Mountain Plains Journal of Business and Technology

Date Published: 10-1-2008

Development And Validation Of Two Versions Of The Student Prior Attitude Scale (SPA)

James E. Weber
St. Cloud State University

Gary Yoshimoto
St. Cloud State University

James A. Smith
St. Cloud State University

Howard W. Bohnen
St. Cloud State University

Follow this and additional works at: https://openspaces.unk.edu/mpjbt

Part of the Business Commons

DEVELOPMENT AND VALIDATION OF
TWO VERSIONS OF THE STUDENT
PRIOR ATTITUDE SCALE (SPA)

JAMES E. WEBER, GARY YOSHIMOTO, JAMES A. SMITH AND HOWARD W. BOHNEN
ST. CLOUD STATE UNIVERSITY

ABSTRACT

It is not uncommon for students to have preconceptions of a class before it begins. These preconceptions may come from information about a class, often through word of mouth, or knowledge that the class requires skills inconsistent with the student's skill set. These preconceptions of a class may have a positive or negative effect on how a student initially feels about the class, and may, in fact, have an effect on endof-term student evaluations. Since student evaluations are often used for personnel decisions, it is vital that these evaluations be reliable and valid. This study reports on the development and validation of a scale, the Student Prior Attitude (SPA) scale, to measure the degree to which students are engaged/disengaged, interested/uninterested, or excited/unexcited to be taking a class before the class even begins. Both an 11-item and a 4-item scale are investigated for reliability and validity. Possible uses of the scales are discussed and descriptive data from the sample are reported for both scales. Implications for future research are discussed.

I. INTRODUCTION

Student attitudes toward aspects of the classes they take have been widely studied for many years (Eiszler, 2002; Greimel-Fuhrmann & Geyer, 2003). In fact, Cashin, in his 1988 summary of student ratings research, identified at least 1,300 articles and books on the topic. Student attitudes toward their classes are important. These attitudes have been shown to be related to outcomes important to the student, including grades (Beran & Violato, 2005; Cohen, 1981; Eiszler, 2002; Greenwald, 1995, 2002), attendance (Conard, 2004; Gump, 2005, 2006; Zierold, Garmanm, & Anderson, 2005), cheating (Angell, 2006; Bennett, 2005; Newstead, Franklyn-Stokes, & Armstead, 1996; Roig & Neaman, 1994), and attrition (Daugherty & Lane, 1999) among others.

But these evaluations of student attitudes are used for other purposes, including instructor development, tenure and promotion decisions (Marsh, 1984). Since student attitudes have important outcomes for multiple constituencies, it is vital that the

measures of these attitudes be valid and reliable (Marsh, 1984). Though disagreement exists, student attitudes measured in well-developed student evaluation forms are generally considered to be valid, reliable, strongly related to the instructor who teaches a course, and useful feedback to faculty for developmental purposes and to administrators in making personnel decisions (Greenwald, 2002; Marsh, 1984).

Measures of student attitudes are not without controversy, however. Numerous studies indicate that the grade that a student expects to receive from the class has an effect on student attitudes toward aspects of that class (Beran & Violato, 2005; Cashin, 1988; Eiszler, 2002; Greenwald, 1995, 2002). In fact, Greenwald (1995) has strongly suggested that the effects of the grades that students expect to receive be removed statistically from these measures of student attitudes toward aspects of their classes.

Additionally, measures of student attitudes also show potential bias in several areas. Cashin (1988, p. 3) defines bias in the context of using ratings for personnel decisions or instructor improvement as "variables not a function of the instructor's teaching effectiveness". Examples of these variables from the Cashin summary include faculty rank, student motivation, prior interest in the subject matter, required or elective nature of the class, and academic field. In addition, some classes known to require quantitative skills typically receive lower student evaluations (Cashin & Clegg, 1987).

It is now common for students to investigate classes before they enroll in them. Websites like RateMyProfessors.com provide information on a class and professor that is available to the student before they take a class and can have an effect on student attitudes before the class even begins. Given the widespread belief that attitudes lead toward behaviors (Schultz, 2002) there is every possibility that prior information and attitudes about a class may have an effect measurable in student evaluations at the end of a term, independent of instructor effectiveness. Negative attitudes are of particular concern if they are reflected in negative student evaluations and these evaluations are used for personnel decisions.

