Spring 1994

Further Comments on the Use of the LSI in Research on Student Performance in Introductory Accounting

Marshall A. Geiger
University of Richmond, mgeiger@richmond.edu

Edmund J. Boyle

Accounting Educators' Journal
Volume 6, Number 1

Further Comments on the Use of the LSI in Research on Student Performance in Introductory Accounting

Marshall A. Geiger
Edmund J. Boyle
University of Rhode Island

ABSTRACT

In this note we respond to the comments of Ruble and Stout (1993) concerning the use of the revised Learning Style Inventory (LSI) developed by Kolb (1985). While our more recent psychometric research on the LSI leads us to conclude that the standard version should no longer be used, unlike Ruble and Stout, we see promise for new or modified versions in future research. We also indicate where several of their comments on our work, as well as the work of others, are not well founded.

Introduction

In this note we respond to Ruble and Stout (1993), hereafter R&S, concerning our earlier research: Geiger (1992) and Geiger and Boyle (1992). These two studies utilized Kolb's (1985) revised Learning Style Inventory (LSI) in the assessment of student and teacher learning preferences and the effects those preferences have on student performance and satisfaction. While some of R&S's comments have merit, others are not well founded.

In general, based on some of our more recent psychometric work (Geiger et al. 1992, 1993), we conclude that the standard version of the 1985 LSI needs to be modified to improve its construct validity in order to further its application in accounting education research. However, we do not concur that the LSI and it's foundation in the Experiential Learning Model (ELM) should be so quickly and unequivocally abandoned as contended by R&S. While they have attempted only to deconstruct and largely criticize the instrument, we and other recent researchers (i.e., Geiger et al. 1993; Romero et al. 1992; Veres et al. 1991) have attempted to develop more psychometrically sound instruments that offer potential gains in the enhancement of measuring individual learning preferences.

In responding to R&S's comments, we will first offer some gen-

eral remarks and then address each of the four issues set forth in their paper: 1) measurement error and the effects of a response set, 2) internal consistency reliability, 3) temporal consistency reliability, and 4) instrument validity.

R&S contend that our discussions of the psychometric properties of the LSI were "incomplete and misleading" (R&S, p. 3). Yet, they continue their remarks by acknowledging that this is an area of "emerging research" (R&S, p. 3). This quickly evolving body of research includes both their work as well as our own. In fact, the data for the Geiger (1992) paper was collected in the Spring of 1989 and data for the Geiger and Boyle (1992) paper was collected in the fall of 1989. Analysis was performed shortly after and the papers were submitted for publication approximately one year from data collection. The discussion of the psychometric limitations of the LSI in Geiger (1992) was added after the paper was submitted and accepted for publication in ***The Accounting Educators' Journal***, and the discussions in both studies cite the critical works of R&S in both educational psychology and accounting education journals. More importantly, new research findings, including our own, that shed new light on a subject area do not merit labeling earlier works "incomplete and misleading." In fact, R&S do not cite our most recent work (i.e., Geiger et al. 1993) in their paper, leaving one to believe, based on their own definition, that their discussion of the LSI is also "incomplete and misleading." If that be the case, then it is not unlikely that most, if not all, published research in fast growing, emerging or dynamic areas would be inappropriately characterized as such, i.e., hindsight is everything. We believe, however, that given the "time lag" in publishing research, that these inevitable failures should not be viewed as sins of omission, but rather the result of the publication process.

In footnote number 2, R&S inappropriately lead the reader to believe that students used for the Geiger (1991) study were the same as those in the Geiger (1992) study. In fact, the students used for the Geiger (1991) study were the initial sample for the early longitudinal work reported in Pinto and Geiger (1991) and Geiger and Pinto (1991). In that same footnote, R&S charge that covariate adjusted means should be used to test for differences between groups when the covariate demonstrates a significant effect in an ANCOVA model. Such a test on adjusted means produced similar results as reported in Geiger (1992), Table 4, i.e., assimilators significantly outperformed accommodators ($p < .01$) even after adjusting for differences in GPA. This finding is consistent with the original discussion, as well as prior research which found accommodators signifi-

cantly under performed compared to their peers (Togo and Baldwin, 1990).

R&S also criticize our work in footnote 3 by stating, "Given the small sample of teachers represented, there is no reason to have confidence that these results would be replicated."(R&S, p. 3). We disagree. First, the original paper called for future studies to try and replicate our findings regarding student/teacher cognitive interactions (Geiger and Boyle, 1992, pp. 97-99). Second, the 12 instructors from two universities who participated in the study appear to be an adequate number to report on; especially in lieu of the typically small number of instructors employed in a vast number of educational research articles. Third, as pointed out in our original study (Geiger and Boyle, 1992, p. 99) other researchers have already replicated our general findings with different instruments. For example, Cooper and Miller (1991) used the Myers-Briggs Type Indicator and found results consistent with ours regarding the interaction of students and instructors in terms of both performance and satisfaction.

