

A Psychometric Analysis of the Student Leadership Competencies Inventory

David M. Rosch, Ph.D.

University of Illinois at Urbana-Champaign

Corey Seemiller, Ph.D.

Wright State University

Abstract

The Student Leadership Competencies Inventory consists of eight scales, each corresponding to its relevant leadership construct within the Student Leadership Competencies framework (Seemiller, 2013). Due to the increasing use of the framework and associated inventory in leadership development programs in higher education, we conducted a thorough analysis of the psychometric properties within each scale. Specifically, using a national dataset of university student responses, we analyzed internal reliability statistics, and conducted exploratory factor analysis with varimax rotation and maximum likelihood confirmatory factor analysis for each of the eight scales. Our results suggested that all scales, overall, possess sufficient strength to be considered valid measures of the leadership constructs within the Inventory, with some notably high co-variances between certain sub-scale factors in several scales.

Introduction

The use of leadership competencies is not a new phenomenon, dating back to the late 1960s and early 1970s with McClelland's assertion of the need to look at one's ability rather than intelligence as a qualifier for school admission or employment (McClelland, 1973). McClelland's work has been built upon over the years, resulting in our modern-day conception of competencies which focuses on knowledge, values, abilities, and behaviors that contribute to one's effective performance in a task or role (Seemiller, 2013).

In the late 1990s to early 2000s, a focus on the competencies required for the effective practice of leadership gained traction in the field of leadership education, challenging the notion that competencies were only reflective of the technical requirements of one's particular job (Bolden & Gosling, 2006). It is now common to find leadership competencies embedded within professional associations, businesses, and schools as a way to guide individual training, development, and evaluation (Seemiller & Murray, 2013). However, using leadership competencies within student leadership programs and courses in higher education remains a relatively new phenomenon. For instance, the initial list of what later became known as the Student Leadership Competencies (SLC) framework was developed in 2008 specifically for college student leadership educators (Seemiller, 2013). In 2009, the National Association of Campus Activities (NACA) created *The Competency Guide for College Student Leaders* (Brill et al., 2009), offering their own framework of leadership competencies. Both frameworks were designed to aid in designing and assessing intentional college student leadership development

experiences. Using such a framework allows leadership educators to identify specific intended competencies to focus on, design learning experiences around those competencies, and then measure student development in the context of those competencies (Seemiller, 2013). In addition to aiding leadership educators, a competency framework supports students' efforts as they engage in their own development. For example, being aware of one's own proficiencies and critical areas for enhancement allows for students to seek out and participate in opportunities that aid in the development of competencies.

Since the emergence of leadership competencies, many inventories, assessments, and self-evaluations have emerged in an attempt to effectively measure one's competency proficiency. Despite the vast number of these instruments available generally, few are designed specifically for college students; and of those, very few have actually been statistically analyzed. Within this group, the Emotionally Intelligent Leadership (EIL) Inventory offers students an opportunity to self-report their level of engagement with 19 capacities (sub-scales). The inventory includes 183 items across the 19 capacities with a general goal of three items per capacity (Miguel & Allen, 2016). Another instrument used with college students is the Student Leadership Practices Inventory (SLPI), developed to assess students' self-perception of their behavior across five practices of effective leadership. Each practice (sub-scale) includes 6 items for a total of 30 items in the instrument (Kouzes & Posner, 2013). In 2011, three years after the emergence of the Student Leadership Competencies framework, the Student Leadership Competencies Inventory, a self-evaluation and reflection tool, was developed for students to assist in their identification of competency proficiency levels across the 60 Student Leadership Competencies.

While both the EIL Inventory and the SLPI include competencies associated with their particular frameworks, the measurements are different from the Student Leadership Competencies Inventory. The items on the EIL Inventory measure frequency of engagement in a competency related to a particular capacity (never to always on a seven-point scale), and the items on the SLPI measure frequency of engagement in a competency related to a particular practice (rarely to very frequently on a five-point scale). Both of these instruments reflect a performance aspect of leadership in which each item measures how often a student engages in a particular leadership behavior. The Student Leadership Competencies Inventory, on the other hand, incorporates a 7-point agree/disagree scale for each item measuring to what extent students believe they have the ability to effectively utilize particular competencies, regardless if they actually engage in related behaviors. Because not all students might have had an opportunity to engage in the use of every competency, thus making it difficult to self-assess performance, the Inventory measures to what extent students believe they have the ability to effectively engage in each competency when called upon to do so. Measuring one's ability, rather than performance, also aligns with the purpose of the Student Leadership Competencies in serving as a developmental tool to help educators design learning experiences that help prepare students for engaging in leadership.

Validating Psychometric Scales. Given that the Inventory and Competency framework have been in use for a number of years, we sought to validate their conceptual structure using commonly accepted psychometric tools. While space does not exist for a comprehensive explanation of available statistical tools or commonly accepted practices at various levels of

rigor, we can summarize such information. The first and most common tool for reporting psychometric validity of a quantitative survey measure in social sciences is the Cronbach alpha statistic (Santos, 1999), which is a measure of the inter-reliability of individual items within a scale. The higher the alpha result, the stronger the correlation among participants' responses to items within the scale. Given that the alpha is a correlation statistic, results can range from 0 (no correlation among items) to 1 (perfect correlation); within this process, negatively-scored items are converted to keep correlations positive. In validating survey scales, a common standard for what is a borderline acceptable alpha result has long been 0.7, with results labeled as "strong" if higher than 0.8 (Nunnally, 1978). In past scale research in the leadership education discipline, the range of alpha results for the initial study of the Emotionally Intelligent Leadership for Students scale ranged from .81 to .93 (Miguel & Allen, 2016), while the ranges for the initial study of the Student Leadership Practices Inventory ranged from .73 to .89 for the self-reported version (Posner, 2004).

