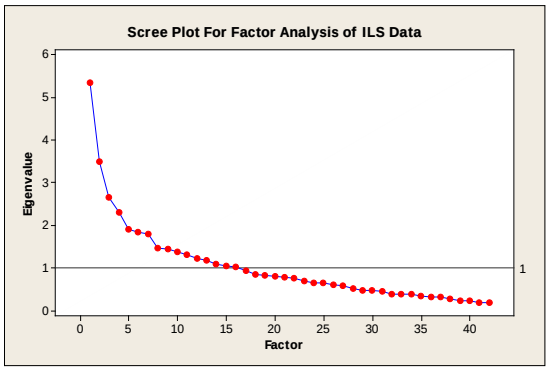
Scale	Cases	Items	Scale Mean	Scale STD	Avg. Inter-Item Correlation	Avg. Item-Total Correlation	Standardized α
Active	144	11	6.49	2.10	0.085	0.409	.506
Sensing	145	11	7.33	2.94	0.284	0.590	.813
Visual	144	11	7.76	2.27	0.154	0.478	.667
Sequential	144	11	6.48	2.18	0.104	0.429	.557

We also performed factor analysis as a method of reducing the dimension of the problem down to a manageable number of common factors that explain the overall variation in the 44 dimensions. Because the ILS questionnaire is theoretically designed to measure variation in learning styles in the four dimensions, we would expect the number of factors that explain the data's variation to reduce to four dimensions.

Results of the factor analyses are summarized in Tables 6 and 7. In Table 6, an orthogonal varimax rotation was chosen, and in Table 7, an oblique (promax) rotation was chosen due to potential dependence of the scales of the ILS with each other. Support for this decision is given in Johnson and Wichern (1998), Section 9.4.

A scree plot (Figure 2) of the factors shows the eigenvalues of the factors leveling off after about the eighth factor, but that the eigenvalues of each factor are greater than one through the first sixteen factors.

Figure 2:
Scree Plot for Factor Analysis of ILS data.



Focusing on the first five factors in Table 6, we see that the Sensing scale emerges in factor 1, the Visual scale emerges in factor 2, and the Active scale emerges in factor 4, although there is some overlap with two other scales. Similarly, the Sequential scale emerges in factor 5 but shows overlap with the Sensing and Visual scales as well. Due to these overlaps, the oblique promax rotation was also performed.

Table 6:
Distribution of High Loading items, Scree Method, Varimax Rotation

Factors	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Active	0	0	0	4	0	0	0	0
Sensing	8	0	3	1	1	1	0	0
Visual	0	6	0	1	1	1	3	0
Sequential	0	0	3	0	3	1	1	2

In Table 7, under the Promax rotation, we see similar results to that of the Varimax rotation. Interpretation of Factors 1 through 3 is similar to that of the Varimax rotation. The only clarification is that the Active scale more clearly is identified as the basis of Factor 4. Finally, the Sequential scale seems to be overlapped a bit with the Sensing scale in factors 3 and 5, before providing the two largest factor loadings in Factor 8. Our conclusions from the factor analysis, then, are similar to those of Zywno (2003) and Van Zwanenberg, et al. (2000).

Table 7:
Distribution of High Loading items, Scree Method, Oblique Rotation

Factors	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Active	0	0	0	4	0	1	0	0
Sensing	4	0	3	0	1	1	1	0
Visual	0	6	0	1	1	1	2	0
Sequential	0	0	3	0	3	1	0	2

Finally, a check of the correlations of each of the factor scores with each other suggests some potential overlap between the Visual and Active scores, and a more notable correlation between the Sensing and Sequential scales (which corresponds to results seen in Tables 6 and 7 for the Factor analysis).

