Development of Student Service-Learning Course Survey (SSLCS) to Measure Service-Learning Course Outcomes

Yan Wang, Feifei Ye, Golden Jackson, Robert Rodgers, Susan Jones
The Ohio State University

Abstract

Service-learning courses focus on both service experience and academic learning. Academic content is covered in both the classroom and the service experience, and the service experiences are reflected upon and processed in the classroom. Based on educational values, potential outcomes can be classified as development of personal competence, interpersonal relationship, and perception of community service as a responsibility of charity or perception of community service as a responsibility of social justice. The Student Service-Learning Course Survey (SSLCS) is designed to measure these four outcome domains. It draws attention to the dichotomy between the two kinds of citizenship and operationalizes the concepts by developing questions to measure the differences. The present study explored the factorial validity of SSLCS and the factorial invariance across gender groups using confirmatory factor analysis. The results of our study indicate that the four factors of SSLCS are validly measured and the partial factorial invariance across gender groups lends support for comparison between female and male groups.

Context and Purpose of Study

Although service-learning as a teaching method gained acceptance in secondary and higher education curricula in the 1980s, the approach is rooted in government programs such as the 1933 Civilian Conservation Corps, the Peace Corps and Vista programs of the 1960s, and the Youth Conservation Corps of the 1970s (O'Grady, 2000). Inspired by educational theorists such as John Dewey, Jean Piaget, David Kolb, and Paolo Freire, each of whom emphasized the importance of integrating experiential learning into the academic curriculum and one of whom (Freire) emphasized achieving social change through reflection on action, contemporary service-learning has an intentional focus on both service experience and academic learning. Academic content is covered in both the classroom and the service experience. The service experiences are reflected upon and processed in the classroom.

The intended outcomes of service-learning courses include academic achievement, personal competence and interpersonal relationship development, and citizenship. Research on the impact of service-learning courses builds on general theories of student development, and theories are used to identify the outcome domains of service-learning courses. For example, Chickering's seven-vector psychosocial theory has been used to delineate such outcome domains as personal competence development, interpersonal relationship development, developing purpose, career planning, etc. (Greene, 1996; Rhodes, 1999). Kohlberg and Gilligan's theories have been used to explore moral development as well as social responsibility development as results of service-learning programs (Delve, Mintz, & Stewart, 1990). King and Kitchener's theories have been used to study and explain cognitive development (Eyler & Giles, 1999).

Based on relevant theories, many studies have supported the effectiveness of service-learning on major outcome domains such as the development of citizenship and social responsibility (Delve, Mintz, & Stewart, 1990; Eyler, Giles, & Braxton, 1997; Giles & Elyer, 1994; Hudson, 1996; Johnson & Bozeman, 1998; Kendrick, 1996; Kollross, 1997; Markus, Howard, & King1993; Myers-Lipton, 1998; Olney & Grande, 1995; Payne, 2000), as well as personal competence and interpersonal relationship development (Cram, 1998; Greene, 1996; Osborne, Hammerich, & Hensley, 1998; Ostrow, 1995; Rhodes, 1999; Wang, 2000). However, different service-

learning programs, depending on intention and design, can have different emphases on the primary outcomes of the programs. One difference centers on citizenship development, with two possible definitions. One definition is to foster a charitable attitude and commitment; the other is to foster a social justice attitude and commitment. For example, Kahne and Westheimer (1999) believe that "citizenship in a democratic community requires more than kindness and decency; it requires engagement in complex social and institutional endeavors," and it "requires that individuals work to create, evaluate, criticize, and change public institutions and programs" (p. 34). To them, there are two kinds of citizenship, one is for change, and the other is for charity. The citizenship for change is to participate in political action and provide solutions to structural problems to achieve social justice. The citizenship for charity emphasizes the importance of altruism and the joy that comes from giving to people less fortunate.

Many theorists emphasize the importance of citizenship for social justice (Barber, 1994; Berman, 1997; Delve, Mintz, & Stewart, 1990; Kahne & Westheimer, 1999; Olney & Grande, 1995; Reardon, 1994). Reardon (1994), for example, believes that if the students do not get adequate opportunity to reflect on social issues creating the need for community service, the student "may end up embracing stereotypic beliefs about the community residents with whom they are working. While such service may enhance the students' feelings of self-worth and moral virtue, it may contribute little to their intellectual and practical understanding of social justice and racial inequality" (p. 53).

Most research on citizenship as a desired outcome of service-learning courses does not specify an underlying meaning of citizenship or develop instruments to reflect the distinction between charitable responsibility and social justice responsibility as two kinds of citizenship. For example, Moely, Mercer, Ilustre, Miron and McFarlan (2002) developed a questionnaire that has six factors, of which one is civic action and another is social justice attitude. However, among the items in civic action, the question "I plan to become involved in my community" did not specify whether it is a charitable involvement or social justice involvement.

In addition, most of the studies on service-learning developed questionnaires to examine only one outcome domain, and only a few studies developed instruments to detect impact of service-learning courses on more than one outcome domain (Eyler & Gile, 1999; Eyler, Giles, & Braxton, 1997; Moely et al., 2002). Among these studies, Moely et al. (2002) used Principle Component Analysis with Varimax rotation to extract several factors representing different outcome domains. They then summed the observed item scores of each domain for each individual and found correlations among different outcome domains.

The current study extends previous work on measuring the outcome of service-learning courses. Based on aforementioned research, we developed one instrument to measure four major outcome domains of service-learning courses: Personal Competence, Interpersonal Relationship, Charitable Responsibility, and Social Justice Responsibility. Through the design of the questions, we tried to distinguish between Charitable Responsibility and Social Justice Responsibility. The purpose of this paper is twofold. The first purpose is to test for the factorial validity of this instrument in a confirmatory approach. The second purpose is to examine the invariance in the factorial structure across gender. Between-group difference is usually of interest to researchers, but test of between-group difference assumes that the survey items mean the same to different groups, and this assumption can be investigated in the name of factorial invariance (Rensvold & Cheung, 1998). Factorial invariance is defined within the more general notion of measurement equivalence. Measurement equivalence exists at several different levels. Factorial invariance is a necessary condition for comparing latent means between groups as well as for other levels of measurement equivalence. Factorial invariance implies that, across groups, item responses are associated with the same constructs and the strength of the association is the same.

Method
Participants

Students enrolled in 22 service-learning courses from Winter 2002 to Winter 2003 (N=487) at The Ohio State University participated in this study. Among the group 30.2% were male, and 69.2% were female, and 0.6% did not report gender. Of the group, 16.8% were younger than 20, 68.0% were between 20 and 25, and 15.2% were older than 25; 82.5% were Caucasian, and 17.5% were African American and other minority groups. In addition, 9.5% were freshmen, 6.8% were sophomores, 13.4% were juniors, 32.2% were seniors, 10.0% were fifth year students, 26.6% were graduate students, and 1.5% did not report their level..