While internal reliability serves as an important indicator of validity, other tools should also be applied when initially investigating the use of a survey scale, including conducting exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) (Pearson, 2009). The overarching goal of EFA, for a single-scale survey, is to reduce the number of necessary items within the scale, and if multiple sub-scales exist situated within a larger omnibus scale, to additionally ensure that each item within each scale "loads" (i.e. is associated) only to that single sub-scale (Kremelberg, 2010). In an EFA, the researcher does not pre-determine within an analysis the number of latent factors that the survey scale is designed to assess. Rather, the researcher conducts a post-hoc analysis comparing (i.e. "exploring") the number of factors that emerged in the statistical analysis to the theoretical arrangement of latent factors the survey was designed to assess (Kremelberg, 2010). To statistically "extract" factors within an EFA, researchers typically use "principle components analysis," which attempts to determine which individual survey items are most strongly correlated with which groups of survey items (i.e. factors) that are themselves uncorrelated with other groups (Jolliffe, 2002). To accomplish this task, researchers must decide if and how to "rotate" survey items when examining the resulting factors, which may make the groupings of survey items easier to interpret. The most common method within an EFA conducted for this purpose is a "varimax" rotation, which often makes small loading results for individual survey items smaller and large results larger (Kremelberg, 2010). In this type of analysis, it is commonly accepted practice that survey sub-scales should map directly to the latent factors that they are designed to assess, and that each survey item should load at a level of .40 or higher only onto its respective scale (Kremelberg, 2010). Using this method, individual survey items that do not load appropriately or highly enough to their respective scales can be deleted from the scale, and scales themselves can be checked for the strength with which they map to latent theoretical factors.

A common final step, once factors and survey items are shown to be appropriately included in the framework, is to conduct confirmatory factor analysis. Within a CFA, researchers "specify" prior to the analysis the number of factors that are being measured and which survey items map to which factors. Once specified, researchers examine the resulting "goodness of fit" statistics to make judgements about the appropriateness of the scale as a whole to represent the underlying factors in which the scale is designed to assess (Spicer, 2005). As with EFA, while perfect consensus does not exist regarding how good the fit of certain statistics

should be to consider a scale “validated,” it is typically accepted for samples that are relatively large (i.e. more than 500 participants) that the Comparative Fit Index (CFI) should be .90 or higher and the Root Mean Square Error of Approximation (RMSEA) should be .07 or lower (Bentler, 1990).

Purpose

The purpose of this study was to assess the psychometric strength of the Student Leadership Competencies Inventory.

Research Questions. Given the rising significance of the Student Leadership Competencies framework and its corresponding inventory for assessing leadership effectiveness, our interest was to thoroughly assess the psychometric strength of the inventory instrument using a national population of students that represent typical users of the tool. Therefore, we possessed the following research questions:

1. Does each of the eight scales within the Student Leadership Competencies Inventory possess sufficient internal reliability for broader validation?
2. Does each scale retain a factor structure within an exploratory factor analysis that corresponds to the theoretical structure of the Competencies framework, and do individual items load strongly enough within their corresponding sub-scale to justify inclusion?
3. Within each scale, does a confirmatory factor analysis yield “goodness of fit” statistics strong enough to suggest that the scale is a valid measure of the leadership construct it is designed to assess within the Student Leadership Competencies framework?

Methods

Instrument. Version 3 of the Student Leadership Competencies Inventory (Seemiller, 2011), which serves as the focus of our analysis, emerged from two prior iterations. Although the focus of this article is not on the development of the instrument itself, understanding the evolution of the Inventory can be informative. Table 1 includes a description of each version, what changes were employed for the subsequent version, and how those changes were determined.

Table 1. *Changing structure of SLC Inventory.*

	Version 1 (2011)	Version 2 (2015)	Version 3 (2016)
Total Competencies	61	60	60
Total Survey Items Across Sub-scales	132	166	181
Response Scale	5 point Likert	5 point Likert	7 point Likert
Modality	Paper	Online-freestanding	Online-Qualtrics
Total Participants in Initial Analysis	122	1597	1398
Analyses Conducted	Cronbach's Alpha	Cronbach's Alpha; EFA	Cronbach's Alpha; EFA; CFA
Post-hoc Changes Made	Deleted 9 items, reworded 8 items, and added 43 items	Reworded 62 items, added 15 items, changed to 7-point scale	

Version 3 of the Inventory includes eight freestanding assessment scales, one for each competency cluster from the SLC framework. Each assessment includes only the sub-scales (competencies) associated with the competency cluster.

