A Psychometric Study of the Index of Learning Styles©

THOMAS A. LITZINGER
The Leonhard Center for the Enhancement of Engineering Education
Pennsylvania State University

SANG HA LEE
Educational Psychology
Pennsylvania State University

JOHN C. WISE
Engineering Instructional Services
Pennsylvania State University

RICHARD M. FELDER
Department of Chemical and Biomolecular Engineering
North Carolina State University

ABSTRACT

A study was conducted on the Felder-Soloman Index of Learning Styles© (ILS) to assess reliability, factor structure, and construct validity as well to determine whether changing its dichotomous response scale to a five-option response scale would improve reliability and validity. Data collected in this study had internal consistency reliability ranging from 0.55 to 0.77 across the four learning style scales of the ILS. Factor analysis revealed that multiple factors were present within three of the learning style scales, which correspond to known aspects of the scale definitions. The factor analysis and direct feedback from students on whether they felt their scores accurately represented their learning preferences provide evidence of construct validity for the ILS. Changing the response scale improved reliability, but it did not change the factor structure substantially nor did it affect the strength of the evidence for construct validity based on student feedback.

Key words: learning styles, reliability, validity

I. INTRODUCTION

A. Learning Styles

Learning styles are characteristic preferences for alternative ways of taking in and processing information. The concept arose with the work of Kolb, whose learning styles instrument is credited by some as the first to be created in the U.S. [1] Since the introduction of learning styles, the number of studies using them has grown dramatically. In a recent review of the learning styles literature, Hall and Moseley report that they identified 71 different learning styles instruments and reviewed 800 papers relevant to the post-16 education [2]. In a related document, Coffield et al. report that there are literally thousands of papers related to learning styles in the literature [3]. The use of learning styles in engineering education has become fairly widespread. A search of the term "learning styles" in the ASEE conference proceedings from 2000 through 2006 turned up hundreds of articles covering a wide range of topics including the effects of learning styles on academic performance [4] and on retention [5] as well as on the design of technology-based learning tools [6]. Among the ASEE articles referring to learning styles are nearly 50 that use the Index of Learning Styles© (ILS) of Felder and Soloman [7].

From their review of the literature, Hall and Moseley [2] created a "continuum of families of learning styles" based on the extent to which the developers of the learning style instrument believe that learning styles are fixed traits. They discuss the pedagogical implications of the various theories of learning styles across this continuum. On one end of the continuum are "fixed trait" theories. Practitioners of this class of learning style theories tend to identify the type of student and design a specific curriculum for them; they increasingly look to information technology to create individual learning "prescriptions" for students based on their type. Hall and Moseley note that this approach makes it likely "that learners will accept labels and be reluctant to move beyond their 'comfort zones' to develop new skills." At the opposite end of the continuum are developers who see learning styles as "fluid traits" that can change over time and may depend upon the specific learning task and context. Practitioners of fluid traits approaches do not focus on determining a student's type and matching pedagogy to it, but rather focus on increasing teachers' and students' understandings of their own learning process. Further, practitioners of this approach consider various "strategies, motivation and processes particular to each learning experience."

The fixed trait approaches to learning styles that diagnose styles and design individualized instruction based on it have come under intense criticism in the literature. One of the strongest critics of learning styles used in this manner is Stahl who writes that there has been an "utter failure to find that assessing children's learning styles and matching to instructional methods has any effect on their learning" [8]. In keeping with Stahl's observation, Coffield et al. report that they found little convincing evidence in the literature that individualized instruction based on learning styles leads to significant gains in learning [3]. While the fluid trait approach is open to some of the same general criticisms as the fixed trait approach, e.g., weak evidence for reliability and validity of the related learning style instruments, its potential to have a positive impact on learning was acknowledged by Coffield et al. In summary observations they write: "A reliable and valid instrument which measures learning styles and approaches could be used as a tool to encourage self-development, not only by diagnosing how people learn, but by showing them how to enhance their learning."

The theory and philosophy behind the development and use of the Index of Learning Styles are firmly in the fluid trait category. It was developed based on the belief that the principal value of a learning styles model is to provide guidance to instructors on developing and using a balanced teaching approach. Once a model has been chosen to serve as a basis for instructional design, the instructor's goal should be to make sure that instruction sometimes addresses each learning style preference defined by the model. The appropriate balance between opposite preferences depends on the course subject and on the background and experience of the students taking the course, but there must be a balance.

What learning style assessments should *not* be used for, in our view, is to label individual students for the purposes of prescribing their curriculum or career choices or to draw inferences about their potential ability to succeed at any endeavor. A student's learning style may provide clues about strengths and areas that might call for additional concentration, but no more than that. Students with any learning style preference have the potential to succeed at any endeavor; the fact that a student prefers, say, visual presentation of information implies nothing about his/her ability to process verbal information, or for that matter his/her ability to process visual information.

Most of the published criticisms of learning styles—such as the one by Stahl—are predicated on the assumption that the purpose of learning styles is to facilitate design of individualized instruction that addresses students' preferences, a goal we have categorically rejected. Once again, we believe that teachers should strive for balanced instruction, making sure that the learning needs of students with all preferences are addressed to some extent, an objective that we have not seen criticized by any detractor of the learning style concept. As for the reliability and validity of the Index of Learning Styles, a review of evidence for both was presented in [9] and more is provided in this article.

B. The Index of Learning Styles

The Index of Learning Styles (ILS) is an online questionnaire designed to assess preferences on four dimensions of a learning style model formulated by Felder and Silverman [10]. The ILS consists of four scales, each with 11 items: sensing-intuitive, visual-verbal, active-reflective, and sequential-global. Felder and Spurlin [9] summarize the four scales as follows:

- *sensing* (concrete, practical, oriented toward facts and procedures) or *intuitive* (conceptual, innovative, oriented toward theories and underlying meanings);
- *visual* (prefer visual representations of presented material, such as pictures, diagrams, and flow charts) or *verbal* (prefer written and spoken explanations);
- *active* (learn by trying things out, enjoy working in groups) or *reflective* (learn by thinking things through, prefer working alone or with one or two familiar partners);
- *sequential* (linear thinking process, learn in incremental steps) or *global* (holistic thinking process, learn in large leaps).