There are few studies of the effect of student knowledge, attitudes and expectations prior to the class on end-of-term evaluations. One exception is Gump's (2006) study of the effect of prior attitudes about attendance in general on attendance in specific classes. This study found a significant relationship between prior attitudes and attendance behavior. Given these results regarding the potential for prior attitudes to affect student behaviors in evaluating or rating classes, along with the dearth of research in this area, it seems reasonable to examine research that examines prior expectations on other dependent variables. One profitable area is the study of

alienation or disengagement on student outcomes in education. Studies have shown significant relationships between alienation or disengagement with student outcomes such as cheating (Roig & Neaman, 1994) and attrition (Daugherty & Lane, 1999).

Measuring student alienation or engagement/disengagement from class presents a problem. Although numerous scales exist to measure alienation (Blumenkrantz & Tapp, 1977; Mau, 1992; Pruden, Shuptrine, & Longman, 1974; Ray, 1982), none capture the specific alienation or withdrawal of a student from a class prior to the beginning of the class. For example, Blumenkrantz & Tapp (1977) developed the 22-item Classroom Alienation Scale as a measure of alienation oriented toward the classroom experience, but the items could not be evaluated by a student prior to taking the class.

The purpose of this study is to develop and validate a functional, useful scale of student attitudes prior to class attendance. The Student Prior Attitude (SPA) scale draws broadly from existing literature from a variety of areas and is developed and validated using standard scale development and validation techniques.

II. METHOD

A literature review was conducted to locate existing scales on alienation, engagement/disengagement, and attitudes associated with classes. In general, alienation is thought to be related to the school experience through powerlessness, meaninglessness, normlessness and social estrangement (Mau, 1992), so scales associated with these constructs were sought. Over the course of a semester the authors of this study collected and evaluated scale items for relevance to the central construct of prior positive and negative student attitudes toward a class (Blumenkrantz & Tapp, 1977; Eison, 1981; Lastovicka & Gardner, 1979; Mau, 1992; Pruden et al., 1974; Ray, 1982; Richins, 1997; Saucier, 1994). Given the intent to measure prior attitudes, the great majority of items collected were not relevant to the construct and none of the items could be used without modification. Possible items were extensively modified, and new items were developed. Four professors within the college, familiar with the scale development process, were recruited to help evaluate the items for construct validity. Feedback from these outside evaluators led to the elimination of one item and the modification of a number of others. All items were then associated with anchors scored on a 5-point Likert-type scale ranging from Strongly Disagree = 1 to Strongly Agree = 5. These 32 items along with items intended to assist in validating the proposed scale were combined in a survey that was administered the first day of the semester to 13 sections of business classes at a mid-sized comprehensive Midwestern university. Given the sensitive nature of two of the validation questions, students were assured of anonymity and that their professors would not see the results

of their surveys. The 442 respondents were 57.8% male, had an average age of 21.5 years (SD = 3.2), and an average selfreported GPA of 3.3. Validation proceeded, roughly following well-accepted validation procedures suggested in the business literature (Churchill, 1979).

III. RESULTS

The 32 items were subjected to a Principal Components analysis, which resulted in one interpretable component and 7 total components with Eigenvalues above 1. A scree plot indicated that only the single interpretable component should be retained (the 3 Eigenvalues defining the knee were 12.33, 2.07 and 1.48). Item-total correlations were computed for all items and the results examined. The 11 items with the highest item total correlations, all of which were elements of the interpretable component, were retained for further study. Items were split 6-5 between positively and negatively worded items. Cronbach's Alpha was calculated for the 11 items in order to assess reliability, and it was extremely high (α = .95). All item-total correlations were above r = .69 and the reliability of the proposed scale could not be improved by deleting any of the items. Additional Principal Components analysis of only the 11 items yielded a single 55 component with Eigenvalue greater than 1 (Eigenvalue = 7.37) explaining 67.03% of the variance, providing strong evidence of the unidimensionality of the proposed scale.

Discriminant validity of the 11-item SPA scale was assessed using the correlation between SPA and a six-item Fashion Leadership scale (Goldsmith, Freiden, & Kilsheimer, 1993). This scale was chosen for its brevity, reliability given its length (α = .85 in this sample) and total lack of conceptual relationship with the SPA-11 scale. The correlation between SPA-11 and the Fashion Leadership scale was minimal (r = .08, ns) as expected, providing support for discriminant validity of SPA-11. Based on variance within the results shown, student responses were not a result of identifiable response bias that artificially inflates reliability estimates.