Table 2. *SLC Inventory.*

Competency Cluster (Scale)	Competencies Included (Sub-Scales)	Number of Items
Learning and Reasoning	Research	3
	Other Perspectives	3
	Reflection and Application	3
	Systems Thinking	3
	Analysis	3
	Synthesis	3
	Evaluation	3
	Idea Generation	3
	Problem Solving	3
	Decision Making	3
Self-Awareness and Development	Self-Understanding	3
	Personal Values	3
	Personal Contributions	3
	Scope of Competence	3
	Receiving Feedback	3
	Self-Development	3
Group Dynamics	Organizational Behavior	3
	Power Dynamics	3
	Group Development	3
	Creating Change	3
Civic Responsibility	Diversity	3

	Others' Circumstances	3
	Inclusion	3
	Social Justice	3
	Social Responsibility	3
	Service	3
Interpersonal Interaction	Productive Relationships	3
	Appropriate Interaction	3
	Helping Others	3
	Empathy	3
	Mentoring	3
	Motivation	3
	Others' Contributions	3
	Empowerment	3
	Providing Feedback	3
	Supervision	3
	Collaboration	3
Communication	Verbal Communication	3
	Nonverbal Communication	3
	Listening	4
	Writing	3
	Facilitation	3
	Conflict Negotiation	3
	Advocating for a Point of View	3
Strategic Planning	Mission	3
	Vision	3
	Goals	3
	Plan	3
	Organization	3
Personal Behavior	Initiative	3
	Functioning Independently	3
	Follow-Through	3
	Responsibility for Personal Behavior	3
	Ethics	3
	Responding to Ambiguity	3
	Responding to Change	3
	Resiliency	3
	Positive Attitude	3
	Confidence	3
	Excellence	3

Each item in the Inventory included a 7-point response scale: Strongly Disagree, Moderately Disagree, Slightly Disagree, Neither Agree Nor Disagree, Slightly Agree, Moderately Agree, and Strongly Agree. A sample item in the Others' Contributions sub-scale, for example, was: "I have the skills needed to ensure that everyone in a group is operating from their strengths." Version 3 of the Inventory was moved onto the Qualtrics platform because of its ease in use, improved appearance to students, and ability to both provide a report to the user after completing the Inventory and allow researchers to collect and analyze data from within the Qualtrics data collection system. Students provided consent to participate in our study through

the main link to any of the eight assessment scales that included both a downloadable consent form as well as a statement that says, “By completing any or all of the Student Leadership Competencies Inventory assessments, you are agreeing to the terms laid out in the Informed Consent.”

Data Collection. In fall of 2015, the Student Leadership Competencies Inventory version 2 was released, which included both an overhaul of sub-scales from version 1 as well as a transition to an online platform as outlined in Table 1. During fall of 2015, an email was sent to a listserv of approximately 125 educators who had previously expressed interest in staying up-to-date on information related to the Student Leadership Competencies. A link was also posted on the Lead365 conference website and shared with educators who had previously sent students to Lead365. And, a pre-conference session discussing the Student Leadership Competencies was offered at the International Leadership Association conference.

As evidenced in Table 2, version 2 had a high number of participants. This surge is significant as it relates to version 3, which is the focus in this article. By the time version 3 had been released in February 2016 (with data collection from February 14, 2016 through August 9, 2017), some educators had already implemented the Inventory into their curriculum. Because the URL for the Inventory remained the same, despite moving platforms for version 3, some educators simply continued having their students complete the Inventory after it switched to version 3.

More recruiting for participants, however, did take place specifically for version 3 in fall of 2016. Information about the Inventory was shared directly with professionals through the National Clearinghouse for Leadership Programs and the National Association of Student Personnel Administrators (NASPA) Student Leadership Programs Knowledge Community (SLPKC) listservs, SLPKC social media accounts, and the American College Personnel Association’s Commission for Student Involvement Newsletter. In addition, information about the Inventory was shared at professional presentations on the Student Leadership Competencies including a NASPA pre-conference program in spring 2016, NASPA webinar in summer 2016, and a Lead365 conference session in fall 2016. And, during fall 2016, information about the Inventory was posted on the creator of the Inventory’s LinkedIn account with a potential reach of more than 500.

During fall 2016, links to the Inventory were also sent directly to students, including those following the Wright State University Leadership Studies Twitter account as Wright State University was the hosting institution as well as collegiate members of Omicron Delta Kappa Leadership Honorary. The link on the Lead365 website remained on the home page, ushering students to the new version.

Some additional recruiting took place in 2017 as well. During the spring of 2017, the author of the Student Leadership Competencies participated in a joint webinar on Student Affairs Live discussing leadership competencies, and during the summer of that year, listserv messages about the Student Leadership Competencies in general (including the Inventory) were sent to the Association of College and University Housing Officers - International, the Great Lakes Association of College and University Housing Officers, and Ohio Housing Officers.

Given that the Inventory served as a self-evaluation tool open to anyone interested in completing it, no intact population can be identified. Educators could forward the hyperlink to the Inventory to students to complete on their own or formally incorporate it in a class or program. In addition, students who saw the link on the Lead365 website, the Student Leadership Competencies website, on the Omicron Delta Kappa listserv, or Wright State University Twitter account could take the Inventory on their own. Further, because the link was open with no password necessary, it could be forwarded to other individuals to also complete.