Table 8:
Pearson Correlations for each of the ILS scales; p-values for testing Ho: ρ = 0 versus Ha: ρ ≠ 0 in parentheses

	Active-Reflective	**Sensing-Intuitive**	**Visual-Verbal**	**Sequential-Global**
Active-Reflective	1			
Sensing-Intuitive	-0.07493 (0.3721)	1		
Visual-Verbal	**0.17701 (0.0344)**	-0.04854 (0.5634)	1	
Sequential-Global	-0.07066 (0.4017)	**0.4007 (<0.0001)**	-0.10328 (0.2196)	1

Section 4: Conclusions

We believe this study provides reasonable evidence that, for Business Calculus students, the ILS instrument is a reliable and valid instrument for assessing learning styles. Because different student populations may differ in learning style, the ILS instrument needs to be verified across multiple populations. Our study primarily contributes to the body of work suggesting that the consistency of earlier findings with the ILS instrument as a valid and reliable instrument persists across different student groups.

Remaining questions after our study include multiple test-retest occasions for a student group over a longer period of time. Our study occurred over an 11-week period, and ideally a comparable group of students could be retested over a longer period of time to obtain more evidence for the long-term reliability of the ILS instrument.

As has been concluded in Van Zwanenberg et al. (2000), the ILS tool is not predictive of academic performance. However, valid questions are raised about this conclusion simply because the learning environments of the students in that study were not necessarily homogeneous (Zywno, 2003). We did not use the ILS to predict academic achievement, specifically because the students in our study had a variety of instructors, and each instructor may favor a particular learning style in his or her teaching style. Thus, for our purposes, the usefulness of the ILS tool is the ease to which the instrument can be completed and scored, and in helping students understand their learning styles. In our experience, most students found the ILS easy to complete with minimal questions, and a strong majority (74%) re-took the ILS at the semester's end.

Another limitation of the study pertains to the generalizability of Business Calculus students to the population of all Business majors. Our participants were generally at the beginning of their college careers. Therefore, some attrition from our student group away from Business majors is certainly expected.

We believe that more verification of the ILS is needed across different disciplines. As in other research (Zywno, 2003; Van Zwanenberg et al., 2000), we find that the Sequential

scale does not form the clear basis for one of the primary factors in our factor analysis. More research is needed in this area.

Ultimately, we believe the ILS is very useful in helping students understand their learning style strengths, and in developing strategies to succeed in their courses. The ILS is a valid and reliable instrument, in our study and in others (Zywno, 2003; Van Zwanenberg et al., 2000). This finding, combined with the user-friendly nature of the ILS, makes it a good tool for both students and educators.

References

Construct Validity [On-line] Available:

http://www.socialresearchmethods.net/kb/constval.htm

http://www.uni.edu/chfasoa/reliabilityandvalidity.htm

Index of Learning Styles:

http://www4.ncsu.edu/unity/lockers/users/f/felder/public/ILSpage.html

Felder, R.M. and Silverman, L.K. (1988). "Learning and Teaching Styles in Engineering Educations." *Engineering Education*, 78(7), 674-681.

Felder, R.M. and Spurlin, J. (2005). "Applications, Reliability and Validity of the Index of Learning Styles." *International Journal of Engineering Education*, 21(1), 103-112.

Johnson, R.A. and Wichern, D.W., Applied Multivariate Statistical Analysis, 4th edition, Upper Saddle River, NJ: Prentice Hall, 1998.

Litzinger, T.A., Lee, S.H., and Wise, J.C. (2005). "A Study of the Reliability and Validity of the Felder-Soloman Index of Learning Styles." *Proceedings of the 2005 ASEE Annual Conference*, ASEE.

Tuckman, B.W., Conducting Educational Research, 5th edition, Belmont, CA: Wadsworth Group, 1999.

Van Zwanenberg, N., Wilkinson, L.J., and Anderson, A. (2000) "Felder and Silverman's Index of Learning Styles and Honey and Mumford's Learning Styles Questionnaire: how do they compare and do they predict academic performance?" *Educational Psychology*, 20(3), 365-380.

Zywno, M.S. (2003) "A Contribution to Validation of Score Meaning for Felder Soloman's Index of Learning Styles." *Proceedings of the 2003 ASEE Annual Conference*, ASEE.