Evidence of convergent validity was provided by an evaluation of the correlations between SPA-11 and two self-reported intention to cheat items included in the survey. In responding to these two items, subjects were assured of their anonymity and that their professors would not see the results of these questions. Since prior research had shown a relationship between alienation and attitudes toward cheating (Roig & Neaman, 1994), it was expected that positive, significant correlations between SPA-11 and cheating intentions would provide evidence that SPA-11 was measuring what was intended. The correlations between SPA-11 and the two cheating items were r = .40 and r = .42, with both significant at the p = 0.01 level. The magnitude of these correlations is relatively high in social science research and

provides support for the convergent validity of SPA11. These results are shown in Table 1.

TABLE 1. EVIDENCE OF RELIABILITY, VALIDITY AND DIMENSIONALITY OF THE SPA-11 AND SPA-4 SCALES

Scale	Reliability	Discriminant Validity	Convergent Validity	
	Coefficient Alpha (α)	Correlation with Fashion Leadership	Correlation with 1st Cheating Item	Correlation with 2nd Cheating Item
SPA-11	.949	.083	.402**	.421**
SPA-4	.922	.031	.378**	.401**

Scale	Dimensionality (Component Analysis)		
	Components	Eigenvalue	% of Variance Explained
SPA-11	1	7.374	67.04%
SPA-4	1	3.242	81.05%

** $p < 0.01$

SPA-11 was scored on a 1-5 point scale, so scores could range from 11 (1 x 11) to 55 (5 x 11), with a midpoint of 33. The actual mean was 30.18 with a minimum of 11 and a 56 maximum of 54. There was significant variation in scores (SD = 9.02), with the mean roughly three points below the midpoint and 32.8% of scores exceeding the midpoint. One-way ANOVA was used to investigate whether SPA-11 varied by gender. The mean for females (30.62) was not significantly different than the mean value for males (29.80). These results are presented in Table 2. Items comprising SPA-11 appear in Appendix 1.

TABLE 2. DESCRIPTIVE STATISTICS FOR THE SPA-11 AND SPA-4 SCALES

Scale	Mean	Std. Dev.	Minimum	Maximum	Gender Differences
SPA-11					
Male (n = 254)	29.80	8.94	11	54	
Female (n = 186)	30.62	9.12	11	54	F = 1.808, p = .165
Total (n = 440)	30.18	9.02	11	54	
SPA-4					
Male (n = 254)	11.82	3.46	4	20	
Female (n = 186)	12.26	3.65	4	20	F = 1.464, p = .232
Total (n = 440)	12.02	3.55	4	20	

In examining the inter-item correlation matrix for SPA-11, researchers noted that four items tracked well together. All four of these items were reverse-scored, and they seemed to represent a general, positive expectation that students might have about the class before they actually took the class. Because short, valid instruments

are potentially very valuable to researchers, the authors decided to examine these four items separately as a proposed SPA-4 scale. Identical procedures were followed as were used in the examination of the SPA-11 scale above.

Coefficient Alpha for SPA-4 was .92, a very high value for a 4-item scale, indicating high internal consistency. Principal components analysis yielded a single component, with an Eigenvalue = 3.24, explaining 81.05% of the variance. SPA-4 appears to be reliable and unidimensional.

Discriminant validity was shown by the minimal (r = .03) correlation between SPA-4 and the Fashion Leadership scale. Convergent validity was demonstrated by correlations of r = .38 and r = .40 respectively of SPA-4 with the two self-reported intention to cheat items (both significant at the p = 0.01 level). In summary, the magnitude of these correlations presents strong evidence of discriminant and convergent validity for the SPA-4 scale.

SPA-4 was scored on a 1-5 point scale, so scores could range from 4 (1 x 4) to 20 (5 x 4), with a midpoint of 12. The observed mean in this sample was 12.02 with a minimum of 4 and a maximum of 20. There was significant variation in scores (SD = 3.54), with the mean virtually equal to the midpoint and 58.9% of scores exceeding the midpoint. One-way ANOVA was used to investigate whether SPA-4 varied by gender. The mean for females (12.26) was not significantly different than the mean value for males (11.82). Items comprising SPA-4 appear in Appendix 1 among the SPA-11 items, marked with an asterisk (*).