Among the 22 courses, eight courses were offered in the College of Education: Collaboration in Urban Communities, Leadership in Community Service (4 sessions), Media and Technology in Education, Reading Foundations, and Life Span Motor Development; five, in the College of Humanities: Fist-year English Reading and Writing (4) and English Writing Seminar (1); two, in the College of Human Ecology: Consumer Housing Problem and Benefit in Universal Design; two, in the College of Social and Behavioral Science: Interpersonal and Organizational Communication and Health Communication in Interpersonal Contexts; two, in the College of Medicine and Public Health: Medical Communication with Latinos and Program Planning and

Implementation; two, in the College of Engineering: Architecture Design Studio and Ecosystem Management Policy; and one, in the College of Agriculture, Food, and Environmental Science, Senior Seminar in Agricultural Education.

Students in these courses were involved in various kinds of service experiences that were related to the subject matters of their courses. Some students were tutoring elementary or secondary students in their reading, writing, or physical education. Some students volunteered their time in non-profit organizations to write grants, cook food for AIDS/HIV patients, or design brochures. Other students designed a senior residents or childcare center facility and implemented their designs.

Procedure

The Institutional Review Board approved the procedures for use with human research participants, and the participants provided informed consent. Participants in college service-learning classes completed the SSLCS in the first class (pre-course administration) and the last class (post-course administration) of the service-learning course.

Measuring Instrument

Participants completed the SSLCS. The SSLCS questionnaire includes 33 items as previously mentioned. It was designed to measure four outcome domains: Personal Competence, Interpersonal Relationship, Charitable Responsibility, and Social Justice Responsibility. The responses to each item are on a 7-point Likert type scale, where 1 is "Strongly Disagree" and 7 is "Strongly Agree." Several items were negatively worded, and thus reverse-scored to make the answering direction in accordance with the other questions.

The scale for Interpersonal Relationship was taken from Student Development Task and Life Style Inventory (Winston & Miller, 1987), which measures one of Chickering's vectors called Mature Interpersonal Relationship (MIR) (Chickering & Reisser, 1993). Modification was made to change the response scale from "Yes/No" to Likert-type scale to keep this scale consistent with other scales. This scale had nine items. All the nine items were negatively worded and were reverse-scored for later analysis.

In the scale for Personal Competence, items were developed to measure students' self-confidence, leadership skills and communication skills. Studies have shown positive impact of service-learning courses on students' development in these areas (e.g. Moely et al., 2002). This scale had seven items, one of which was negatively worded, and was reverse-scored for later analysis.

The scale for Charitable Responsibility measured students' willingness to help others for altruistic reasons and the scale for Social Justice Responsibility measured students' awareness of social injustice issues and commitment to work for social change. The development of items for Charitable Responsibility and Social Justice Responsibility was based on Eyler and Giles' (1999) notation that citizenship consists of five components: "Values, Knowledge, Skills, Efficacy, Commitment" (p. 157). For each scale, we designed questions to address the five components plus one question on students' belief in their responsibility. Two more questions in the scale of Charitable Responsibility asked students' willingness to volunteer and to provide community services; hence, this scale had eight items and one of them was negatively worded. Three more questions in the scale of Social Justice Responsibility evaluated student awareness of social injustice and need to increase equity; hence, this scale had nine items and two of them were negatively worded. All the negatively worded items in these two scales were reverse-scored for later analysis.

For each of the four scales, Pearson correlation coefficients were examined for each gender at each administration time to see if the items performed as expected with positive correlations with each other in one scale. For the scale of Social Justice Responsibility, two items were consistently found to correlate negatively with the other items, and they are: *I basically feel that this country is fair to all people* and *I feel that each individual controls whether he or she is poor or wealthy*. For the scale of Personal Competence, the item, *I am not clear on my career goals*, had nearly zero correlations with the other items. All of the three items were negatively worded. Research has shown that negative worded items can be confusing and respondents might make careless mistakes when answering them, which could cause an artifact in factor analysis (Marsh, 1996). Thus, these three items were not included and all subsequent analysis was based on the revised instrument with 30 items. A couple of items in the scale of Interpersonal Relationship also showed a few negative correlations. These items were kept because this scale was externally developed (Winston and Miller, 1987), though the response scale was changed from 'Yes/No' to a 7-point scale. The revised instrument is presented in Table 1, along with the abbreviated names of the items, which were used later for presentation purpose.

Data Analysis

Model. Because the development of each scale for SSLCS was based on prior hypothesis of the four-factor model for possible outcomes of service-learning courses, the hypothesis was incorporated explicitly into model specification and estimation process. Confirmatory factor analysis (CFA) was adopted to examine whether each item measures the corresponding factor and whether the relationship among the factors is as expected. Figure 1

Table 1
Items in the instrument of Student Service-Learning Course Survey (SSLCS)

Outcome Domain	Item Content	Name
Personal Competence	1. I know how to lead in a new situation.	PER1
	2. I know how to lead in a cross-cultural situation.	PER2
	3. I know that I can make a positive difference in the lives of others.	PER3
	4. I know how to communicate my ideas in a situation that is new to me.	PER4
	5. I know how to communicate my ideas in a cross-cultural situation.	PER5
	6. I am very aware of some of my own weaknesses and strengths.	PER6
Charitable Responsibility	1. I have a responsibility to help those individuals who are less fortunate than me.	CHA1
	2. I have a responsibility to provide community service.	CHA2
	3. I have a good understanding of the needs in the community where I am going to provide services.	CHA3
	4. I know how to become involved in helping others who are less fortunate than me.	CHA4
	5. I will act in charitable way to help people in need.	CHA5
	6. We should reach out in charitable way to specific people in need.	CHA6
	7. I am confident that I can help individuals in need.	CHA7
	8. I probably won't volunteer after this course. ®	CHA8
Social Justice Responsibility	1. I will act to work for social justice changes in society.	SOC1
	2. We should create programs and public policies to address social issues.	SOC2
	3. I am confident that I can help in promoting equal opportunities for all people.	SOC3
	4. I have a responsibility to help efforts directed at social justice changes in society.	SOC4
	5. I know how to organize efforts for social changes.	SOC5
	6. I have a good understanding of the social justice issues in the community where I am going to provide services.	SOC6
	7. This society needs to increase social and economic equality.	SOC7
Interpersonal Relationship	1. I avoid groups where I would be of the minority race. ®	INT1
	2. I sometimes used phrases or words such as "Blacks have rhythm", or "Honkie," or "people on welfare are only looking for a free ride". ®	INT2
	3. I would prefer not to room with someone who is from a different culture or race. ®	INT3
	4. I find it annoying when I hear people speaking in a language I don't understand. ®	INT4
	5. There are some topics that should never be discussed in college classrooms. ®	INT5
	6. I think most women tend to respond to situations emotionally, while men respond by thinking. ®	INT6
	7. I deal with students who are different from me (for example, of another race or who speak a different language) by being polite and staying away from them as much as possible. ®	INT7
	8. I generally keep my beliefs to myself in order to avoid offending others. ®	INT8
	9. I become annoyed with people who frequently try to change the rules. ®	INT9
Note: ® indicates that this item is reverse-scored for data analysis.		

displays the proposed factorial structure. Each item was allowed to load only on its corresponding factor. The four factors were allowed to covary because their interrelationships are of interest.

Most confirmatory factor analyses are performed as part of a methodological approach known as Structural Equation Modeling (SEM). We used the SEM software called AMOS (Arbuckle & Wothke, 1999) to fit the proposed model in Figure 1 to responses obtained from each gender at each administration time separately. All analyses were conducted on the sample covariance matrices using maximum likelihood estimation. Covariance matrices were used because SEM is for the analysis of covariance structure and its application to sample correlation matrix may produce incorrect values of the omnibus test statistic or standard errors (Cudeck, 1989).