As [10] indicates, each scale of the instrument is related to a scale of one or more other instruments (and the sensing-intuitive scale of the ILS is intended to be identical to the same scale of the Myers-Briggs Type Indicator). The combination is unique, however, and was chosen to be consistent with style differences observed by one of the authors in two decades of teaching engineering students.

In the ILS, students complete a sentence by choosing one of two options representing opposite ends of one of the learning style scales. This dichotomous structure was chosen to force a decision between the two alternatives, or as Tuckman says to avoid "fence sitting," [11] thereby increasing the chances that the instrument response will detect preferences. Exclusion of a middle or "no opinion" response option is not uncommon in design of surveys that seek to determine opinions [12]. The number of items on each learning style scale and the scoring method were also chosen such that a score of zero, indicating no preference, was not possible.

C. Research Objectives

The primary objective of this study was to investigate reliability of data collected using the ILS and to seek evidence for the validity of the instrument. An additional objective was to investigate the effect of the dichotomous response format on reliability of the data. The investigation of the effect of response scale structure was motivated by the fact that dichotomous items are likely to result in reduced reliability because two alternatives may be insufficient to discriminate differences consistently [13].

D. Summary of Terminology and Methods

1) Reliability: Reliability can be interpreted as a measure of the extent to which an instrument gives similar results when used in repeated tests [14]. All measurements of the type administered in this study will be affected by the presence of random errors. Thus, when repeated measurements are made some, perhaps even most of the scores for individual subjects will change. It is also likely that the mean score and the variance will be different between different administrations of the instrument. When an instrument with high reliability is used for such repeated testing, the changes in the scores between administrations should be less than the changes that occur when an instrument with lower reliability is used. In this sense the data from the measurement with the higher reliability instrument are considered to be more "reliable."

Classical test theory makes the role of random error in reliability explicit. In this theory, reliability is defined as the variance in the "true" score divided by the sum of the variance in the true score and the variance due to random errors that occur during the measurement. The relationship of reliability to the variance in the measured score makes explicit the fact that reliability is a property of the data from the measurement, rather than a property of the instrument used to make the measurement.

Reliability plays a central role in the development of measurement instruments because adequate reliability is a necessary, but not sufficient, condition for validity [15]. It is also critical to interpretation of data obtained from an instrument because the square root of reliability is the estimated correlation of test scores with the true scores [15]. Additional interpretations and uses of reliability in instrument development are described in [15].

Reliability can be estimated by many approaches [14]. Internal consistency, the approach used in this study, may be the most common of them. It provides an estimate of reliability for a single administration of an instrument and is based on the average correlation between items, hence the name internal consistency. The internal consistency reliability is calculated using the following expression:

$$r = \frac{N\overline{\rho}}{1+\overline{\rho}(N-1)}$$

where N is the number of items and $\overline{\rho}$ is the average correlation between items. It can also be calculated from variances using the following equation, generally referred to as Cronbach's coefficient alpha:

$$\alpha = \frac{N}{(N-1)}\left[1 - \frac{\sum \sigma_i^2}{\sigma_x^2}\right],$$

where $\Sigma\sigma_i^2$ is the sum of the variances of individual items, and σ_x^2 is the variance in the score. For dichotomous items, Cronbach's coefficient alpha takes on a special form:

$$\alpha = \frac{N}{(N-1)}\left[1 - \frac{\sum p_i(1-p_i)}{\sigma_x^2}\right]$$

where p_i is the fraction of the students answering "positively." [14]

2) Validity: In the introduction to his chapter on validity, Nunnally writes, "In a very general sense, a measuring instrument is valid if it does what it is intended to do" [15]. Later in the same chapter he notes that "Strictly speaking, one validates not a measuring instrument but rather some use to which it is put." In Standards for Educational and Psychological Testing published by the American Psychological Association, validity is defined as "the appropriateness, meaningfulness, and usefulness of the specific inferences made from test scores" [16]. Still it is common to speak of validating an instrument.

Many types of validity are discussed in the literature; the most important for this study is construct-related validity, often referred to as simply construct validity. In the social science measurement literature, "construct" refers to the thing that is to be measured; the term acknowledges the fact that in most cases what is being measured is an abstraction "constructed" in the mind of the researcher as opposed to being concrete and readily observable [15]. Evidence of construct validity means evidence that the instrument is measuring the construct that it is intended to measure, so that inferences drawn on the basis of measurement are valid.

In his discussion of instrument validation, Nunnally states that "there is no way to prove validity of an instrument purely by appeal to authority, deduction from a psychological theory, or any type of mathematical proof…validation is an unending process." Thus, the process of validating an instrument consists of gathering evidence over many studies; no single study should be considered as a complete proof of validity.

3) Factor Analysis: Factor analysis is performed to identify clusters of items for which responses have common patterns of variation. "Each such cluster, or factor, is denoted by a group of variables whose members correlate more highly among themselves than they do with variables not included in the cluster" [15]. Factor analysis assumes that responses to individual items in an instrument are linear combinations of the factors and it produces a factor model that relates the item responses to the factors in linear combinations.

There are two types of factor analysis, confirmatory and exploratory [17]. In confirmatory factor analysis, substantive constraints, derived from a theory related to the construct being measured, are placed on the factor model. These constraints include selecting the number of factors, defining which factors are correlated with each other, and which factors affect which measured variables. Factor analysis is then performed to "confirm" that the model is correct. In exploratory factor analysis, no *a priori* constraints are placed on the factor model. Consequently, in exploratory factor analysis, interpretation of the results is required to identify the most appropriate factor model. (Readers interested in an introduction to factor analysis are referred to [18].)