IV. DISCUSSION

Student evaluations of their classes and professors have become ubiquitous in higher education. Much like performance appraisals in the business world, these student evaluations have come to be used for instructor development and personnel decision-making purposes. As in the business world, there is a concern for the reliability, validity and potential bias of any instrument used in making personnel decisions. In fact, student evaluations have been criticized for leading to more lenient grading on the part of instructors who rely on student evaluations to aid their promotion and tenure efforts (Eiszler, 2002).

Current research has concentrated on characteristics of the class or instructor that may influence or bias student evaluations. The literature is extensive, with factors such as the instructor's rank, expressiveness, age, gender, personality, and research productivity along with class size, time of the day offered, course level, difficulty, workload and academic field being studied (Cashin, 1988). Student variables have

also been studied, including student motivation to take a course, interest in the class and expected grade in the class (Cashin, 1988). These studies have failed, however, to develop a consistent, reliable measure of student predisposition toward a class. This is what this study has attempted to do in developing the SPA-11 and SPA-4 scales.

The development of the SPA scales is an important step forward in an attempt to examine how student predispositions toward a class may affect outcomes such as student evaluations. Both SPA scales are reliable and are validated against a construct likely related to student evaluations. As would be expected from a longer scale, the SPA11 exhibits slightly better psychometric properties than the SPA-4. What is surprising is how close the reliability and validity figures actually are between the two scales. Although further study is necessary, it is possible that the SPA-4 could be used successfully in place of the longer scale. Short, valid and reliable scales in this area would be a great step forward. Though this research is an important first step, additional, confirmatory validation efforts are needed using a different sample.

If SPA is shown to be significantly related to student evaluations with an effect size large enough to be important, using the SPA scales in a properly designed study should make it possible to evaluate the impact of prior attitudes on student evaluations. This could potentially be of great benefit to instructors that teach unpopular, quantitative classes that have been shown to receive lower student ratings.

The relationship demonstrated by the correlations between the SPA scales and the intention to cheat items in this study point toward needed study of other dysfunctional student behaviors and SPA. Examples might be absenteeism, tardiness, turning in assignments late, or lack of participation in class or in assigned group activities in class. It is easy to see the unfairness of assigning a disgruntled or 58 disengaged student to graded group activities if the instructor knows that this student is unlikely to pull their weight in these group activities.

If prior attitudes toward a class predict student evaluations, other pre-existing traits should also be studied. Especially for quantitative classes, pre-existing quantitative skills predict student performance. Given the relationship between grades and student evaluations, do pre-existing quantitative skills predict student evaluations? If so, a well-designed evaluation system should allow the separation of the effects of prior skill development from the effects of poor current teaching, making the use of student evaluations for personnel decisions in this instance more appropriate.

In summary, this study has provided initial evidence of the reliability, validity, dimensionality and initial descriptive values of two new scales, SPA-11 and SPA-4.

Additional research is needed to confirm the practical usefulness of these scales and to confirm their psychometric properties in a different sample.

REFERENCES

Angell, L. R. (2006). The relationship of impulsiveness, personal efficacy and academic motivation to college cheating. College Student Journal, 40(1), 118-131.

Bennett, R. (2005). Factors associated with student plagiarism in a post-1992 university. Assessment & Evaluation in Higher Education, 30(2), 137-162.

Beran, T., & Violato, C. (2005). Ratings of university teacher instruction: How much do student and course characteristics really matter? Assessment & Evaluation in Higher Education, 30(6), 593-601.

Blumenkrantz, D., & Tapp, J. T. (1977). Alienation and education: A model for empirical study. The Journal of Educational Research, 77(2), 104-109.

Cashin, W. E. (1988). Student ratings of teaching: A summary of the research (pp. 1-6). Kansas State University: Center for Faculty Evaluation & Development.

Cashin, W. E., & Clegg, V. L. (1987). Are student ratings of different academic fields different? Paper presented at the Annual meeting of the American Educational Research Association.

Churchill, G. A. (1979). A paradigm for developing better measures of marketing constructs. Journal of Marketing Research, 16(1), 64-73.