Model fit. Whether the model provides a good fit to the data could be assessed by goodness-of-fit indexes. The most commonly used goodness-of-fit index is the chi-square statistic. A non-significant value of chi-square indicates evidence of good fit. A problem with this statistic is its dependence on sample size as stated by Cheung and Rensvold (2002): "For large sample sizes, it provides a very sensitive statistical test, but not a practical test, of model fit." It is hard to determine how large a sample is when the chi-square statistic is considered not appropriate. The common practice is to assess alternative goodness-of-fit indexes as well, among which CFI (Comparative Fit Index), TLI (Tucker Lewis coefficient), and RMSEA (Root Mean Square Error of Approximation) are commonly reported in the literature. The first two indexes range from 0 to1. We employed CFI and TLI guidelines of 0.95 as standards of good fitting models (Hu & Bentler, 1999). Browne and Cudeck (1993) and MacCallum, Browne, and Sugawara (1996) presented guidelines of assessing model fit with RMSEA: values less than .05 indicate close fit, values ranging from 0.05 to 0.08 indicate fair fit, values from .08 to .10 indicate mediocre fit, and values greater than .10 indicate poor fit. A confidence interval of RMSEA provides information regarding the precision of RMSEA point estimates and was also employed as suggested by MacCallum et al. (1996).

Convergent and discriminant validity. We assessed convergent validity by examining whether the factor loadings of most items related to the same underlying construct are significantly different from zero (Anderson & Gerbing, 1988). Items significantly loaded on the same factor indicate the presence of a common construct and hence evidence of convergent validity of these measures. Regarding discriminant validity for any two estimated constructs, we followed suggestions by Anderson and Gerbing (1988) by constraining the estimated correlation between each pair of constructs to 1.0 and conducting a chi-square difference test between the constrained and unconstrained models. If the constrained model has a significantly higher chi-square value (i.e. significantly worse fit), it is likely that the correlation for the pair of factors is not one and discriminant validity is achieved.

Model diagnostics. Parameter estimates of item factor loadings and communalities were consulted to identify ways that each subscale might be improved from an empirical perspective. To be a good measure of its underlying construct, an item needs to have a significant factor loading and a decent communality (Byrne, 1998). A factor loading is significantly different from zero with its value greater than twice its standard error. The communality is percentage of the item variance attributed to its respective latent factor(s), which serves as an indicator of the extent to which each observed indicator adequately measures its construct (Byrne, 1998). Items

with non-significant factor loading and/or small communalities (relative to other items) were targeted for removal to enhance model parsimony. However, theoretical concerns related to item content were given first priority in model modifications.

Factorial Invariance. Factorial invariance across gender was investigated for pre-course and post-course responses respectively. First, the baseline model (as shown in Figure 1) was fit to the data for both male and female groups simultaneously without imposing between-group constraints. The fit of this simultaneously estimated model could provide the baseline value against which all subsequently specified models are compared. The second step tested factorial invariance, i.e. the invariance of all factor loadings, across gender by imposing equality constraints on corresponding factor loadings for male and female students. The next step depended on whether factorial invariance was established. With factorial invariance, further levels of measurement equivalence would be tested, including the equality of variance-covariance matrices of the latent factors to compare the correlations of the factors across groups, and the equality of error variances to compare scale reliabilities across groups (Cheung & Rensvold, 1999). However, as shown later in the results section, across-gender factorial invariance was not established for either pre-course or post-course administration, and hence, the following procedure was adopted.

With factorial non-invariance, steps were taken to determine the sources of non-invariance. A sequence of models were tested, and in each model, loadings of items related to one factor were constrained equal, allowing the test for the invariance of all item loadings of this factor. When for a certain factor, items were not found to have equal loadings across gender, steps were taken to examine respectively the invariance of each item loading related to this factor. It is important in the whole process that, noted by Byrne (2004), "as factor-loading parameters are found to be invariant across groups, their specified equality constraints are maintained *cumulatively*, through the remainder of the invariance-testing process."

To test the invariance of different levels as mentioned above, we measured the significance of the difference in the chi-square statistic between two models, one without equality constraints on certain parameters across groups and one with them. This is legitimate given that the latter model is a nested subset of the first. Significant difference indicates that the constrained model provides poorer fit than the unconstrained model, and the constraint associated with the former is a source of non-invariance. Because several invariance tests would be conducted, a more stringent alpha level of .01 was set for each single test.

Figure 1
Path Diagram of Hypothesized Factor Model for
Student Service Learning Course Survey (SSLCS)
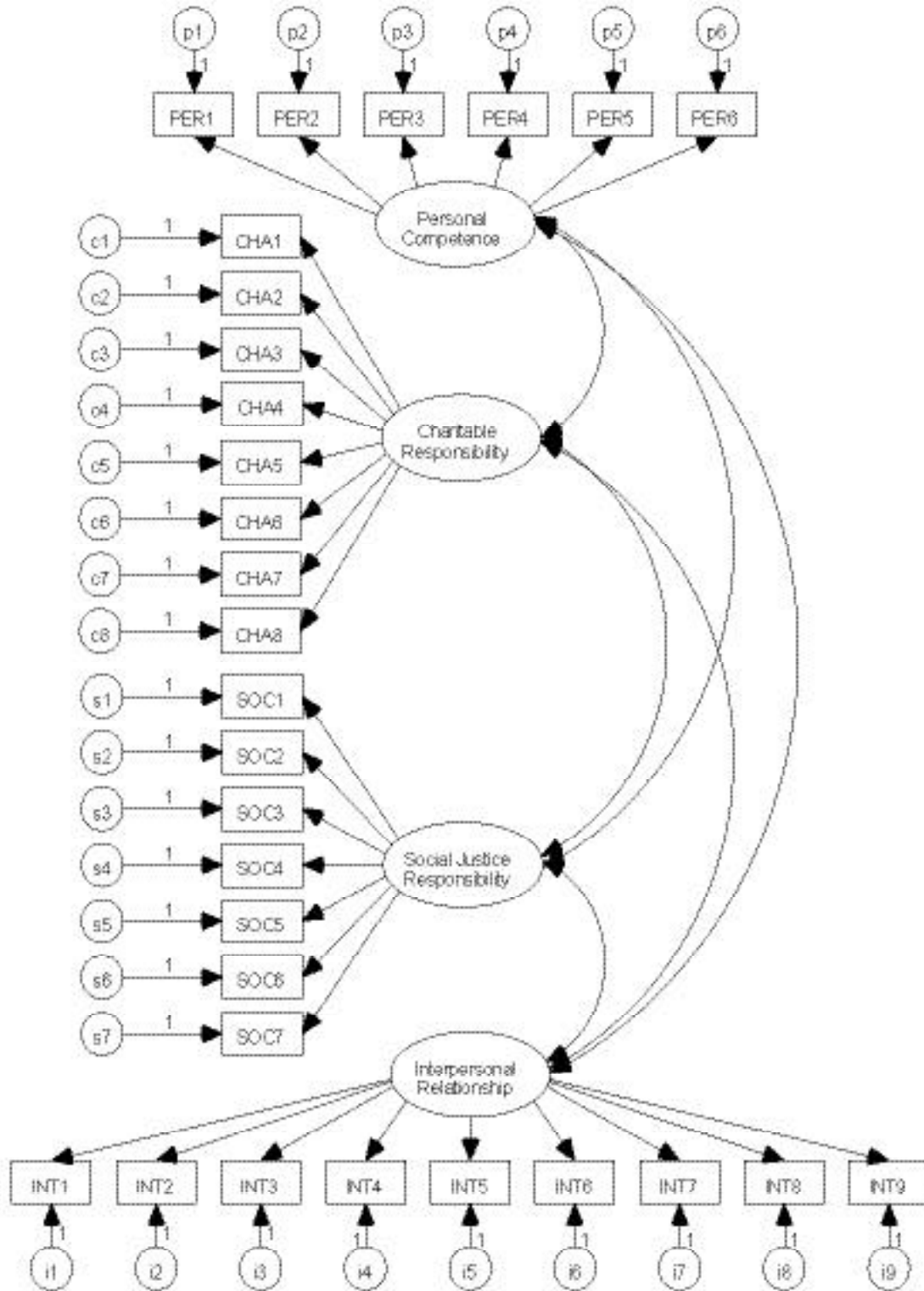
Model identification.   Two types of constraints are commonly used to identify a scale for confirmatory factor models: constrain the variance of a construct to be unity or select an item and fix its factor loading to be unity. In tests for across-group factorial invariance, these identification procedures embody a tacit assumption of across-group invariance: equality of factor variance or equality of the factor loading of the particular item with the unity factor loading (Cheung & Rensvold, 1999). When testing factorial invariance, i.e. the equality of all factor loadings across groups, unity factor variance implies that the test is for a strict factorial invariance, where both factor loadings and factor variances are equal (Meredith, 1993). Cudeck (1989) suggested that unity factor variance could result in a model in which the factors are not scale invariant. Thus, in this study, for the purpose of assessing across-group factorial invariance, we selected one item in each subscale and constrained its factor loading to unity. However, selecting an item arbitrarily could be problematic because it can happen that the selected item itself is not a good indicator of the underlying construct with a non-significant factor loading and/or small communality. To avoid this, we conducted preliminary analysis by constraining the factor variance to unity so that we estimated the factor loadings and communalities of all the items. For each subscale, we selected an item among the items that had significant factor loadings and decent communalities for male and female groups at pre-course and post-course administration. This item was constrained to have unity factor loading in all the subsequent analysis.