II. METHODOLOGY

A. Instruments

The five-option response scale for the modified ILS is illustrated in Figure 1 along with the corresponding item from the ILS. The new scale, which is similar to a five-point Likert scale, makes two changes to the original scale. It introduces a neutral response option for those who feel that they have no preference, and it offers two levels of strength of preference. Although its options are similar to a Likert scale, the new scale does not use the traditional Strongly Agree to Strongly Disagree descriptors due to the nature of the ILS. In the items of the ILS, students complete a sentence by choosing one of two options representing opposite ends of one of the learning style scales. With the modified scale, the students were instructed to consider the scale as a continuum and to mark where they fell on the continuum by selecting one of the five positions. The specific instructions for the modified instrument were: "Each of the following statements has two possible endings that represent opposite ends of a continuous scale. Please mark the box that best indicates where you fall on the scale."

Evidence for construct validity was obtained by giving students feedback based on their scores and asking them to assess the match

Original ILS

I understand something better after I

☐ (a) try it out.

☐ (b) think it through.

Modified ILS

I understand something better after I

try it out. ☐ ☐ ☐ ☐ ☐ think it through.

Figure 1. Format of items in ILS and modified ILS.

between the measured style and their perception of their style. The feedback provided to the students was a short description of the characteristics of their measured learning style. Separate feedback was written for each of the eight style categories within the ILS, two for each of the four scales. For each of the eight categories, different feedback was written for students with a moderate or strong preference, scores 5-7 or 9-11 on the ILS, and for those who were balanced in their style, scores of 1 or 3. In all, sixteen different descriptions of learning style categories were written. In order to assess the match between their measured style and their perception of their style, students were asked to answer the following question: "How well do these statements describe your learning style?" The scale for their responses to this question included three options: Very well/just like me, Fairly well/somewhat like me, and Poorly/not at all like me.

An example of the information provided to the students with a strong or moderate preference for visual learning is presented in Figure 2. For students who did not have a moderate or strong preference, the description of the learning style was modified to reflect a balance between the two styles. The feedback provided to a student who was balanced on the Visual-Verbal scale was: *Your scores on the survey seem to say that you have no clear preference for either verbal or visual learning situations. You may be just as good at getting information from pictures, charts, and graphs as you are at getting information through your reading.* Students were asked to respond to the same question on how well the description matched their style.

Prior to use of the modified response scale and the feedback, a focus group was run with nine undergraduates who completed both forms of the instrument. The group had some suggestions for improvement of the feedback that were incorporated into the final design. A Web site was then prepared with the two forms of the instrument and the feedback, and data collection was initiated.

B. Participants

Random samples of 1,000 students from three colleges, Engineering, Liberal Arts, and Education, were invited to participate in the study via email; both undergraduate and graduate students were included in the samples. Five hundred students in each random sample were asked to complete the ILS first and 500 were asked to complete the modified ILS first. Participants were informed that if they completed the instrument that they would be contacted two weeks later to ask them to complete the second form of the instrument. The incentive provided for participation was entry into a random drawing for $100; one such incentive was awarded in each college. In an attempt to increase the return rate, i.e., the number of students who completed both forms, students were awarded one entry in the random drawings for completing the first form of the ILS and two additional entries if they completed the second form.

In the first round of surveys, a total of 710 students completed at least one form of the instrument; 371 students did the original form of the ILS and 339 did the modified ILS. In the second round, 233 students who received the ILS first and 215 students who received the modified ILS first returned to complete the other form. Thus, a total of 586 students completed the ILS and 572 completed the modified ILS. The return rates were 62.8 percent for the ILS first group and 63.4 percent for modified ILS first group. A Chi-square test indicates that return rates for the two groups are not statistically different (Chi-square = 0.029, Degrees of Freedom = 1, p-value = 0.865). Therefore the form of the response scale, dichotomous versus five-options, did not affect the return rates.

The total number of students who completed both forms of the instrument was 448, giving an overall response rate of 15 percent. Approximately 80 percent of participants were undergraduate students and approximately 50 percent of the participants were female. Students in engineering participated at the highest rate of the three colleges, 39 percent, perhaps because the study originated in engineering.

C. Analysis

Analysis of the data from the two forms of the ILS included response statistics, checks of internal consistency of the four scales using Cronbach's coefficient alpha, and an exploratory factor analysis to identify the components within each scale. The exploratory factor analysis was conducted within the SPSS program using principal component analysis and orthogonal Varimax rotation. The factor analysis results and student feedback were used to seek evidence for construct validity of the ILS scales.

III. Results and Discussion

A. Effect of Scale on Response Characteristics and Scores

The modified ILS response scale included a neutral response option that is not present in the ILS. Across all 44 items, the neutral response accounted for 16 percent or approximately 1 in 7 responses. The use of the neutral response on individual items ranged from zero to 28 percent. The item for which not a single student used the neutral response was in the Active-Reflective scale: "In a study group

According to the survey, you may have a preference for "visual" learning. The following statements describe some characteristics of visual learners:

- I prefer visual demonstrations over verbal explanations.
- I remember best what I see: pictures, diagrams, flow charts, time lines, films, and demonstrations, for example.
- I am much less likely to remember what I hear or text that I read.

How well do these statements describe your learning style?

Very Well/Just like me	Fairly Well/ Somewhat like me	Poorly/Not at all like me
○	○	○

Figure 2. Sample feedback for students with moderate or strong preference.

working on difficult material, I am more likely to: a) jump in and contribute ideas or b) sit back and listen." For eight of the items, use of the neutral response exceeded 20 percent. Four of those eight items were from the Sensing-Intuitive scale, which had the highest usage of the neutral response at 19 percent. For the remaining scales, neutral response usage ranged from 14 to 17 percent.

The other change introduced by the modified ILS response scale was addition of a moderate strength of preference response. Table 1 presents the fraction of responses across the three options: strong, moderate, and no preference, for the four scales. The pattern of usage on the Active-Reflective and Visual-Verbal scales is relatively similar; those expressing a preference use the strong and moderate preference options in roughly equal proportions. For the remaining two scales, the moderate preference response was used at a somewhat higher rate than the strong preference response. The significant usage of the moderate response option along with the use of the neutral response suggest that fewer scores for the modified ILS will be in the strong preference category.