Sample. Version 3 of the Inventory was intentionally divided into 8 individual assessment scales, one for each of the eight clusters of the Student Leadership Competencies Framework. Separating the larger Inventory into smaller components allowed for students to select which of the individual mini-assessments they wanted to take. In some cases, they may have taken all, whereas in others they may have only taken one that is relevant to their context. Because of this, some students opted to take some assessments and not others, making participation rates across assessment scales varied. Thus, participants who consented to provide data for this study did so without needing to complete all eight scales within the Inventory. However, as many potential participants were all associated in some way with formal leadership development programs in postsecondary education, it is likely that overlap exists across participant pools within each scale analysis. As our research goal was to examine the psychometric properties of each scale, only responses that contained a full response set within the scale were examined. Therefore, any participants who had completed less than 100% of the items within the respective scale were eliminated from our analysis. Table 3 displays the sample size and demographic spread for each resulting sample.

Table 3. Scale sample sizes and demographics.

Scale	n	Man (%)	Woman (%)	White (%)	African-American (%)	Asian-American (%)	Latinx (%)	Native American (%)
Civic Responsibility	96 6	35	64	70	10	7	6	1
Communication	10 63	35	63	69	10	7	7	1
Group Dynamics	10 96	37	62	70	9	7	7	1
Interpersonal Interaction	11 42	37	62	70	10	7	7	1
Learning and Reasoning	13 98	33	65	66	11	9	6	2
Personal Behavior	12 26	38	60	69	11	7	8	1
Self-Awareness	13 18	36	62	69	10	7	7	1
Strategic Planning	99 4	35	64	70	11	6	6	2

NOTE – not all percentages add up to 100 due to unreported demographic data

Data Analysis. A comprehensive analysis of the psychometric properties of a survey instrument requires a multi-step investigation. As it is common in survey analysis to report internal reliabilities, often represented by Cronbach alpha statistics in leadership development research (Santos, 1999), we first examined the Cronbach alpha statistics for all scales and each sub-scale found within each respective overall scale. We then investigated the underlying factor structure found in each scale conducting principle components analysis (PCA) exploratory factor analysis (EFA) with varimax rotation, designed to maximize the degree of contrast between items loading to separate sub-scales within each overall scale. Within this set of analyses, we examined the underlying factor structure found within each scale and noted loading statistics for each individual item. Lastly, to investigate the degree to which each scale adequately presented the underlying constructs it was designed to measure, we conducted maximum likelihood confirmative factor analysis (CFA) for each scale. Within this set of analyses, we examined multiple widely-used goodness of fit statistics for each scale to determine each scale's validity in representing its underlying construct.

Methodological Limitations. Any research study is conducted with inherent limitations related to chosen methodology; our research described herein was no different. Perhaps the most significant limitation within this study was the lack of control exerted over student participants. Some percentage of participants were likely formally enrolled in a leadership course, while some were not. Some percentage of participants possessed more formal knowledge of the Competency framework, while some likely possessed very little. Some participants completed multiple assessment scales, while some completed only one. While in other studies a lack of control would disqualify results, we feel that this overall sample is representative of the types of students

who would typically complete the Inventory. At the same time, however, we recognize that our sample was heavily skewed towards those who identify as a woman with regards to gender, and to slightly less of an extent, as White with regard to racial identity. We felt these sample-based limitations would not disqualify the general validity of our findings, yet recognize the need for expanded analysis of more diverse samples in the future.

Another potential limitation in our analysis was our decision to treat each assessment scale (and therefore each competency cluster) separately. This reduces our ability to make claims about the statistical relationships between and across competency clusters. At the same time, given the topical breadth across clusters and the fact that many leadership educators who use the Inventory currently only integrate particular scales within their formal programs, we assert that investigating the full Inventory in one omnibus analysis was unnecessary. With these limitations in mind, we proceeded with the study.

Results

Prior to engaging with our primary research questions, we first examined the overall shape of each dataset by determining scale means and dispersion statistics. The summary of our findings are displayed in Table 4. Most participants' responses to survey items fell between "slightly agree" and "moderately agree" within each scale

Table 4. *Competency scale means and dispersion statistics.*

Scale	M	SD
Civic Responsibility	5.85	1.23
Communication	5.75	1.21
Group Dynamics	5.76	1.10
Interpersonal Interaction	5.88	1.15
Learning and Reasoning	5.88	1.17
Personal Behavior	5.87	1.20
Self- Awareness	5.89	1.17
Strategic Planning	5.80	1.24

The remainder of our results will be nested within each scale, where we describe each of our analyses with respect to each individual scale.

Civic Responsibility. We first investigated the Cronbach alpha reliability statistics for the overall scale and each respective subscale. These results are displayed in Table 5. The overall scale and each respective sub-scale analysis suggested adequate inter-item reliability appropriate for widespread use (Santos, 1999).

Table 5. *Civic Responsibility scale and sub-scale reliability statistics.*

Scale/sub-scale	Reliability
Civic Responsibility (overall)	.93
Diversity	.84
Others' Circumstances	.85
Inclusion	.84
Social Justice	.83
Social Responsibility	.78
Service	.90

We then conducted a Principle Components Analysis Exploratory Factor Analysis (EFA), which is appropriate for item analysis and potential item reduction in survey analysis (Conway & Huffcutt, 2003). The EFA resulted in four factors emerging with eigenvalues over 1.0, which together accounted for 68% of the variance in scale results. Individual item loadings can be found in Table 6. For clarity, loadings lower than 0.40 are not displayed. Each survey item loaded on only one factor (with the exception of the third "Social Responsibility" item, which loaded relatively weakly on two factors), while all items within a subscale loaded collectively onto a common factor. However, significant loading overlap emerged across the Diversity, Others' Circumstances, and Inclusion sub-scales.