Measurement Error and Response Set

One of the continuing criticisms by R&S regarding the standard 1985 LSI is the possible existence of a response-set bias due to the one ability per column ipsative scoring format. In fact, while some researchers have argued that they have found a response-set bias (Ruble and Stout, 1990 and 1991; Veres, et al. 1991), they have not examined the LSI using the same individuals. To appropriately isolate any response-set bias, the two instruments utilized in these studies (i.e., scrambled and standard LSI versions) should have been administered to the same individuals. Comparisons between two different groups to assess an individual's tendency to respond in a biased pattern represents a serious flaw in research design. Our most recent work (Geiger et al. 1993) does not suffer from such a design flaw and is the first such study to administer the standard ipsative LSI and a normative 48-item scrambled version to the same individuals (N=455). Our results found average alpha coefficients for the four learning abilities of .833 and .825 for the standard and normative versions, respectively. Hence, the reduction in alpha levels from the standard LSI to the scrambled normative reversion does not reflect the existence of any significant response-set bias. These more recent findings further support the arguments presented in our earlier work that there does not appear to be a significant response-set bias inflating the alpha estimates of internal consistency reliability.

Internal Consistency Reliability

Our findings indicate that the LSI has consistently measured the posited learning abilities in a sufficiently reli-

able manner. One of Nunnally's (1967) early discussions of internal consistency suggested a reliability of .50 or .60 for instruments in early stages of development. He later increased that threshold to .70 (Nunnally, 1978) and also cautions that for some applied settings reliabilities of even .80 are not sufficiently high.

While we agree with R&S that advising students is of serious import, we do not agree that the internal consistency of the LSI shows: 1) lower levels than intended by Nunnally (1978) for reliability measures of this type, and 2) significantly inflated coefficient alphas. First, an applied context for the LSI would be to measure an individual's preferences for multiple ways of learning, and then to offer course and/or career counseling based on their learning style preferences. As important as that is, it is not equivalent to reliably assessing whether an individual possesses tendencies for societal harm (e.g., murder, child molestation, rape, arson, etc.)--personality traits for which Nunnally's comments were also directed. These "more serious" psychological assessments certainly require much higher thresholds of reliability, as Nunnally (1978) affirmed in his discussion of instrument reliability. For assessments of learning preferences in research, an alpha coefficient threshold of .70 or .80 appears entirely adequate. Accordingly, our results, and those of R&S and other researchers, have typically found coefficient alphas well in excess of .70 and typically closer to .80 and above.

Second, based on our properly designed research to assess any artificially inflated alpha levels (i.e., Geiger et al. 1993) discussed earlier), we find no merit in their contention of a response-set bias significantly inflating alpha levels of the standard LSI.

Temporal Consistency Reliability

R&S incorrectly indicate that test-retest reliability and classification stability of the LSI has been performed for periods of up to one year between administrations (R&S, p. 7). In fact, the cited Geiger and Pinto (1991) study examined changes in learning styles of students over time and readministered the LSI at one and two year intervals. While the ELM posits that learning styles of adult learners are "relatively enduring" traits, several studies have found that college age individuals tend to alter their learning preferences more dramatically than other age groups (Price, 1987), and that the same individuals even prefer different learning modes in different types of learning situations (Talbot, 1983). Hence, aggregating a 9-day test/retest reliability calculation with data from a multi-year study of college students is, at best, simplistic and inappropriate.

Instrument Validity

As every introductory student of instrument construction can articulate, adequate instrument reliability is a necessary but not sufficient condition for construct validity--i.e., actually measuring what the instrument purports to measure. Accordingly, after conducting additional psychometric work over the last two years, our general concern with the standard LSI rests not with its reliability, but with issues of construct validity. The ELM posits a learning preference trade-off between the four learning style abilities (i.e., thinking vs. feeling, and doing versus watching) which is measured by the LSI. In particular, the ipsative scoring format of the standard LSI, by design, pits one learning ability against another to form the ELM's proposed learning trade-offs and the resultant two-dimensional learning plane espoused by Kolb (1985).

Our first factor analytic study using the standard ipsative LSI could not adequately support the two bi-polar learning dimensions as posited by ELM (Geiger et al. 1992). Our second study presented the same individuals with the standard ipsative LSI and a scrambled 48-item normatively scored (1=not like me; 7=very much like me) instrument (Geiger et al. 1993). While we replicated our earlier findings with the standard LSI, the modified LSI presented consistent and very strong support for the four individual learning abilities, but provided little support for the two learning dimension trade-offs. These more recent results confirm that the posited separate learning abilities can be validly assessed by individuals, however the actual trade-off of learning abilities imposed by the ipsative scoring format is not as apparent.

Additionally, other recent work by Romero et al. (1992) presented 662 subjects a modified 14-item normatively scored LSI that used behavioral and characteristic anchors representing the four learning abilities. Half of the items were anchored by the abstract conceptualization (thinking)/concrete experience (feeling) trade-off and half by the active experimentation (doing)/reflective observation (watching) tradeoff. Unlike our results, their two-factor results lend strong support for the two-dimensional trade-off posited by Kolb and the ELM.