Measurement Scale.   There might be arguments as to the appropriateness of analyzing students' responses on a 7-point scale because such data are ordinal, instead of on an interval scale, which is a requirement in confirmatory factor modeling. There are different opinions regarding the variables on an ordinal scale in SEM. It has been common practice in SEM to treat the ordinal variables as if they were on a continuous scale and there is evidence that the more categories there are, the less severe the bias. For reasonably large samples, when the number of ordinal categories is 4 (or 5) or higher, use of maximum likelihood estimation is justified (Byrne, 1998). In this study, the items were measured on a 7-point scale. An examination of the response frequencies of each item showed that there were observations at each scale point. It is not unreasonable to treat the survey responses as on an interval scale.

Incomplete Data.   There are missing values in both pre-course and post-course responses. Listwise deletion would have reduced the sample size dramatically from 487 to 230 for pre-course data and 235 for post-course data. We used the direct maximum likelihood estimation algorithm with missing data in the analysis, as provided by AMOS (Arbuckle & Wothke, 1999). For detailed information with respect to the way AMOS deals with missing data, please refer to Arbuckle (1996).

Results

Internal Consistency

Table 2 presents Cronbach's Coefficient Alpha (Cronbach, 1951) for each subscale by gender and administration time. Coefficient alpha represents a lower bound to the composite reliability of an item set (Raykov, 1997). The observed coefficients were all in an acceptable range for Personal Competence, Charitable Responsibility, and Social Justice Responsibility. For the subscale of Interpersonal Relationship, there was evidence for its reliability in the responses obtained from male students, but not those from female students; that is, male students seemed to be more consistent in answering these 9 items than female students.

Table 2
Cronbach Coefficient Alpha for the four scales of Student Service Learning Course Survey

Scale	Pre-course			Post-course		
	Male	Female	Total	Male	Female	Total
Personal Competence	.835	.802	.813	.774	.758	.763
Charitable Responsibility	.860	.829	.847	.888	.880	.885
Social Justice Responsibility	.804	.777	.790	.825	.829	.828
Mature Interpersonal Relationship	.729	.656	.698	.743	.598	.683

Baseline Model

Model Fit.   The model displayed in Figure 1 was tested against the data obtained from male and female students at the beginning and at the end of the courses separately. Table 3 shows the resulted goodness-of-fit statistics for each group at each administration time. With degrees of freedom equal 399, all the chi-square statistics were significant, indicating the hypothesized factor models might be untenable. However, it is known that chi-square value is sensitive to the sample size and decisions cannot be made based on this single index. The other goodness-of-fit indexes in Table 3 showed that the model provided a good fit with CFI and TLI both greater than .95. Values of RMSEA were in an acceptable range (MacCallum et al., 1996), supporting the fit of the baseline model.

Item Loadings and Communality.   Table 4 provides the factor loadings and communalities of the 30 items for male and female groups at each administration time. Note these are standardized to the first item in each factor and are not interpretable as correlations. All the factor loadings were significantly different from zero for the items in the subscales of Personal Competence,

Table 3
Goodness-of-Fit Indexes of the Baseline Model by Gender and Administration

administration	gender	*df*	X^2	p	CFI	TLI	RMSEA	90% Confidence Interval of RMSEA	
pre-course	male	399	865.535	0.000	0.958	0.951	0.089	0.081	0.098
	female	399	1180.976	0.000	0.971	0.966	0.076	0.071	0.081
post-course	male	399	620.673	0.000	0.975	0.971	0.062	0.052	0.071
	female	399	931.518	0.000	0.979	0.976	0.063	0.058	0.068

No. of Male: 147; No. of Female: 337; CFI = Comparative Fit Index; NFI = Normed Fit Index; TLI = Tucker

Lewis coefficient; RMSEA = root mean square error of approximation

Charitable Responsibility, and Social Justice Responsibility. However, some items had low communalities for both groups at both administration times, such as item 6 in Personal Competence, or for a particular group at pre-course or post-course administration, such as items 5 and 6 in Social Justice Responsibility. These items might not be reliable indicators of their respective factors.

On the subscale of Interpersonal Relationship, item 5 was not significantly loaded for female students at post-course administration, item 7 was not significantly loaded for male students at post-course administration, and item 8 was not significantly loaded for male students at both administration times. An examination of the communalities of these items showed that variances of items 5 and 8 were not explained much by this factor either. For male students, items 5, 7, and 8 had low communalities at both administration times. For female students, items 5, 6, 8, and 9 had low communities at both administration times. This suggests that these two sets of items might not adequately measure the underlying construct for male and female groups respectively.