Table 6. *Civic Responsibility EFA individual item loadings.*

Scale Item	Factor 1	Factor 2	Factor 3	Factor 4
Diversity1	.64			
Div2	.57			
Div3	.57			
Others' Circumstances1	.68			
OC2	.63			
OC3	.44			
Inclusion1	.75			
Inc2	.80			
Inc3	.82			
SocialJustice1			.58	
SJ2			.92	
SJ3			.88	
SocialResponsibility1				.56
SR2				.66
SR3		.42		.40
Service1		.96		
Ser2		.94		
Ser3		.85		

Lastly, we conducted a maximum likelihood confirmatory factor analysis, which suggested that the empirical data was a good representation of the underlying factor structure described within the Competency framework (CFI=0.95; RMSEA=0.07).

Communication. The alpha reliability statistics that emerged from our analysis are displayed in Table 7. Most reliabilities fell well within acceptable ranges, while the Non-verbal Communication statistic should be interpreted as marginally acceptable.

Table 7. *Communication scale and sub-scale reliability statistics.*

<u>Scale/sub-scale</u>	<u>Reliability</u>
Civic Responsibility (overall)	.91
Verbal Communication	.81
Non-verbal Communication	.67
Listening	.85
Writing	.90
Facilitation	.70
Conflict Negotiation	.76
Advocating a Point of View	.79

The results of the EFA analysis on the Communication scale can be found in Table 8. Five factors emerged from the analysis possessing eigenvalues higher than 1.0, which collectively predicted 64% of the total scale variance. All items other than the third Non-verbal Communication item loaded strongly onto a single factor, while all items within a subscale loaded collectively onto a common factor. Significant loading overlap emerged between Verbal and Non-verbal Communication items, as well as between Facilitation and Conflict Negotiation items.

Table 8. *Communication EFA individual item loadings.*

Scale Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
VerbalCommunication1	.72				
VC2	.78				
VC3	.66				
Non-verbalCommunication1	.59				
NC2	.60				
NC3					
Listening1				-.80	
Lis2				-.90	
Lis3				-.82	
Lis4				-.63	
Writing1		-.95			
Wri2		-.94			
Wri3		-.90			
Facilitation1			.61		
Fac2			.54		
Fac3			.77		
ConflictNegotiation1			.71		
CN2			.63		
CN3			.57		
AdvocatingPointofView1					-.52
APV2					-.86
APV3					-.79

We conducted a maximum likelihood confirmatory factor analysis for the Communication scale, which suggested that the empirical data was a good representation of the underlying factor structure described within the Competency framework (CFI=0.94; RMSEA=0.06).

Group Dynamics. All alpha reliability statistics for the Group Dynamics scale are displayed in Table 9. All reliabilities fell well within acceptable ranges.

Table 9. *Group Dynamics scale and sub-scale reliability statistics.*

Scale/sub-scale	Reliability
Group Dynamics (overall)	.90
Organizational Behavior	.84
Power Dynamics	.82
Group Development	.87
Creating Change	.77

The results of the EFA analysis on the Group Dynamics scale can be found in Table 10. A total of three factors emerged with eigenvalues greater than 1.0, predicting 70% of the total variance found within the analysis. All items loaded strongly onto a single factor, while all items within a

subscale loaded collectively onto a common factor. Significant loading overlap emerged only between Organizational Behavior and Power Dynamics items.

Table 10. *Group Dynamics EFA individual item loadings.*

Scale Item	Factor 1	Factor 2	Factor 3
OrganizationalBehavior1	.83		
OB2	.85		
OB3	.85		
Power Dynamics 1	.81		
PD2	.74		
PD3	.65		
Group Development1		.89	
GD2		.91	
GD3		.87	
Creating Change1			.92
CC2			.90
CC3			.52

We conducted a maximum likelihood confirmatory factor analysis for the Group Dynamics scale, which suggested that the empirical data was a good representation of the underlying factor structure described within the Competency framework (CFI=0.97; RMSEA=0.06).

Interpersonal Interaction. All alpha reliability statistics for the Interpersonal Interaction scale are displayed in Table 11. Most reliabilities fell well within acceptable ranges, while the Productive Relationships sub-scale reliability emerged as marginally reliable, and the Collaboration sub-scale did not achieve adequate internal reliability.

Table 11. *Interpersonal scale and sub-scale reliability statistics.*

Scale/sub-scale	Reliability
Interpersonal Interaction (overall)	.94
Productive Relationships	.63
Appropriate Interaction	.76
Helping Others	.81
Empathy	.83
Mentoring	.85
Motivating Others	.83
Others' Contributions	.87
Empowerment	.79
Providing Feedback	.74
Supervision	.83
Collaboration	.58

The results of the EFA analysis on the Interpersonal Interaction scale can be found in Table 12. Six factors emerged from the analysis possessing eigenvalues higher than 1.0, which collectively predicted 61% of the total scale variance. All items other than the second Productive Relationships item and the third Providing Feedback item loaded onto a single factor. All items within most subscales loaded collectively onto a common factor, with the exception of the Collaboration sub-scale. Significant loading overlap occurred between Productive Relationships, Helping Others, and Empathy items; Mentoring and Providing Feedback items; and Motivating Others and Others' Contributions items.