	Act-Refl	Sen-Int	Vis-Verb	Seq-Glob
Strong preference	42%	36%	41%	36%
Moderate Preference	44%	45%	42%	48%
No Preference	14%	19%	17%	16%

Table 1. Fraction of responses in the three response categories for the modified ILS.

Learning style scores on the ILS are obtained by taking the difference of the number of items chosen for each of the two styles in a scale. For example, a student who selects 10 visual items and 1 verbal item would have a score of 9 on the visual scale. This scoring approach limits scores to odd numbers from −11 to 11. The modified ILS is scored by assigning a value of zero for the neutral position on the response scale and values to 0.5 and 1 for the moderate and strong preference options, respectively. The scale score is determined just as it is for the ILS, by subtracting the total points for each of the two learning styles in a scale. For the modified ILS, the scale scores have a range of −11 to 11 in increments of 0.5.

Score distributions for the visual-verbal scale on the two forms of the ILS are illustrated in Figure 3; the top histogram is that for the ILS and the bottom is for the modified ILS. Comparison of histograms shows that fewer students have scores indicating strong visual preference on the modified ILS; consistent with the expectation based on the patterns of use of the no preference and moderate preference responses. For the modified ILS, only 4 percent of the students have scores of 9 or higher whereas for the ILS, 21 percent of the students have scores of 9 or higher. Inspection of responses on the other three scales shows similar trends.

Table 2 compares the mean and variance of the scale scores for the two forms of the ILS. With the exception of the Active-Reflective scale for which the means are close to zero, the modified instrument has smaller mean scale scores. However, none of the

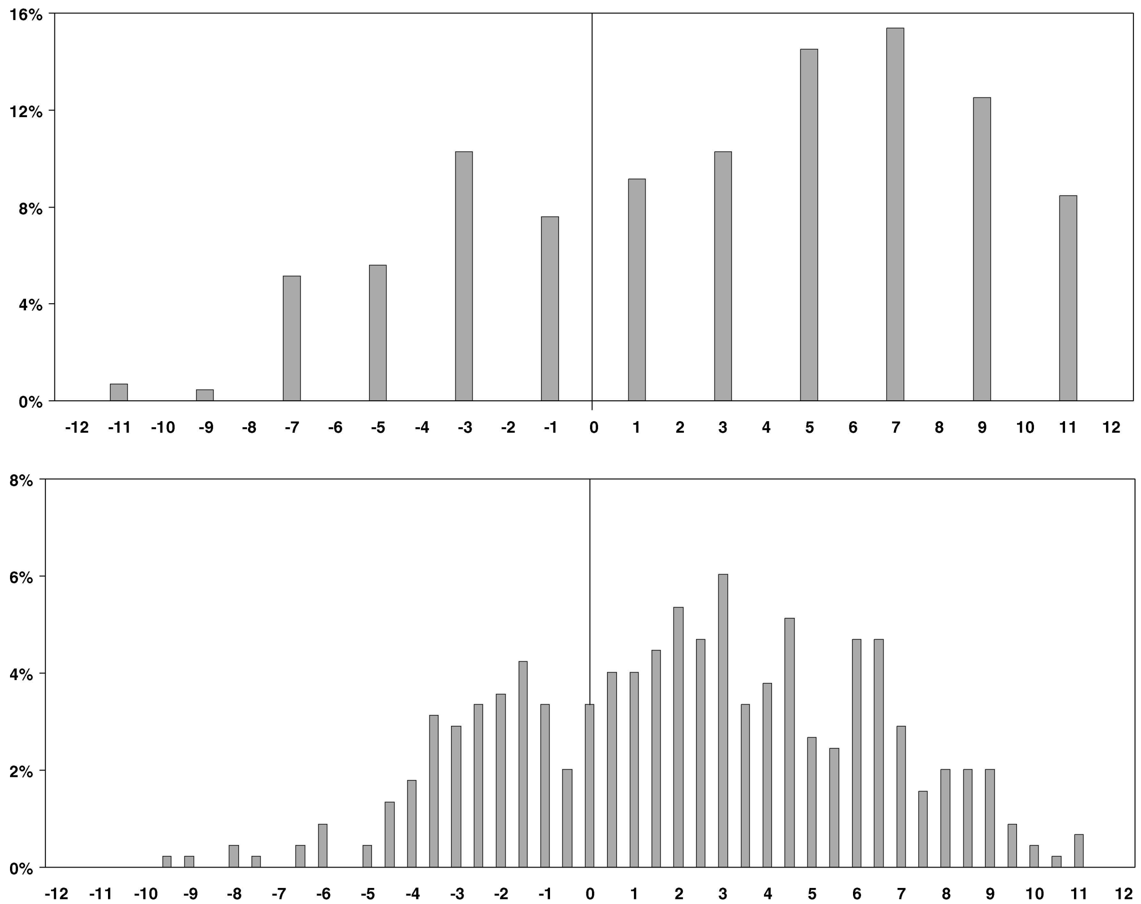

Figure 3. Histograms of scores for Visual (+)/Verbal scale.

differences in the means are statistically significant. The variances in the scale scores for the modified ILS are substantially smaller than those of the ILS; F-tests showed that the differences in the variances are statistically significant with p-values less than 0.01. Thus, the two forms of the instrument have equivalent means, but the modified ILS has narrower distributions of scores.

B. Internal Consistency Reliability

Table 3 presents the internal consistency reliability coefficients for the ILS from the current study along with past studies reported by Felder and Spurlin [9]. Because reliability is a characteristic of the measurement on the specific sample under study, values of reliability coefficients for the ILS vary among the studies. The Cronbach's coefficient alpha values obtained in this study show a similar pattern to that of past studies and are comparable in magnitude to the values obtained for three of the four scales. The Sensing-Intuitive (S-N) scale and the Visual-Verbal (V-V) scale generate data with reliability in excess of 0.7, whereas data from the Active-Reflective (A-R) and Sequential-Global (S-G) scales had reliability coefficients of 0.61 and 0.55, respectively. The reliability coefficients all exceed the minimum standard of 0.50 suggested by Tuckman for attitude and preference assessments [11].