Table 5
Variances and Intercorrelations of the Four Factors by Gender and Administration

		Male				Female			
	Factor	PC	CR	SJR	IR	PC	CR	SJR	IR
Pre-course	PC	1.039				1.132			
	CR	0.540	1.489			0.628	0.992		
	SJR	0.610	0.903	1.477		0.486	0.790	1.265	
	IR	0.357	0.635	0.719	1.136	0.400	0.429	0.509	0.775
Post-course	PC	0.792				0.757			
	CR	0.666	1.524			0.753	1.300		
	SJR	0.712	0.968	1.174		0.770	0.891	1.266	
	IR	0.463	0.487	0.562	0.862	0.434	0.341	0.468	1.148

Note. PC = Personal Competence; CR = Charitable Responsibility; SJR = Social Justice Responsibility; IR = Interpersonal Relationship. All coefficients are significant.

Variances and Intercorrelations of Factors. Covariance coefficients among the factors are significant and positive, supporting the hypothesis that the relationships between the factors is positive. To show the strength of these relationships, Table 5 presents estimated intercorrelations along with their variances on the diagonal, for male and female groups at pre-course and post-course administrations separately. There were medium to high correlations between the four factors, while the correlation between Charitable Responsibility and Social Justice Responsibility was uniformly the highest.

Convergent and Discriminant Validity. As shown in Table 4, all the items except items 5, 7, and 8 in Interpersonal Relationship, were significantly loaded on their corresponding underlying factors. This showed evidence for convergent validity of these measures (Anderson & Gerbing, 1988).

For discriminant validity, we compared the chi-square statistic between the constrained model where the correlation of a pair of factors was fixed to unity and the unconstrained model with the correlation freely estimated. Altogether, six pairs of factors were examined for each gender at each administration time. With *df* =1 all of the obtained chi-square difference values were substantial and statistically significant at the .001 level with only one exception, where the chi-square difference was 1.76 (ρ =.18) for the male students at post-course administration and the constraint was the unity correlation between Charitable Responsibility and Social Justice Responsibility. Except for this particular value, the chi-square difference values for the six pairs of comparisons ranged from 12.22 to 102.06 for male students and from 79.11 to 231.16 for female students at pre-course administration, from 42.85 to 76.87 for male students and from 44.76 to 180.37 for female students at post-course administration. These chi-square statistics

Table 4
Factor Loadings and Communalities of Items by Gender and Administration

Factor	Item	Factor Loading				Communality			
		pre-course		post-course		pre-course		post-course	
		male	female	male	female	male	female	male	female
Personal Competence	PER1	1.00	1.00	1.00	1.00	0.58	0.65	0.61	0.55
	PER2	1.16	0.92	0.91	1.03	0.59	0.43	0.41	0.48
	PER3	0.75	0.46	0.73	0.54	0.42	0.24	0.37	0.23
	PER4	0.85	0.90	0.85	0.94	0.45	0.61	0.37	0.45
	PER5	0.98	0.86	1.04	1.10	0.58	0.49	0.65	0.64
	PER6	0.46	0.37	0.41	0.44	0.17	0.13	0.06	0.07
Charitable Responsibility	CHA1	1.00	1.00	1.00	1.00	0.53	0.51	0.59	0.67
	CHA2	1.13	1.03	1.01	0.84	0.62	0.48	0.64	0.64
	CHA3	0.64	0.71	0.66	0.65	0.27	0.22	0.39	0.22
	CHA4	0.53	0.89	0.88	0.90	0.24	0.44	0.56	0.58
	CHA5	0.83	0.89	0.69	0.77	0.58	0.58	0.47	0.59
	CHA6	0.75	0.61	0.65	0.77	0.43	0.25	0.44	0.58
	CHA7	0.69	0.84	0.74	0.75	0.37	0.50	0.58	0.51
	CHA8	0.92	0.75	0.58	0.60	0.47	0.29	0.25	0.26
Social Justice Responsibility	SOC1	1.00	1.00	1.00	1.00	0.69	0.70	0.64	0.68
	SOC2	0.85	0.80	0.61	0.73	0.53	0.45	0.29	0.45
	SOC3	0.84	0.59	0.88	0.61	0.52	0.30	0.55	0.38
	SOC4	0.96	0.86	1.07	0.94	0.55	0.57	0.64	0.62
	SOC5	0.21	0.55	0.76	0.69	0.03	0.20	0.38	0.31
	SOC6	0.61	0.38	0.40	0.72	0.24	0.08	0.13	0.26
	SOC7	0.76	0.70	0.67	0.53	0.34	0.36	0.23	0.23
Interpersonal Relationship	INT1	1.00	1.00	1.00	1.00	0.38	0.37	0.35	0.51
	INT2	0.86	0.78	1.40	0.51	0.22	0.20	0.38	0.14
	INT3	0.98	0.97	0.82	0.85	0.47	0.29	0.27	0.36
	INT4	0.95	1.36	1.35	0.96	0.30	0.45	0.41	0.34
	INT5	0.62	0.41	0.63	*0.10*	0.12	0.04	0.08	0.00
	INT6	1.08	0.36	1.27	0.35	0.40	0.03	0.37	0.04
	INT7	0.50	0.96	*0.39*	0.65	0.12	0.34	0.04	0.23
	INT8	*0.26*	0.46	*0.37*	0.38	0.03	0.07	0.04	0.06
	INT9	0.78	0.60	0.90	0.47	0.21	0.11	0.25	0.09

Note. All the factor loadings are significant except the loadings in italic and underlined.

suggested the presence of discriminant validity of the measures on the four SSLCS factor scales, with the exception that male students considered the constructs of Charitable Responsibility and Social Justice Responsibility approximately the same after being educated in service learning courses.

Factorial Invariance Across Gender

For the pre-course administration, goodness-of-fit statistics related to the two-group unconstrained model (Model 1), where the factor loadings were freely estimated, are reported in Table 6. The chi-square value of 2047.46, with df = 798, provides the baseline value against which subsequent tests for invariance were compared. Model 1 had CFI and RMSEA values of .967 and .057, indicating that the hypothesized four-factor model of SSLCS represented a good fit for both male and female groups.

Model 2 constrained all the factor loadings to be equal across gender groups. The chi-square difference between Model 1 and Model 2 was 67.55, which was significant (ρ =.00) at df =26. The null hypothesis that all the factor loadings are equal across gender was rejected. We proceeded to find the sources of invariance by testing the equivalence of factor loading of each construct across gender groups.

Models 3, 4, and 5 tested the equality of factor loadings across gender groups for the three subscales of Personal Competence, Charitable Responsibility, and Social Justice Responsibility respectively. With an alpha level of .01, the differences in the chi-square statistic were non-significant, furnishing evidence for the equal factor loadings of items related to these three factors across gender groups.

The factor loadings of items related to Interpersonal Relationship were non-invariant between female and male groups. From Model 6a to 6h, the factor loading of each item in the subscale of Interpersonal Relationship was constrained to be equal across gender in an order from item 2 to item 9. It was found that item 6 has non-invariant factor loading across gender with a significant chi-square difference value of 13.10 at df = 1.

For the post-course administration, the two-group unconstrained model (Model 1) has a chi-square value of 1552.77 with df = 798. CFI and RMSEA values of .979 and .044 indicated that the hypothesized four-factor model of SSLCS represented a good fit across two gender groups.

For post-course administration, an examination of factorial invariance showed similar results to those of pre-course administration. The only difference was that the sources of non-invariance were items 2 and 6 in the subscale of Interpersonal Relationship, while only item 6 was non-invariant for pre-course administration.