Table 12. *Communication EFA individual item loadings.*

Scale Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
ProductiveRelationships1				.49		
PR2						
PR3				.59		
AppropriateInteractions1				.69		
AI2				.70		
AI3				.65		
HelpingOthers1	.79					
HO2	.67					
HO3	.77					
Empathy1	.62					
Emp2	.49					
Emp3	.50					
Mentoring1						.73
Men2						.70
Men3						.55
MotivatingOthers1		-.77				
MO2		-.76				
MO3		-.67				
OthersContributions1		-.68				
OC2		-.77				
OC3		-.70				
Empowerment1			.42			
Empo2			.87			
Empo3			.81			
ProvidingFeedback1						.79
PF2						.71
PF3						
Supervision1					-.47	
Sup2					-.58	
Sup3					-.54	
Collaboration1			.51			
Col2					-.42	
Col3			.84			

We also conducted a maximum likelihood confirmatory factor analysis for the Interpersonal Interaction scale, which suggested that despite the loading overlap seen within the EFA results, the empirical data was a good representation of the underlying factor structure (CFI=0.91; RMSEA=0.06).

Learning and Reasoning Scale. All alpha reliability statistics for the Learning and Reasoning scale are displayed in Table 13. Most reliabilities fell well within acceptable ranges, while the Decision-making sub-scale did not achieve adequate internal reliability.

Table 13. *Interpersonal scale and sub-scale reliability statistics.*

Scale/sub-scale	Reliability
Learning and Reasoning (overall)	.94
Research	.76
Others Perspectives	.72
Reflection and Application	.84
Systems Thinking	.75
Analysis	.72
Synthesis	.83
Evaluation	.70
Idea Generation	.87
Problem-solving	.82
Decision-making	.57

The results of the Learning and Reasoning EFA analysis can be found in Table 14. Six factors emerged from the analysis possessing eigenvalues higher than 1.0, which collectively predicted 61% of the total scale variance. All items other than the first Systems Thinking survey item loaded onto a single factor, while all items within each subscale loaded collectively onto a common factor. Significant loading overlap occurred between Systems Thinking, Analysis, and Synthesis items; Evaluation and Decision-making items; and Idea Generation and Problem-solving items.

Table 14. *Learning and Reasoning EFA individual item loadings.*

Scale Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Research1					.73	
Res2					.68	
Res3					.65	
OthersPerspectives1				.59		
OP2				.56		
OP3				.69		
ReflectionApplication1						-.86
RA2						-.86
RA3						-.80
SystemsThinking1						
ST2	.52					
ST3	.43					
Analysis1	.61					
Ana2	.40					
Ana3	.71					
Synthesis1	.69					
Syn2	.72					
Syn3	.78					
Evaluation1			.61			
Eva2			.74			
Eva3			.64			
IdeaGeneration1		-.88				
IG2		-.85				
IG3		-.89				
ProblemSolving1		-.44				
PS2		-.49				
PS3		-.46				
DecisionMaking1			.53			
DM2			.53			
DM3			.49			

We also conducted a maximum likelihood confirmatory factor analysis for the Learning and Reasoning scale, which suggested that the empirical data was a good representation of the underlying factor structure (CFI=0.92; RMSEA=0.06).

Self-Awareness. The alpha reliability statistics that emerged from our analysis of the Self-Awareness scale are displayed in Table 15. All sub-scale reliabilities fell well within acceptable ranges.

Table 15. *Self-Awareness scale and sub-scale reliability statistics.*

Scale/sub-scale	Reliability
Self-Awareness (overall)	.89
Self-Understanding	.77
Personal Values	.82
Personal Contributions	.77
Scope of Competence	.75
Receiving Feedback	.77
Self-Development	.79

The EFA results for the Self-Awareness scale can be found in Table 16. Five factors emerged from the analysis possessing eigenvalues higher than 1.0, which collectively predicted 65% of the total scale variance. All items loaded strongly onto a single factor, while all items within a subscale loaded collectively onto a common factor. Significant loading overlap emerged between only Self-Understanding and Personal Values items.

Table 16. *Self-Awareness EFA individual item loadings.*

Scale Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Self-Understanding1	.63				
SU2	.52				
SU3	.53				
PersonalValues1	.83				
PV2	.80				
PV3	.72				
PersonalContributions1				.77	
PC2				.84	
PC3				.85	
ScopeofCompetence1		.64			
SC2		.82			
SC3		.84			
ReceivingFeedback1					-.74
RF2					-.81
RF3					-.71
Self-Development1			-.62		
SD2			-.68		
SD3			-.66		

We conducted a maximum likelihood confirmatory factor analysis for the Self-Awareness scale, which suggested that the empirical data was a very good representation of the underlying factor structure described within the Competency framework (CFI=0.97; RMSEA=0.04).

Strategic Planning. All Strategic Planning alpha reliability statistics are displayed in Table 17. All reliabilities fell well within acceptable ranges.

Table 17. *Strategic Planning scale and sub-scale reliability statistics.*

Scale/sub-scale	Reliability
Strategic Planning (overall)	.90
Mission	.84
Vision	.86
Goals	.84
Planning	.83
Organization	.87

The results of the EFA analysis on the Strategic Planning scale can be found in Table 18. Three factors emerged with eigenvalues greater than 1.0, predicting 68% of the total variance found within the analysis. All items loaded strongly onto a single factor, while all items within a subscale loaded collectively onto a common factor. Significant loading overlap emerged between Mission and Vision items, as well as between Goals and Planning items.