In sum, recent studies that have used scrambled versions of the LSI (Veres et al. 1991) and modified versions to alleviate the ipsative scoring difficulties (Geiger et al. 1993; Romero et al. 1992) appear to be fruitful avenues of future inquiry. In fact, R&S, without properly discussing this "emerging research" (R&S, p. 3) regarding the psychometric properties of several modified LSI instruments, alluded to this line of research in their footnote number 9. However, they failed to properly discuss these

more recent studies aimed at enhancing the measurement properties of the LSI.

Moreover, although we disagree with several points made in their paper, we agree that the use of the standard LSI should be curtailed. However, we do not believe that the tenants of the ELM should be so quickly abandoned. Recent works have indicated that modified versions of the LSI may prove useful in reliably and validly assessing learning preferences of individuals. Unlike the work of R&S, and discussed in this reply, our and other researchers' most recent efforts have focused on improving the measurement properties of the LSI. Although promising, further study of these and other instruments is necessary before definitive statements regarding their utility can be made.

Notwithstanding future findings, the recent conclusions of Veres et al. (1991), once ardent critics of the standard LSI, warrant repeating: "... the modified version argues against dismissal of the LSI as an instrument for the study of learning styles" (p. 143). Finally, although we disagree with several of the comments presented by R&S, we support their efforts to make accounting education researchers more cognizant of the psychometric issues effecting research instruments employed for analysis.

References

Cooper, S. E., & Miller, J. A. (1991). MBTI learning style-teaching style discongruencies. ***Educational and Psychological Measurement***, 51, 699-706.

Geiger, M. A. (1991). Performance during the first year of college: Differences associated with learning style. ***Psychological Reports***, 58, 633-4.

Geiger, M. A. (1992). Learning styles of introductory accounting students: An extension to course performance and satisfaction. ***The Accounting Educators' Journal***, 4-1, 22-39.

Geiger, M. A., & Boyle, E. J. (1992). Learning styles of students and instructors: An analysis of course performance and satisfaction. ***The Accounting Educators' Journal***, 4-2, 86-101.

Geiger, M. A. Boyle, E. J., & Pinto, J. K. (1992). A factor analysis of Kolb's Revised Learning Style Inventory. ***Educational and Psychological Measurement***, 52, 753-59.

Geiger, M. A. Boyle, E. J., & Pinto, J. K. (1993) An examination of ipsative and normative versions of Kolb's Revised Learning Style Inventory. ***Educational and Psychological Measurement***, 53, 717-26.

Geiger, M. A., & Pinto, J. K. (1991). Changes in learning style preference during a three-year longitudinal study. ***Psychological Reports***, 69, 755-62.

Kolb, D. A. (1985). ***Learning Style Inventory: Self-scoring inventory and interpretation manual.*** Boston: McBer and Company.

Nunnally, J. C. (1967) ***Psychometric Theory.*** New York: McGraw-Hill.

Nunnally, J. C. (1978) ***Psychometric Theory.*** 2nd Ed. New York: McGraw-Hill.

Pinto, J. K., & Geiger, M. A. (1991). Changes in learning-style preferences: A prefatory report of longitudinal findings. ***Psychological Reports***, 68, 195-201.

Price, G. E. (1987). Changes in learning style for a random sample of individuals 18 and older who responded to the productivity environmental preference survey, 1986. Paper presented at the Annual Convention of the American Association of Counseling and Development, April 21-25.

Romero, J. E., Tepper, B. J., & Tetrault, L. A. (1992). Development and validation of new scales to measure Kolb's 1985 learning style dimensions. ***Educational and Psychological Measurement***, 52, 171-80.

Ruble, T. L., & Stout, D. E. (1990). Reliability, construct validity, and response-set bias of the revised Learning-Style Inventory. ***Educational and Psychological Measurement***, 50, 619-29.

Ruble, T. L., & Stout, D.E. (1991). Reliability, classification stability, and response-set bias of alternate forms of the Learning Style Inventory (LSI-1985). ***Educational and Psychological Measurement***, 51, 481-9.

Ruble, J. L., & Stout, D. E. (1993) Comments on the LSI in research on student performance in introductory accounting. ***Accounting Educators' Journal***, 5-2, xxx.

Talbot, R. (1983), Learning styles in different situations, ***Creativity and Innovation Network***, 9, 64-70.

Togo, D. F. and B. A. Baldwin (1990). Learning Style: A determinant of student performance for the introductory financial accounting course. ***Advances in Accounting***, 8, 189-99.

Veres, J. G., III, Sims, R. R., & Locklear, T. S. (1991). Improving the reliability of Kolb's revised Learning Style Inventory. ***Educational and Psychological Measurement***, 51, 143-50.