Conclusions and Discussion

The questionnaire of SSLCS was developed to measure four major outcome domains of service-learning courses — personal competence, interpersonal relationship, charitable responsibility, and social justice responsibility. The present study explored the factorial validity of SSLCS. Confirmatory factor analysis was performed on data collected from students enrolled in service-learning courses at Ohio State University. To prepare for examining between-gender difference, analysis was conducted on female and male groups separately at pre- and post-course administrations. Results showed that our proposed four-factor model provided satisfactory fit to the data. The significance of factor loadings of most items supported the convergent validity of the factor structure.

As evidence of discriminant validity of these survey measures, the correlation was found not to be 1 for all but one pair of factors. The two factors found to be highly correlated were Charitable Responsibility and Social Justice Responsibility for male students at post-course administration. It should be noted that one purpose of the newly developed instrument of SSLCS was to specify the distinction between charitable responsibility and social justice responsibility as two kinds of citizenship. Such effort resulted in partial success with high but not perfect correlations between these two factors for female students at both pre- and post-course administration and for male students at pre-course administration. However, for male students measured after the service-learning course, the two constructs were found to be almost the same. This empirical evidence suggests that it is hard to distinguish between these two constructs, though theoretically the two constructs are considered different. From the correlation coefficients between these two factors, we saw that male students had higher correlations than female students, and so did post-course administration than pre-course administration. It seems that male students behave more similarly to the two constructs than female students. The service-learning courses appear to intensify the relationship between these two constructs. Such results were only based on this study, and we suggest more study on the relationship between these two constructs. If they are found to be non-separable, the items related to the two constructs might be combined to measure the construct of citizenship.

The covariance structure among factors in this study partly confirmed the relationship among similar constructs found by Moely et al. (2002). In Moely et al.'s study, the two subscales of civic action and social justice attitudes are conceptually similar to Charitable Responsibility and Social Justice Responsibility in this study respectively. In Moely et al.'s study, the civic action and social justice attitudes have almost the highest correlation among all the factors, which is similar to the results of this study.

Table 6
Goodness-of-Fit Statistics for Tests of Factorial Invariance: A Summary

Model	Model Description	Comparative Model	x^2	df	Δx^2	Δdf	ρ
	pre-course						
Model1	Combined baseline models (Male and Female)		2047.46	798			
Model2	Factor loadings constrained equal	Model 1	2115.01	824	67.55	26	0.00
Model3	Factor loadings on PC constrained equal	Model 1	2056.59	803	9.14	5	0.10
Model4	Model3 with factor loadings on CR constrained equal	Model 3	2072.76	810	16.16	7	0.02
Model5	Model4 with factor loadings on SJR constrained equal	Model 4	2088.19	816	15.44	6	0.02
Model6a	Model5 with factor loadings of Item 2 on IR constrained equal	Model5	2088.32	817	0.12	1	0.73
Model6b	Model5 with factor loadings of Items 2 and 3 on IR constrained equal	Model6a	2088.32	818	0.01	1	0.93
Model6c	Model5 with factor loadings of Items 2, 3, and 4 on IR constrained equal	Model6b	2092.18	819	3.86	1	0.05
Model6d	Model5 with factor loadings of Items 2, 3, 4, and 5 on IR constrained equal	Model6c	2093.66	820	1.48	1	0.22
Model6e	Model5 with factor loadings of Items 2, 3, 4, 5, and 6 on IR constrained equal	Model6d	2106.76	821	13.10	1	0.00
Model6f	Model5 with factor loadings of Items 2, 3, 4, 5, and 7 on IR constrained equal	Model6d	2098.73	821	5.08	1	0.02
Model6g	Model5 with factor loadings of Items 2, 3, 4, 5, 7, and 8 on IR constrained equal	Model6f	2099.28	822	0.54	1	0.46
Model6h	Model5 with factor loadings of Items 2, 3, 4, 5, 7, 8, and 9 on IR constrained equal	Model6g	2101.54	823	2.26	1	0.13
	post-course						
Model1	Combined baseline models (Male and Female)		1552.77	798			
Model2	Factor loadings constrained equal	Model 1	1606.14	824	53.37	26	0.00
Model3	Factor loadings on PC constrained equal	Model 1	1556.18	803	3.41	5	0.64
Model4	Model3 with factor loadings on CR constrained equal	Model 3	1562.09	810	5.91	7	0.55
Model5	Model4 with factor loadings on SJR constrained equal	Model 4	1577.81	816	15.72	6	0.02
Model6a	Model5 with factor loadings of Item 2 on IR constrained equal	Model5	1588.10	817	10.29	1	0.00
Model6b	Model5 with factor loadings of Item 3 on IR constrained equal	Model5	1577.83	817	0.03	1	0.87
Model6c	Model5 with factor loadings of Items 3 and 4 on IR constrained equal	Model6b	1579.85	818	2.01	1	0.16
Model6d	Model5 with factor loadings of Items 3, 4, and 5 on IR constrained equal	Model6c	1583.06	819	3.21	1	0.07
Model6e	Model5 with factor loadings of Items 3, 4, 5, and 6 on IR constrained equal	Model6d	1593.87	820	10.81	1	0.00
Model6f	Model5 with factor loadings of Items 3, 4, 5, and 7 on IR constrained equal	Model6d	1585.34	820	2.28	1	0.13
Model6g	Model5 with factor loadings of Items 3, 4, 5, 7, and 8 on IR constrained equal	Model6f	1585.38	821	0.05	1	0.83
Model6h	Model5 with factor loadings of Items 3, 4, 5, 7, 8, and 9 on IR constrained equal	Model6g	1587.66	822	2.28	1	0.13

Note. Δx^2=difference in chi-square values between models; Δdf= difference in number of degrees of freedom between models; PC = Personal Competence; CR = Charitable Responsibility; SJR = Social Justice Responsibility IR = Interpersonal Relationship.

However, this study is different from Moely et al.'s study in that in their study, the two factors focus on Action and Attitude respectively while in this study both factors have indicators on Knowledge, Attitudes, Belief, Self-confidence, and Responsibility.

An examination of the intercorrelations among these four factors showed that the correlations increased from the pre-course administration to the post-course administration among the three factors of Personal Competence, Charitable Responsibility, and Social Justice Responsibility. This was not the case for the relationship between Interpersonal Relationship and the constructs of Charitable Responsibility and Social Justice Responsibility. To explore why such difference exists, we generated composite scores of the observations of each subscale and conducted t-test to compare the difference between pre-course and post-course administrations for male and female students respectively (results not included). It was found that the service-learning courses had positive influence on students' personal competence, charitable responsibility, and social justice responsibility, while the only scale that did not change significantly was interpersonal relationship. While changes in means do not result in changes in correlations, the apparent lack of change in Interpersonal Relationship might suggest why the correlations between Interpersonal Relationship and the factors of Responsibilities did not increase while those among the other three factors did. Previous studies (Greene, 1996; Rhodes, 1999) also did not find any significant change in interpersonal relationship using the same instrument but with a different scale (Yes/No). This suggests that although this scale has been proved to be useful in other context, it might not be appropriate in testing effectiveness of service-learning on students' interpersonal relationship development. This non-effectiveness might be because of the quality of its indicators in measuring this factor. In our study, it had the lowest coefficient of internal consistency. The communalities for the indicators of this factor had lower values than expected, especially items 5, 7, 8 for male and 5, 6, 8, 9 for female, implying that these items did not have substantial relationship with the underlying construct. This scale might be multidimensional instead of measuring one construct. Further research using this scale could remove these items or add another dimension composed of these items in the instrument.