Table 18. *Strategic Planning EFA individual item loadings.*

Scale Item	Factor 1	Factor 2	Factor 3
Mission1		-.74	
Mis2		-.85	
Mis3		-.83	
Vision1		-.79	
Vis2		-.88	
Vis3		-.80	
Goals1	.83		
G2	.80		
G3	.83		
Planning1	.75		
Pla2	.70		
Pla3	.80		
Organization1			.82
Org2			.91
Org3			.90

We conducted a maximum likelihood confirmatory factor analysis for the Strategic Planning scale, which suggested that the empirical data was a good representation of the underlying factor structure described within the Competency framework (CFI=0.95; RMSEA=0.07).

Discussion

The goal of this research study was to analyze the psychometric strength and appropriateness of the Student Leadership Competencies Inventory as a measure of the Student Leadership Competencies framework (Seemiller, 2013). We conducted measures of internal reliability, exploratory factor analysis, and confirmatory factor analysis. Our results provided support, in general, for the use of each scale within the Inventory, while also indicating some

specific weaknesses that should be addressed before considering each scale within the larger Inventory a “valid” measure of the overall theoretical framework.

The internal reliability results proved strong overall. Out of the 60 total sub-scales within the overall Competency framework, only four – Non-Verbal Communication (0.67), Productive Relationships (0.63), Collaboration (0.58), and Decision-making (0.57) – emerged as marginally reliable (between 0.60 and 0.69) or unreliable (below 0.60). These results suggest that other than these four sub-scales, participants who complete one of the eight individual assessment scales within the Inventory respond to survey items within each sub-scale relatively consistently.

The results of our EFA analyses suggested a more complex picture. Overall, very few items (eight in total, which includes all three of the Collaboration sub-scales within the Interpersonal Interaction scale) did not load to an acceptable level onto their respective sub-scales or too strongly onto another sub-scale. This indicated that these items – four among the 169 total – are not appropriate measures to continue to use within the eight measures. The remaining 161 individual survey items loaded to appropriate levels within relevant subscales. However, sub-scales designed to measure conceptually similar leadership competencies (e.g. Diversity and Inclusion subscales within the Civic Responsibility scale) loaded onto common factors rather than distinct sub-scale factors. Such common loading suggests that the degree of conceptual overlap across similar factors may have been too great for survey respondents to differentiate their competencies between and across distinct sub-scales. These results imply that students who completed the Diversity and Inclusion sub-scales, for example, received conceptually valid scores representing their capacity to lead with competence in the areas of diversity and inclusion, but that their separate scores between these two sub-scales may be nothing more than random variation. Additionally, competencies such as diversity and inclusion, mission and vision, and goals and planning, are often taught in tandem even though they are not identical, which may be a factor in why they factored together.

The results of our CFA analyses provide further support for the overall conceptual framework of the Student Leadership Competencies framework and the scale that has been developed to measure capacity for its implementation. All goodness of fit statistics fell within acceptable ranges ($CFI \geq .090$; $RMSEA \leq 0.07$) appropriate for concluding that the respective scale is, overall, a valid measure of the relevant underlying latent construct for which it was designed to assess.

Implications for Scale Use. Our results provide a nuanced picture of the validity of the Student Leadership Competencies Inventory, as well as of the arrangement of competencies within the framework (Seemiller, 2013) it is designed to assess. From one perspective, we found high internal validity across most sub-scales; EFA results that indicate that 95% of all items loaded onto their appropriate sub-scales, and CFA results that unambiguously suggest a good fit of the survey scales to their latent factors. These findings might simply suggest that the non-loading survey items be dropped, substitute items be inserted if relevant, and subsequent study be conducted to assess whether the new items load strongly enough on their appropriate factors. From another perspective, however, the degree of overlap between and across conceptually related competencies suggests the Competency framework itself might be too complex, or perhaps the scale items within overlapping competencies do not serve as appropriate

differentiation for survey completers. Recommendations based on these findings might include simplifying the framework (i.e. reduce the number of sub-scales so they no longer overlap), re-constructing the sub-scales that commonly load together, or offering transparency in the related competencies that factor together.

As the scale is currently constructed, students who complete the Group Dynamics scale within the Inventory, for example, would receive overall scale scores that, according to these results, are likely a statistically valid representation of their competence in regard to the management of group dynamics. At the same time, their Organizational Behavior and Power Dynamics sub-scale scores are likely not a valid representation of the differences between these two competencies, but instead an aggregate representation of both Organizational Behavior and Power Dynamics together. Therefore, scores for these individual competencies should be utilized more to help students reflect on areas for their growth in general rather than in evaluating themselves on the differences in scores.

In an effort to address the issues that emerged, we recommend that the five items that did not factor with their associated competencies be re-written using language that differentiates them from the wording of the items of related competencies.