Cohen, P. A. (1981). Student ratings of instruction and student achievement: A metaanalysis of multisection validity studies. Review of Educational Research, 51(3), 281- 309.

 Conard, M. A. (2004). Conscientiousness is key: Incentives for attendance make little difference. Teaching of Psychology, 31(4), 269-272.

Daugherty, T. K., & Lane, E. J. (1999). A longitudinal study of academic and social predictors of college attrition. Social Behavior and Personality, 27(4), 355-362.

Eison, J. A. (1981). A new instrument for assessing students' orientation towards grades and learning. Psychological Reports, 48, 6.

Eiszler, C. F. (2002). College students' evaluations of teaching and grade inflation. Research in Higher Education, 43(4), 483-501.

Goldsmith, R. E., Freiden, J. B., & Kilsheimer, J. C. (1993). Social values and female fashion leadership: A cross-cultural study. Psychology & Marketing, 10, 399-412.

Greenwald, A. G. (1995, August 11-15). Applying social psychology to reveal a major (but correctable) flaw in student evaluations of teaching. Paper presented at the 103rd Annual Meeting of the American Psychological Association, New York, NY.

Greenwald, A. G. (2002). Constructs in student ratings of instructors. In H. I. Braun, D. N. Jackson & D. E. Wiley (Eds.), The role of constructs in psychological and educational measurement (pp. 277-297). Mahwah: Lawrence Erlbaum Associates.

Greimel-Fuhrmann, & Geyer, A. (2003). Students' evaluation of teachers and instructional quality - analysis of relevant factors based on empirical evaluation research. Assessment & Evaluation in Higher Education, 28(3), 229-238.
Gump, S. E. (2005). The cost of cutting class: Attendance as a predictor of student success. College Teaching, 53(1), 21-26.

Gump, S. E. (2006). Guess who's (not) coming to class: Student attitudes as indicators of attendance. Educational Studies, 32(1), 39-46.

Lastovicka, J. L., & Gardner, D. M. (1979). Components of Involvement. In J. C. Maloney & B. Silverman (Eds.), Attitude Research Plays for High Stakes (pp. 53-73). Chicago: American Marketing Association.

Marsh, H. W. (1984). Students' evaluations of university teaching: Dimensionality, reliability, potential biases and utility. Journal of Educational Psychology, 76, 707-754.

Mau, R. Y. (1992). The validity and devolution of a concept: Student alienation. Adolescence, 27(107), 731-741.

Newstead, S. E., Franklyn-Stokes, A., & Armstead, P. (1996). Individual differences in student cheating. Journal of Educational Psychology, 88(2), 229-241.

Pruden, H. O., Shuptrine, F. K., & Longman, D. S. (1974). A measure of alienation from the marketplace. Journal of the Academy of Marketing Science, 2, 610-619.

Ray, J. J. (1982). Toward a definitive alienation scale. The Journal of Psychology, 112, 67-70.

Richins, M. L. (1997). Measuring emotions in the consumption experience. Journal of Consumer Research, 24(2), 127-146.

Roig, M., & Neaman, M. A. W. (1994). Alienation, learning or grade orientation, and achievement as correlates of attitudes toward cheating. Perceptual & Motor Skills, 78, 3.

Saucier, G. (1994). Mini-Markers: A brief version of Goldberg's unipolar Big-Five markers. Journal of Personality Assessment, 63(3), 506-516.

Schultz, D. E. (2002). Mastering brand metrics: The task is to ask the right questions. Marketing Management, 11(3), 8-9.

Zierold, K. M., Garmanm, S., & Anderson, H. A. (2005). A comparison of school performance and behaviors among working and nonworking high school students. Family & Community Health, 28(3), 214-224.

APPENDIX 1. SPA-11 AND SPA-4 ITEMS.

I'm not interested in this class.
I'm only interested in my grade in this class.
I am frustrated that I have to take this class.
I am irritated that I have to take this class.
I am excited to be taking this class. – (R)
I am enthusiastic about taking this class.* – (R)
I am eager to be taking this class.* – (R)
I am pleased to be taking this class.* – (R)
I am only taking this class because it is required.
This class interests me.* – (R)
Having to take this class is a pain.

* SPA-4 items
– (R) reverse scored items
Items were scored on a 5-point scale, with Strongly Disagree = 1 and Strongly Agree = 5.