The investigation of factorial invariance lent support to the equality of factor loadings across gender groups for all the constructs except Interpersonal Relationship. In this subscale, the association between item 6, *I think most women tend to respond to situations emotionally, while men respond by thinking,* and the underlying construct was different across gender groups at both pre-course and post-course administrations: the relationship was much stronger for male than female.

The relationship between item 2, *I sometimes used phrases or words such as "Blacks have rhythm", or "honkie," or "people on welfare are only looking for a free ride,"* and the underlying construct was also stronger for male than for female, but only at post-course administration. These results suggest that female and male groups have different perception of interpersonal relationship, presenting a problem in comparing these two groups on this construct using latent means. Given the difficulty to achieve factorial invariance, Byrne, Shavelson, and Muthén (1989) proposed partial factorial invariance, where non-invariant items are allowed to vary when analyzing between-group differences. Byrne et al. argued that if the non-invariant items constitute only a small portion of the model, they should not affect cross-group comparisons to any significant extent. In our case, only 2 out of 9 items had non-equivalent factor loadings across gender groups. It is reasonable to compare the gender difference with these two items' loading freely estimated for both groups.

Regarding limitations and future directions, our study was a pretest-posttest design without a control group in which courses that cover similar subject matter but do not have service-learning components are selected. With a control group, we could have assessed whether similar results from factor analysis would be obtained by administering the same instrument. The control group also could be used to see whether the positive influence of service-learning on students comes from the course per se or is the combination of test effect and course impact. Further studies need to investigate this area.

In addition, service-learning courses have impacts on students beyond the outcome domains measured by SSLCS. For example, students' academic learning and cognitive development have shown to be other areas where service-learning courses can benefit students. Further development of SSLCS should consider these factors as well.

This study extends the current efforts of creating a valid instrument to measure service-learning course outcomes in a comprehensive way. It draws attention to the distinction between the two kinds of citizenship and tries to operationalize the concepts by developing questions to measure the differences. Modest evidence was found for the distinction between these two constructs.

Another approach to investigating the convergent and discriminant validity of measures on the SSLCS would be to examine data within a multitrait-multimethod (MTMM) matrix framework as originally proposed by Campbell and Fiske (1959), where measures of multiple traits are assessed by multiple methods and all trait-method correlations are arranged in a MTMM matrix. A CFA approach to the analysis of data in a MTMM matrix should be explored in future studies. The convergent validity of measures on the SSLCS factor scales could be supported if the correlations of measures on the same

construct when assessed by different methods are high. Discriminant validity would be substantiated if the intercorrelations of measures on the factor scales reflecting different constructs are relatively low.

The results of our study indicate that the four factors of SSLCS are validly measured and the partial factorial invariance across gender groups lends support for comparison between female and male groups. The question of immediate interest is whether there is effect of service-learning courses, i.e., whether students' responses are different between pre-course and post-course administrations. We suggest that such pretest-posttest differences not be examined until the factorial structure is proved invariant from pre-course administration to post-course administration.

Editor's Notes

This article by Wang, Ye, Jackson, Rodgers, and Jones has much to provide to those in several types of endeavors. In terms of who would measure the differential aspects of service-based learning, it provides the first draft of a survey developed around the elements of concepts in the field. For those who look at the psychometric issues in developing any such instrument, they deal with some of the key issues in such an instrument that can have profound effects on the resulting instrument. Finally, for those interested in using various methodologies with simultaneous equations, they demonstrate using AMOS, which is a Simultaneous Equation Model (SEM) to perform confirmatory factor analysis and to test various hypotheses associated with comparative and hieratical modeling.

In terms of measuring the impact of classroom service-based learning, they develop an instrument based on the current thinking that the aspects of service-based learning include Personal Competence, Charitable Responsibility, Social Justice Responsibility, and Interpersonal Relationship. Their major innovation here seems to be splitting Responsibility into Social Justice and Charitable perceptions. While their statistical evidence of the uniqueness of these two factors is not going to be persuasive to all based on the correlation of the two aspects when viewed as underlying factors, it is important to remember that the results obtained in this study are evidence of the statistical relationship based on this specific set of responses. There is at least an indication that the two aspects are worthy of further study and further consideration of their divergent validity for other studies.

The general methodology of empirically looking for an appropriate and statistically parsimonious confirmation of their factor model is a very good demonstration of the types of questions that should be viewed by those who measure outcomes in our various educational experiences. Regardless of your opinion of the need for

factorial invariance, they build an argument for its importance and also demonstrate a possible way to look at the invariance of factors over the pre-and post-test.

The discussion also points out some of the realities of having items scaled in two different directions and points to some of the issues discussed in considering social desirability and acquiescence response sets as well as the disturbing likelihood that students (or whoever – for that matter) respond to a survey they adopt a "respond to the left side" or "respond on the right side" of the set of responses for the questions. The discussion of Cronbach's Alpha is also a good reminder that scales without internal consistency are not likely to support conclusive results from studies. There is a need to have a caution of modifying the set of items and reporting the final results as being statistically significant at the stated probability level.

This gets into the use of the SEM program in AMOS. The key in looking at the statistical test is to realize that the strategy is one of starting from a full or "saturated model and then moving to remove aspects of the model. Each aspect that is removed increases the "Poorness of Fit" and the increase in the poorness can be viewed relative to the increase in the error degrees of freedom. The degrees of freedom come from the model, not from the size of the sample. A complex model can have a large number of degrees of freedom. The discussion of the null model may or may not be persuasive. For those who use this methodology, there are also Modification Indices in AMOS that would suggest which of the earlier assumptions of the initial factor model might be relaxed as represented by the inclusion of additional paths to the initial model. One might suspect that these indices would suggest a loading of some of the Responsibility items on both Charitable and Social Responsibility.

The issue of the initial model not-withstanding, the reader is given evidence in terms of some of the other indices for goodness or poorness of fit. Note these are only two of several alternatives.

The key aspect of the testing is that a systematic strategy is crucial and this is based on the use of "next best" and nested or hieratical tests. While these fit the requirements for the change in the unexplained variance and the change in the degrees of freedom, the large number of alternative requires that, as the authors note, the probabilities be viewed as guides to next steps rather than correct probabilities. This is similar to the issues encountered when one does sequential regression modeling of adding or deleting variables.

Finally this article raises a multitude of key questions for both the methodologist and the individual involved in the concepts of service-based learning. Is the evidence sufficiently compelling to accept the four-factor model without some of the items loading on multiple factors? Is the evidence sufficiently compelling to delete the scale of

Interpersonal Relations from future studies? If the scale is not deleted, what should be done with the problematic items on the scale such as 5, 6, 7, 8, and 9? Does the evidence that the mean for Interpersonal Relations does not change from pre- to post-test give any real insight into the stability of the correlations between factors? If you were measuring the result of a service based learning course – what measures would you use? Are there other measures that would broaden the view of the outcomes?

