

The Measurement of Socially Responsible Leadership: Considerations in Establishing Psychometric Rigor

John P. Dugan

Loyola University Chicago (USA)

doi: 10.7358/ecps-2015-012-duga

jdugan77@gmail.com

LA MISURAZIONE DELLA LEADERSHIP
SOCIALMENTE RESPONSABILE: CONSIDERAZIONI
SULLA DEFINIZIONE DEL RIGORE PSICOMETRICO

ABSTRACT

As educational institutions increase attention toward the preparation of future citizen leaders, the assessment of the efficacy of these interventions must increase in sophistication as well. Socially responsible leadership, and the social change model of leadership development from which it is derived, are among the most applied conceptual frameworks informing leadership education. Educational gains associated with this approach are typically assessed using the «Socially Responsible Leadership Scale» (SRLS), an instrument that has been used widely in educational research in the United States. This paper extends the psychometric understanding of the SRLS using findings from the Multi-Institutional Study of Leadership (MSL), an international research program examining developmental and educational influences on socially responsible leadership. To date, MSL has collected data in five countries at over 300 institutions of higher education with over 300,000 participants. Implications from this paper explore how the psychometric rigor of the SRLS can continue to evolve as well as considerations for the theoretical evolution of the social change model and socially responsible leadership.

Keywords: Assessment, Development, Leadership, Measurement social responsibility.

Although the preparation of future citizen leaders has long served as a key educational outcome, the past three decades have seen unprecedented attention to the topic repositioning leadership development as much more than just an educational byproduct but one that must be purposefully cultivated (Dugan, 2011; Komives, 2011). This attention, coupled with social outcry for leaders well prepared to address the complex and often intractable issues facing an increasingly global society, contributed to a surge in curricular and co-curricular educational interventions, scholarship, and theory building (Komives, 2011). Much of this work centers on advancing a particular approach to leadership grounded in ethics, shared responsibility, and democratic citizenship (Komives & Dugan, 2010). Socially responsible leadership, as derived from the social change model of leadership development (HERI, 1996), is consistent with this and among the few approaches employed in education with clear definitional/theoretical foundations, empirically supported literature, and practical applications. However, as the literature continues to expand there is a need to ensure that the alignment between conceptual/theoretical models, research, and practice is one characterized by psychometric rigor and continued refinement. The purpose of this paper is to provide insights into the measurement of socially responsible leadership and considerations for refining the conceptual and theoretical bases that inform it.

1. WHAT IS SOCIALLY RESPONSIBLE LEADERSHIP?

Socially responsible leadership is defined as a «purposeful, collaborative, values-based process that results in positive social change» (Komives, Wagner, & Associates, 2009, p. xii). This definition is consistent with contemporary leadership theory, which is characterized by a movement away from management, production, and command and control toward social justice, reciprocal relationships, and a concern for the common good (Komives & Dugan, 2010; Northouse, 2012). Indeed, leadership defined in this way aligns well with calls in both the United States and Europe for educational systems to better prepare citizen leaders ready to engage in a diverse, representative democracy (National Association of Student Personnel Administrators & American College Personnel Association, 2004; Association of American Colleges & Universities, 2007; Hoskins & Mascherini, 2009; European Commission, 2012).

Socially responsible leadership is theoretically grounded in the social change model of leadership development, which was created specifically for use in educational contexts and is among the most applied models in higher education (HERI, 1996; Kezar, Carducci, & Contreras-McGavin, 2006; Owen, 2012).

Table 1. – Value definitions for the social change model of leadership development.

VALUE	DEFINITION
<i>Individual Domain</i>	
Consciousness of self	Awareness of the beliefs, values, attitudes, and emotions that motivate one to take action.
Congruence	Thinking, feeling, and behaving with consistency, genuineness, authenticity, and honesty towards others; actions are consistent with most deeply-held beliefs and convictions.
Commitment	The psychic energy that motivates the individual to serve and that drives the collective effort; implies passion, intensity, and duration, and is directed toward both the group activity as well as its intended outcomes.
<i>Group Domain</i>	
Collaboration	To work with others in a common effort; constitutes the cornerstone value of the group leadership effort because it empowers self and others through trust.
Common purpose	To work with shared aims and values; facilitates the group's ability to engage in collective analysis of issues at hand and the task to be undertaken.
Controversy with civility	Recognizes two fundamental realities of any creative group effort: that differences in viewpoint are inevitable, and that such differences must be aired openly, but with civility. Civility implies respect for others, a willingness to hear each others' views, and the exercise of restraint in criticizing the views and actions of others.
<i>Community/Societal Domain</i>	
Citizenship	The process whereby an individual and the collaborative group become responsibly connected to the community and the society through the leadership development activity. To be a good citizen is to work for positive change on the behalf of others and the community.
Change	Change serves as the «hub» of the model reflecting the process of engaging in leadership to contribute to a better world.

Adapted from: HERI, 1996.