The values of Cronbach's alpha reported for this study in Table 3 are slightly different than those reported previously [19]. The previously published values were calculated using all of the 572 students who completed the ILS. In order to compare the internal consistency reliability of the two forms of the ILS used in this study, it was necessary to use identical samples because reliability is related to the sample taking the instrument. Therefore, Cronbach's coefficient alphas in Table 3 are estimated based on the 448 students who completed both ILS forms.

Table 4 compares the Cronbach's alpha values for the ILS and the modified ILS. As expected, the changing from the dichotomous to the five-option response scale increases the reliability on all four learning style scales. A common, alternative approach to increasing reliability is to increase the number of items in an instrument [15]. Estimates of the number of items that would have to be added to the ILS to match the reliability of the modified ILS are also presented in Table 4. The required number of items gives an indication of the substantial improvement in reliability.

The number of items is calculated based upon the Spearman-Brown prophecy formula, which relates the reliability of the new test (r_{New}) with a different number of items, to the reliability of the existing test (r_{Old}). The Spearman-Brown prophecy formula is

$$r_{New} = \frac{k \times r_{Old}}{1 + (k-1)r_{Old}} = \frac{k \times r_{Old}}{k \times r_{Old} + (1 - r_{Old})},$$

		Mean	Variance
Active(+)	ILS	0.04	22.6
/Reflective	Modified ILS	0.32	13.7
Sensing (+)	ILS	1.21	33.7
Intuitive	Modified ILS	0.70	19.1
Visual (+)	ILS	3.13	29.0
Verbal	Modified ILS	2.16	15.8
Sequential(+)	ILS	0.74	19.8
Global	Modified ILS	0.47	10.5

Table 2. Mean and variance of scale scores for the two forms.

Act-Refl	Sen-Int	Vis-Verb	Seq-Glob	N	*Source*
0.61	0.77	0.76	0.55	448	Current Study
0.56	0.72	0.60	0.54	242	Livesay *et al.* [20]
0.51	0.65	0.56	0.41	284	Van Zwanenberg *et al.* [21]
0.60	0.70	0.63	0.53	557	Zywno [22]

Table 3. Cronbach's alpha coefficients for the ILS.

	Cronbach's alpha			Additional items required to achieve the same increase in reliability
	ILS	Modified ILS	Ratio	
Active-Reflective	0.61	0.69	1.42	5
Sensing-Intuitive	0.76	0.82	1.38	4
Visual-Verbal	0.75	0.77	1.06	1
Sequential-Global	0.55	0.60	1.26	3

Table 4. Comparison of Internal Consistency Reliability.

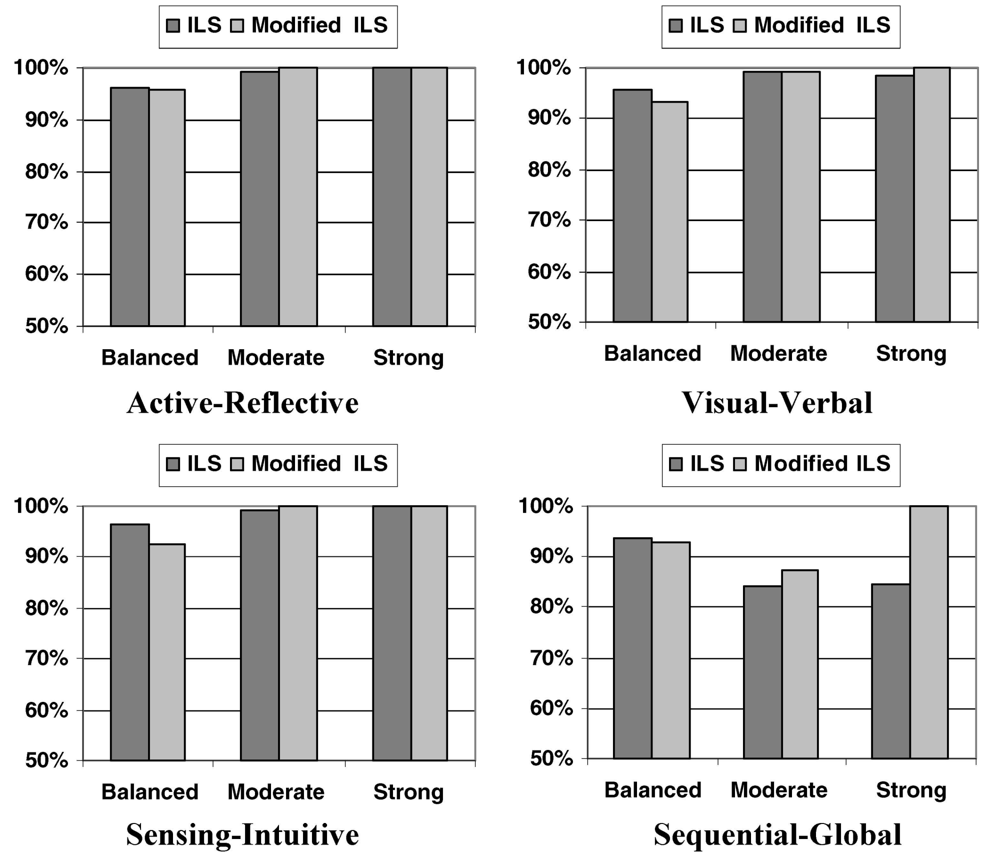

Figure 4. Percentage of students indicating that the description of their learning style matched their preferences very well or fairly well.

where *k* is the ratio of the number of items in the new test to the number of items in the old test [23]. The values of reliability for the ILS and modified ILS were used to determine k for each scale. Then (k-1) was multiplied by the number of items in each scale of the ILS (11) to determine the entry in the last column of Table 4. Thus as many as five additional items would be required to achieve the same increase in reliability obtained by modifying the response scale.