References

Anderson, J. C., & Gerbing, G. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103,* 411-423.

Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. Marcoulides, & R. Schumacker (Eds.), *Advanced Structural Equation Modeling: Issues and Technique.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Arbuckle, J. L., & Worthke, W. (1999). *AMOS 4.0 User's Guide.* Chicago, IL: Smallwaters Corp.

Barber, B. R. (1994). A proposal for mandatory citizen education and community service. *Michigan Journal of Community Service Learning, 1 (1),* 86-93.

Berman, S. (1997). *Children's social consciousness and the development of social responsibility.* New York: State University of New York Press.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing Structural Equation Models* (pp. 136-162). Newbury Park, CA: Sage.

Byrne, B. M. (1998). *Structural Equation Modeling with LISREL, PRELIS, and SIMPLIS: Basic Concepts, Applications, and Programming.* Mahwah, NJ: Lawrence Erlbaum.

Byrne, B. M. (2004). Testing for multigroup invariance using AMOS Graphics: A road less traveled. *Structural Equation Modeling, 11(2),* 272-300.

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105,* 456-466.

Campbell, D. T., and Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81-105.

Cheung, G. W., & Rensvold, R. B. (1999). Testing factorial invariance across groups: a reconceptualization and proposed new method. *Journal of Management, 25(1),* 1-27.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indices for testing measurement invariance. *Structural Equation Modeling, 9,* 233–255.

Chickering, A. W., & Reisser, L. (1993). *Education and identity.* San Francisco, CA : Jossey-Bass Publishers.

Cram, S. B. (1998). *The impact of service learning on moral development and self-esteem of community college ethics students.* Unpublished Dissertation, University of Iowa.

Cronbach L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16,* 197-334.

Cudeck, R. (1989). Analysis of correlation matrices using covariance structure models. *Psychological Bulletin, 105(2),* 317-327.

Delve, C. I., Mintz, S. D., & Stewart, G. M. (1990). Promoting values development through community service: A design. In C. I. Delve, S. D. Mintz, & G. M. Stewart (Eds.), *Community service as value education. New directions for Students Services, No. 50,* (pp.7-29). San Francisco: Jessey-Bass.

Eyler, J., & Giles, D. E., Jr. (1999). *Where is the learning in service-learning?* San Francisco, CA: Jossey-Bass, Inc.

Eyler,J. S., Giles, D. E., Jr., & Braxton, J. (1997). The impact of service-learning on college students. *Michigan Journal of Community Service Learning, 4,* 5-15.

Giles, D. E., & Eyler, J. S. (1994). The impact of a College Community Service Laboratory on students' personal, social and cognitive outcomes. *Journal of Adolescence, 17,* 327-339.

Greene, D. P. (1996). *Moral reasoning, student development, reciprocity and quality of life in a service learning experiment.* Unpublished dissertation, Colorado State University.

Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55.

Hudson, W. (1996). Combining community service and the study of American public policy. *Michigan Journal of Community Service Learning, 3,* 82-91.

Johnson, S. D., & Bozeman, M. (1998). *Service learning and the development of social responsibility.* Paper presented at the Annual Convention of the Central States Communication Association, Chicago, IL.

Kahne, J., & Westheimer, J. (1999). In the service of what? The politics of service learning. In Claus, J. & Ogden, D. (Eds.), *Service-learning for youth empowerment and social change,* (pp. 25-42). New York: Perter Lang.

Kendrick, J. R. (1996). Outcomes of service-learning in an Introduction in Sociology Course. *Michigan Journal of Community Service Learning, 3,* 72-81.

Kollross, C. A. (1997). *Service learning and citizenship: Is there a connection?* Unpublised Thesis, California State University, Long Beach.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1,* 130-149.

Markus, G. B, Howard, J. P. F., & King, D. C. (1993). Integrating community service and classroom instruction enhances learning: Results from an experiment. *Educational Evaluation and Policy Analysis, 15 (4),* 410-419.

Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifacts? *Journal of Personality and Social Psychology, 70(4),* 810-819.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58,* 525-543.

Moely, B, D., Mercer, S. H., Ilustre, V., Miron, D., & McFarlan, M. (2002). Psychometric properties and correlates of the Civic Attitudes and Skills Questionnaire (CASQ): A measure of students' attitudes related to service-learning. *Michigan Journal of Community Service Learning, 8,* 15-26.

Myers-lipton, S. J. (1998). Effect of a comprehensive service-learning program on college students' civic responsibility. *Teaching Sociology, 26,* 243-258.

O'Grady, C. R. (2000). Integrating service leaning and multicultural education: An overview. In O'Grady, C. R. (Ed.), *Integrating service-learning and multicultural education in colleges and universities,* (pp. 1-21). Mahwah, NJ: Lawrence Erlbaum Associates.

Olney, C., & Grande, S. (1995). Validation of a Scale to Measure Development of Social Responsibility. *Michigan Journal of Community Service Learning, 2,* 43-53.

Osborne, R. E., Hammerich, S., & Hensley, C. (1998). Student effects of service-learning: Tracking change across a semester. *Michigan Journal of Community Service Learning, 5,* 5-13.

Ostrow, J. M. (1995). Self-consciousness and social position: On college students changing their minds about the homeless. *Qualitative Sociology, 18 (3),* 357-375.

Payne, C. A. (2000). Changes in involvement as measured by the Community Service Involvement Preference Inventory. *Michigan Journal of Community Service Learning, 7,* 41-46.

Raykov, T. (1997). Scale reliability, Cronbach's coeffeicent alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate Behavioral Research, 32,* 329–354.

Reardon, K. M. (1994). Undergraduate research in distressed urban communities: An undervalued from of service-learning. *Michigan Journal of Community Service Learning, 1,* 44-54.

Rensvold, R. B., & Cheung, G.W. (1998). Testing measurement models for factorial invariance: A systematic approach. *Educational and Psychological Measurement, 58,* 1017–1034.

Rhodes, C. (1999). *Psychological changes in student development of college sophomore women along Chickering's seven vectors with service-learning as an institutional effect.* Unpublished Dissertation, Peabody College, Vanderbilt University, Nashville, TN.

Wang, W. (2000). *Service learning: Is it good for you?* Paper presented at the Annual Meeting of the AERA Conference Roundtable, New Orleans, LA.

Winston, R. B., & Miller T. K. (1987), Student developmental task and lifestyle inventory manual. *Athens, GA: Student Development Associates, Inc.*

IR Applications is an AIR refereed publication that publishes articles focused on the application of advanced and specialized methodologies. The articles address applying qualitative and quantitative techniques to the processes used to support higher education management.

Authors can submit contributions from various sources such as a Forum presentation or an individual article. The articles should be 10-15 double-spaced pages, and include an abstract and references. Reviewers will rate the quality of an article as well as indicate the appropriateness for the alternatives. For articles accepted for *IR Applications*, the author and reviewers may be asked for comments and considerations on the application of the methodologies the articles discuss.

Articles accepted for *IR Applications* will be published on the AIR Web site and will be available for download by AIR members as a PDF document. Because of the characteristics of Web-publishing, articles will be published upon availability providing members timely access to the material.

Please send manuscripts and/or inquiries regarding *IR Applications* to Dr. Gerald McLaughlin.