Suggestions for Leadership Educators. The Student Leadership Competencies Inventory has been used since 2011 for educational purposes to help students uncover their self-perceived competency levels for the 60 Student Leadership Competencies (Seemiller, 2013). The main intent was to serve as a practical tool for student education and reflection and not a data collection instrument for research purposes (Seemiller). With these intents in mind, many practical uses emerge for the Inventory given our results. In general, using inventories can be an important component to a leadership program or course because they can provide a means for self-reflection and evaluation (Sitzmann, Ely, Brown, & Bauer, 2010). The Student Leadership Competencies Inventory fits this description, and as 93% of its sub-scales accurately reflected student competency levels, it overall can offer students a tool to self-assess their competency proficiency.

Having students complete the Inventory allows for the discovery of their relative strengths and weaknesses. Although students may over-inflate their responses (Kruger & Dunning, 1999), the relative ranking of their sub-scale scores can still provide insight into which competencies they should focus on for development. Students can then develop goals and plans around developing competencies critical for their success.

Campuses using the Student Leadership Competencies framework for program design, curriculum, and assessment, can integrate the self-assessment tool into their programmatic framework to offer a more comprehensive approach to leadership development. Without a tool such as this, educators would likely seek out an assessment that measured different competencies or capacities and then try to retrofit the assessment for use within a curricular framework based upon the Student Leadership Competencies. Finally, having an online tool that is available without a code, cost, or login allows for educators and students to easily access and instrument for self-awareness.

Conclusion. This research study was designed as a comprehensive, but initial, psychometric examination of the Student Leadership Competencies Inventory, a free online assessment tool designed to assess competency within the Student Leadership Competencies framework (Seemiller, 2013). Using a national dataset of university student participants, our results suggest that with minor exceptions regarding a small number of sub-scales, the Inventory serves as a valid measure of the framework.

References

- Bentler, P M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246.
- Bolden, R., & Gosling, J. (2006). Leadership competencies: Time to change the tune? *Leadership*, 2, 147–163.
- Brill, K., Croft, L., Ogle, J., Russell, S., Smedick, B., Hicks, M., & Coats, S. (2009). *Competency guide for college student leaders*. Columbia, SC: NACA.
- Conway, J. M., & Huffcutt, A. I. (2003). A review and evaluation of exploratory factor analysis practices in organizational research. *Organizational research methods*, 6(2), 147-168.
- Jolliffe, I. T. (2002). Principal component analysis and factor analysis. *Principal component analysis*, 150-166.
- Kouzes, J. M. & Posner, B. Z. (2013). *Student leadership practices inventory: Self report*. Retrieved from http://www.studentleadershipchallenge.com/userfiles/sample_student_lpi_self_online_self_report.pdf
- Kremelberg, D. (2010). *Practical statistics: A quick and easy guide to IBM® SPSS® Statistics, STATA, and other statistical software*. Thousand Oaks, CA: Sage Publications.
- Kruger, J. & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121-1134.
- McClelland, D. C. (1973). Testing for “competence” rather than intelligence. *American Psychologist*, 28, 1–14.
- Miguel, R.F. & Allen, S.J. (2016). Report of the validation of the Emotionally Intelligent Leadership for Students Inventory. *Journal of Leadership Education*, 15(4), 15-32.
- Nunnally, J. (1978). *Psychometric theory*. New York, NY: McGraw-Hill.

- Pearson, R. W. (2009). *Statistical persuasion: How to collect, analyze, and present data... accurately, honestly, and persuasively*. Thousand Oaks, CA: Sage Publications.
- Posner, B. Z. (2004). A leadership development instrument for students: Updated. *Journal of College Student Development*, 45(4), 443–456. DOI: 10.1353/csd.2004.0051.
- Santos, J. R. A. (1999). Cronbach's alpha: A tool for assessing the reliability of scales. *Journal of Extension*, 37(2), 1-5.
- Seemiller, C. (2011). *Student leadership competencies inventory*. Retrieved from www.studentleadershipcompetencies.com/evaluations/inventory.
- Seemiller, C. (2013). *The student leadership competencies guidebook*. San Francisco: Jossey-Bass.
- Spicer, J. (2005). *Making sense of multivariate data analysis: an intuitive approach*. Thousand Oaks, CA: Sage Publications.
- Croft, L. & Seemiller, C. (2017). Developing leadership competencies. In C. Seemiller (Ed.), *New Directions for Student Leadership*: No. 156. A competency based approach for student leadership development. San Francisco: Jossey-Bass.
- Seemiller, C. (2013). *The student leadership competencies guidebook*. San Francisco, CA: Jossey-Bass.
- Seemiller, C. & Murray, T. (2013). The common language of leadership. *Journal of Leadership Studies*, 7(1), 33-45.
- Sitzmann, T., Ely, K., Brown, K. G., & Bauer, K. N. (2010). Self-assessment of knowledge: A cognitive learning or affective measure? *Academy of Management Learning & Education*, 9, 169–191.

Author Biographies

David M. Rosch (dmrosch@illinois.edu) is an Associate Professor and coordinates the graduate program in Agricultural Education Program at the University of Illinois. He teaches undergraduate and graduate courses in leadership theory and human capacity development, and his research is focused on the effects of leadership education on youth and emerging adults.

Corey Seemiller (corey.seemiller@wright.edu) is a faculty member in Organizational Leadership at Wright State University. She has served as the co-chair for the National Leadership Symposium and the co-chair for the Leadership Education Academy. Dr. Seemiller is the author *The Student Leadership Competencies Guidebook* and *Generation Z Goes to College*.