C. Construct-Related Evidence of Validity

1) Factor Analysis: Exploratory factor analysis was performed to determine the factor structure of the four learning style scales for both forms of the instrument. Results of the exploratory factor analysis of the ILS were reported previously [19]. In that analysis, solutions from four to eight factors were considered. A four factor solution was investigated first because the ILS is based on four learning style scales. The four factor structure showed poor loading of items onto their intended scales, indicating that more than four factors were present. Trends in the clustering of the items as the number of factors was increased led to the conclusion that the eight factor solution was the most meaningful.

Table 5 presents a summary of the items in each factor along with a description of the factors; a copy of the ILS is included in the appendix so that items can be linked to the item numbers. The Sensing-Intuitive items all loaded onto one factor; although three items, 22, 30, and 42, were not most strongly loaded on that factor. (In Table 5, the italicized items are those that did not have their dominant loading onto that factor.) The other scales were found to relate to more than one factor. For the Visual-Verbal scale, all items loaded onto two factors, although item 39 did not have its largest loading on either factor. The Global-Sequential scale also contains two factors and the Active-Reflective scale contains three factors.

The factor structure provides evidence of construct validity for the ILS. For the Sensing-Intuitive scale, all items load onto a single factor with a clear relationship to the sensing-intuitive scale: preference for concrete or abstract information. For the Visual-Verbal scale the evidence of construct validity is also good as there are two factors and they are both related to preference for visual or verbal information. For the Active-Reflective and Sequential-Global scales, the identified factors also appear to be appropriate.

Exploratory factor analysis of the modified ILS also investigated four to eight factor solutions; the eight-factor solution was selected to allow the most direct comparison to the factor structure of the ILS. For the Sensing-Intuitive scale, all eleven items were loaded onto a single factor as they were for the ILS. For the modified ILS, however, all items except for item 42 had their dominant loading on this factor. For the other three scales of the modified ILS, the basic factor structure was the same as that of the ILS, and all items loaded onto factors similar to those given in Table 5. However, the dominant loading of some items switched to a different factor within the scale. The loading of items 39 and 40 did not improve compared to the ILS factor analysis. Overall, the use of the five-option response scale did not have a major impact on the factor structure of the ILS, which indicates that the two forms of the instrument are measuring the same constructs.

2) Students' Perception of their Learning Style: The assessment by the students of whether their measured learning styles match their perception of their styles provides additional evidence for construct

Scale	Items	Factors
Sensing - Intuitive	38, 6, 18, 14, 2, 10, 34, 26, *22, 42, 30*	Preference for concrete information (facts, data, the "real world") or abstraction (interpretations, theories, models)
Visual - Verbal	7, 31, 23, 11, 15	Information format preferred for input
	27, 19, 3, 35, 43, *39*	Information format preferred for memory or recall
Sequential - Global	20, 36, 44, 8, 12, 32, 24	Linear/sequential or random/holistic thinking
	28, 4, 16, *40*	Emphasize details (the trees) or the big picture (the forest)
Active - Reflective	25, 1, 29, 5, 17	Action-first or reflection-first
	37, 13, 9	Outgoing or reserved
	21, 33, 41	Favorable or unfavorable attitude toward group work

Table 5. Factors in the eight factor solution for the ILS [19].

validity. Out of 572 students who completed the ILS, 354 evaluated their measured learning style, and 436 out of 586 who completed the modified ILS evaluated their styles. The plots in Figure 4 show the percentage of students in a given response category, i.e., balanced, moderate, and strong, who felt that the description provided matched them fairly well or very well. In these plots, the feedback for the modified ILS is 'binned' in precisely the same way as the original instrument, i.e., balanced: 1 to 3, moderate: 5 to 7, and strong: 9 to 11, to make the most fair comparison. For the modified ILS, students whose scores fell between these bands were not included, consequently not all of the 436 students who provided feedback on the modified ILS are represented in the plots.

For all of the learning style scales, the data provide strong evidence for construct validity of the ILS. For the Active-Reflective, Sensing-Intuitive, and Visual-Verbal scales, 90 percent or more of the students indicated that the learning style description matched them fairly well or very well. For the Sequential-Global scale the agreement is not as good, but it is still in excess of 80 percent for all three strengths of preference.

Comparing the data from the two different forms of the instrument shows that there is little effect on the fraction of students who agree that their measured style matches their perception of their style. Only the data for the Sequential-Global scale suggest that the modified ILS may be improving the validity; however, the numbers of students in the moderate and strong categories on this scale are too small for the differences to be significant. Therefore it does not appear that the increased reliability associated with the use of a five-option response scale led to any improvement in validity. To some extent the lack of an effect was not unexpected because the items were identical between the two forms of the instrument. On the other hand, increased reliability is associated with less error in the measurement, which could have enhanced validity.

IV. Conclusions

The work presented in this paper was conducted with two goals in mind: to assess evidence for reliability and validity of the Index of Learning Styles, and to determine the effects of modifying the response scale in a way that was expected to increase reliability. The results show that the original version of the ILS generates data with acceptable levels of internal consistency reliability, and that evidence for its construct validity from both factor analysis and student feedback is strong. The modification of the dichotomous response scale format to a five-option scale did not change the mean scores on the four learning style dimensions, but it did result in statistically significant reductions in the standard deviations of the scores for all scales and in substantial improvements in internal consistency reliability for three of the four scales. On the other hand, neither the construct validity nor the factor structure of the instrument was strengthened by the modification of the response format. Furthermore, the modification of the response scale resulted in narrower distributions of scores meaning that more students had scores in the balanced category and fewer had scores in the strong preference category, raising the possibility that strength of assessed preferences could be underestimated by the modified response scale.

If the ILS were a new instrument, further work to investigate this "centering" tendency of the modified ILS response scale, and its possible effects on validity would be in order, as would work to improve some of the items that do not fit well into the factor structure of the instrument. However, the ILS is not a new instrument and has a substantial history of use. Given that it generates data with satisfactory internal consistency reliability and that evidence for its validity is strong, changing the instrument, even to improve it, seems ill-advised. Therefore, we have decided not to move forward with a revision of the ILS based on the results of this study.

Acknowledgment

The authors would like to acknowledge the contributions of the reviewers to this paper; we believe that their comments led to substantive improvements in the final manuscript. Support for this work was provided by the Leonhard Center for the Enhancement of Engineering Education at Penn State.

REFERENCES

[1] Delahoussaye, M., "The Perfect Learner: An Expert Debate on Learning Styles," *Training* Vol. 39, No. 5, 2004, pp. 28-36.

[2] Hall, E. and D. Moseley, "Is There a Role for Learning Styles in Personalised Education and Training?," *International Journal of Lifelong Learning*, Vol. 24, No. 3, 2005, pp. 243-255.

[3] Coffield, F., D. Moseley, E. Hall, and K. Ecclestone, *Learning Styles and Pedagogy in Post-16 Education: A Systematic and Critical Review*, The Learning and Skills Research Centre, London, 2004, http://www.lsda.org.uk/files/PDF/1543.pdf, accessed 25 April 2007.

[4] Gabriele, G., D. Kaminski, B. Lister, and P. Théroux, "Effect Of Learning Style On Academic Performance In An Introductory Thermal-Fluids Course, *Proceedings, 2005 ASEE Annual Conference and Exposition*.

[5] Camp, C., S. Ivey, L. Lackey, A. Lambert, J. Marchetta, and A. Robinson, "Learning Styles And Freshman Retention: What Are The Links?," *Proceedings, 2005 ASEE Annual Conference and Exposition*.

[6] Anderson, E., N. Chandrashekar, J. Hashemi, and S. Kholamkar, "Web-Based Delivery Of Laboratory Experiments And Its Effectiveness Based On Student Learning Style," *Proceedings, 2006 ASEE Annual Conference and Exposition*.

[7] Felder, R.M., and B.A. Soloman, *Index of Learning Styles*, *http://www.ncsu.edu/felder-public/ILSpage.html*, 2004, accessed February 15, 2006.

[8] Stahl, S.A., "Different strokes for different folks?" in L. Abbeduto (Ed.), *Taking sides: Clashing on Controversial Issues in Educational Psychology,* Guilford, CT: McGraw-Hill, 2002.

[9] Felder, R.M., and J. Spurlin, "Reliability and Validity of the Index of Learning Styles: A Meta-analysis," *International Journal of Engineering Education,* Vol. 21, No. 1, 2005, pp. 103-112.

[10] Felder, R.M., and L.K. Silverman, "Learning and Teaching Styles in Engineering Education," *Engineering Education,* Vol. 78, No. 7, 1988, pp. 674-681.

[11] Tuckman, B.W., *Conducting Educational Research*, 5th ed., Fort Worth, TX: Harcourt Brace Publishers, 1999.

[12] Converse, J.M., and S. Presser, *Survey Questions: Handcrafting the Standardized Questionnaire,* Newbury Park, CA: Sage Publications, 1986.

[13] Spector, P.E., *Summated Rating Scale Construction: An Introduction*, Newbury Park, CA: Sage Publications, 1992.

[14] Carmines, E.G., and Zeller, R.A., *Reliability and Validity Assessment*, Newbury Park, CA: Sage Publications, 1979.

[15] Nunnally, J.C., *Psychometric Theory, 2nd Edition,* New York, NY: McGraw-Hill, 1978.

[16] American Educational Research Association, American Psychological Association, and National Council on Measurement, *Standards for Educational and Psychological Testing*, Washington, DC: American Psychological Association, 1985.

[17] Long, J.S., *Confirmatory Factor Analysis*, Newbury Park, CA: Sage Publications, 1983.

[18] Kim, J.O., and C.W. Mueller, *Introduction to Factor Analysis: What It Is and How You Do It*, Newbury Park, CA: Sage Publications, 1979.

[19] Litzinger, T.A., Lee, S.H.; Wise, J.C., and Felder, R.M., "A Study of the Reliability and Validity of the Felder-Soloman Index of Learning Styles," *Proceedings, 2005 ASEE Annual Conference and Exposition*.

[20] Livesay, G.A., K.C. Dee, E.A. Nauman, and L.S. Hites, Jr., "Engineering Student Learning Styles: A Statistical Analysis Using Felder's Index of Learning Styles," *Proceedings, 2002 ASEE Annual Conference and Exposition*.

[21] Van Zwanenberg, N., and L.J. Wilkinson, "Felder and Silverman's Index of Learning Styles and Honey and Mumford's Learning Styles Questionnaire: How Do They Compare and How Do They Predict?" *Educational Psychology,* Vol. 20, No. 3, 2000, pp. 365-381.

[22] Zywno, M.S., "A Contribution of Validation of Score Meaning for Felder-Soloman's Index of Learning Styles," *Proceedings, 2003 Annual ASEE Conference and Exposition*.

[23] Allen, M.J., and W.M. Yen, *Introduction to Measurement Theory*, Monterey, CA: Brooks/Cole Publishing Company, 1979.

AUTHORS' BIOGRAPHIES

Thomas A. Litzinger is director of the Leonhard Center for the Enhancement of Engineering Education and professor of Mechanical Engineering at Penn State, where he has been on the faculty since 1985. His work in engineering education involves curricular reform, teaching and learning innovations, faculty development, and assessment.

Address: 201 Hammond Building, University Park, PA 16802; telephone: (+1) 814.865.4015; e-mail: tal2@psu.edu.

John C. Wise was director of Engineering Instructional Services within the College of Engineering at Penn State, when this study was performed. He is currently employed in the College of Education at Penn State as associate director of the Regional Education Lab - Mid-Atlantic. He received his B.A. in Liberal Arts from The University of the State of New York and his M.S. and Ph.D. in Instructional Systems from Penn State.

Address: 108A Rackley Building, University Park, PA 16802; e-mail: jwise@psu.edu.

Sang Ha Lee was a doctoral candidate in Educational Psychology at Penn State during the study described in this paper. He completed his doctoral studies and is now employed by Korea Institute of Curriculum & Evaluation. He earned his M. Ed. and B.S. in Mathematics Education from the Korea National University of Education and taught mathematics in Korea for eight years.

Richard M. Felder is Hoechst Celanese Professor Emeritus of Chemical Engineering at North Carolina State University. He is the co-developer of both the Index of Learning Styles and the learning style model upon which it is based. Many of his papers on engineering education may be found at *http://www.ncsu.edu/felder-public.*

Address: North Carolina State University, Department of Chemical and Biomolecular Engineering, Engineering Building I, Box 7905, 911 Partners Way, Raleigh, NC 27695; e-mail: rmfelder@mindspring.com.

APPENDIX

Index of Learning Styles © 1996, North Carolina State University

1. I understand something better after I
   (a) try it out.
   (b) think it through.
2. I would rather be considered
   (a) realistic.

(b) innovative.
3. When I think about what I did yesterday, I am most likely to get
   (a) a picture.
   (b) words.
4. I tend to
   (a) understand details of a subject but may be fuzzy about its overall structure.
   (b) understand the overall structure but may be fuzzy about details.
5. When I am learning something new, it helps me to
   (a) talk about it.
   (b) think about it.
6. If I were a teacher, I would rather teach a course
   (a) that deals with facts and real life situations.
   (b) that deals with ideas and theories.
7. I prefer to get new information in
   (a) pictures, diagrams, graphs, or maps.
   (b) written directions or verbal information.
8. Once I understand
   (a) all the parts, I understand the whole thing.
   (b) the whole thing, I see how the parts fit.
9. In a study group working on difficult material, I am more likely to
   (a) jump in and contribute ideas.
   (b) sit back and listen.
10. I find it easier
    (a) to learn facts.
    (b) to learn concepts.
11. In a book with lots of pictures and charts, I am likely to
    (a) look over the pictures and charts carefully.
    (b) focus on the written text.
12. When I solve math problems
    (a) I usually work my way to the solutions one step at a time.
    (b) I often just see the solutions but then have to struggle to figure out the steps to get to them.
13. In classes I have taken
    (a) I have usually gotten to know many of the students.
    (b) I have rarely gotten to know many of the students.
14. In reading nonfiction, I prefer
    (a) something that teaches me new facts or tells me how to do something.
    (b) something that gives me new ideas to think about.
15. I like teachers
    (a) who put a lot of diagrams on the board.
    (b) who spend a lot of time explaining.
16. When I'm analyzing a story or a novel
    (a) I think of the incidents and try to put them together to figure out the themes.
    (b) I just know what the themes are when I finish reading and then I have to go back and find the incidents that demonstrate them.
17. When I start a homework problem, I am more likely to
    (a) start working on the solution immediately.
    (b) try to fully understand the problem first.
18. I prefer the idea of
    (a) certainty.
    (b) theory.
19. I remember best
    (a) what I see.
    (b) what I hear.
20. It is more important to me that an instructor
    (a) lay out the material in clear sequential steps.
    (b) give me an overall picture and relate the material to other subjects.
21. I prefer to study
    (a) in a study group.
    (b) alone.
22. I am more likely to be considered
    (a) careful about the details of my work.
    (b) creative about how to do my work.
23. When I get directions to a new place, I prefer
    (a) a map.
    (b) written instructions.
24. I learn
    (a) at a fairly regular pace. If I study hard, I'll "get it."
    (b) in fits and starts. I'll be totally confused and then suddenly it all "clicks."
25. I would rather first
    (a) try things out.
    (b) think about how I'm going to do it.
26. When I am reading for enjoyment, I like writers to
    (a) clearly say what they mean.
    (b) say things in creative, interesting ways.
27. When I see a diagram or sketch in class, I am most likely to remember
    (a) the picture.
    (b) what the instructor said about it.
28. When considering a body of information, I am more likely to
    (a) focus on details and miss the big picture.
    (b) try to understand the big picture before getting into the details.
29. I more easily remember
    (a) something I have done.
    (b) something I have thought a lot about.
30. When I have to perform a task, I prefer to
    (a) master one way of doing it.
    (b) come up with new ways of doing it.
31. When someone is showing me data, I prefer
    (a) charts or graphs.
    (b) text summarizing the results.
32. When writing a paper, I am more likely to
    (a) work on (think about or write) the beginning of the paper and progress forward.
    (b) work on (think about or write) different parts of the paper and then order them.
33. When I have to work on a group project, I first want to
    (a) have "group brainstorming" where everyone contributes ideas.
    (b) brainstorm individually and then come together as a group to compare ideas.
34. I consider it higher praise to call someone
    (a) sensible.
    (b) imaginative.
35. When I meet people at a party, I am more likely to remember
    (a) what they looked like.
    (b) what they said about themselves.

36. When I am learning a new subject, I prefer to
   (a) stay focused on that subject, learning as much about it as I can.
   (b) try to make connections between that subject and related subjects.
37. I am more likely to be considered
   (a) outgoing.
   (b) reserved.
38. I prefer courses that emphasize
   (a) concrete material (facts, data).
   (b) abstract material (concepts, theories).
39. For entertainment, I would rather
   (a) watch television.
   (b) read a book.
40. Some teachers start their lectures with an outline of what they will cover. Such outlines are
   (a) somewhat helpful to me.
   (b) very helpful to me.
41. The idea of doing homework in groups, with one grade for the entire group,
   (a) appeals to me.
   (b) does not appeal to me.
42. When I am doing long calculations,
   (a) I tend to repeat all my steps and check my work carefully.
   (b) I find checking my work tiresome and have to force myself to do it.
43. I tend to picture places I have been
   (a) easily and fairly accurately.
   (b) with difficulty and without much detail.
44. When solving problems in a group, I would be more likely to
   (a) think of the steps in the solution process.
   (b) think of possible consequences or applications of the solution in a wide range of areas.