The social change model is comprised of seven leadership values and serves as both a process-model to inform group experiences as well as a developmental model that captures individual and collective learning necessary to engage in leadership for social change (HERI, 1996; Table 1). Values in the social change model reflect the intersection of knowledge, attitudes, and skills reflecting one's overall leadership capacity (Komives *et al.*, 2009). The values include consciousness of self, congruence, commitment, collabora-

tion, common purpose, controversy with civility, citizenship, and change. These values operate across three specific domains: the individual (i.e., consciousness of self, congruence, commitment), the group (i.e., collaboration, common purpose, controversy with civility), and broader community and society (i.e., citizenship) with the ability to engage in social change serving as the overarching principle. Table 1 provides definitional parameters for each of the values and domains in which they fall.

2. MEASUREMENT OF SOCIALLY RESPONSIBLE LEADERSHIP

Socially responsible leadership and the social change model from which it was derived were originally created as a conceptual framework for the leadership development of college students. Shortly after their articulation, Tyree (1998) created a measurement tool, the *Socially Responsible Leadership Scale* (SRLS), to operationalize the constructs. Her research followed standards of scale development moving through a process of expert review of 291 potential items, pilot tests of an eight scale instrument with each scale measuring a value associated with the social change model as well as change, and testing of the final 104-item instrument with a random sample of over 600 undergraduate students at a single institution in the United States. The last phase employed Confirmatory Factor Analysis (CFA) along with other analytic post-hoc tests to establish the reliability and validity of the scales. Reliability demonstrated adequate levels ranging from .69 to .92. Test-retest reliability was also established. Validity was largely assessed using expert reviewers in the item selection process for face validity along with correlational techniques to establish the construct validity of items within specific scales.

Unfortunately, use of the SRLS was limited in the years following its creation given the large number of items and need for statistical software to calculate and analyze scores. It was not until the publication of two studies in 2006 that any mention of the SRLS appeared in peer-review, academic journals. Dugan (2006a and 2006b) used a 103-item version of the instrument in two single-campus, exploratory studies examining influences of gender as well as campus-based involvement (i.e., community service, student organization membership, positional leader role attainment, participation in formal leadership training programs) in the measurement of socially responsible leadership. Reliability levels were consistent with those achieved in Tyree's (1998) research verifying that the scales had some degree of transferability across higher education contexts in the United States.

Since then, two large-scale studies in the United States employed the SRLS as a key outcome measure. The Multi-Institutional Study of Leadership (MSL; see: <http://www.leadershipstudy.net>) was initially designed to establish a national normative dataset against which student scores on the SRLS could be compared as well as advance a deeper understanding of influences on socially responsible leadership. The quantitative, cross-sectional study collected data at 52 institutions of higher education in the United States in 2006 with a final sample size exceeding 55,000 participants. The researchers employed a shortened version of the SRLS that was generated during pilot studies using standard data reduction techniques (DeVellis, 2012). This 68-item version, sometimes referred to as the SRLS-R2, performed well, but did demonstrate a reduction in the reliability on the Citizenship Scale albeit not below acceptable thresholds. This led to a recommendation to use a 71-item version of the SRLS in future research. Overall, researchers were pleased with the degree to which the constructs were replicated using CFA techniques and that reliability levels were consistent across all institutions as well as major demographic groups (Dugan, Komives, & Segar, 2008; Dugan & Komives, 2010). The MSL eventually expanded its purpose to examine the influences of higher education on leadership development broadly defined, but continues to use the SRLS and advance the psychometric refinement of the instrument. Currently, the MSL collects data every three years (e.g., 2009, 2012, 2015, 2018) with more than 300 schools in five countries having participated contributing to an overall sample exceeding 350,000 cases.

The Wabash National Study (see: <http://www.liberalarts.wabash.edu>) is a longitudinal, mixed-methods study that ran from 2006-2012 and examined the impact of the college environment on critical college outcomes. Data were collected from 49 institutions in the United States and included the 68-item version of the SRLS. Numerous published studies from this research advance the understanding of socially responsible leadership with most mirroring results found in MSL studies. The study also offered further evidence of the transferability of the scales across institutions and populations replicating factor structures and reliability levels achieved in other research.

3. MAJOR DEVELOPMENTS IN ADVANCING THE PSYCHOMETRIC RIGOR OF THE SRLS

The sections that follow detail major psychometric developments related to the SRLS along with implications for the social change model that grounds it.

3.1. *Considerations of cross-sectional and self-report designs*

The use of both cross-sectional and self-report designs in educational research has rightly elicited concerns regarding quality and accuracy (Pascarella, 2001; Bowman & Seifert, 2011; Herzog & Bowman, 2011; Porter, 2011). Although the SRLS has been used in longitudinal studies, it is most frequently used in programmatic assessment with a cross-sectional design. Of most significant concern in cross-sectional research is the lack of adequate controls to assess change over time (Pascarella, 2001). However, scholars do recognize that not every study can or should be longitudinal in nature stressing the importance of using any form of control over none at all (Pascarella, 2001). This is the approach used in MSL, which employs retrospective questions to capture data about students' perceptions, knowledge, and experiences prior to enrollment in college. The use of retrospective questioning allows for the creation of quasi-pretests affording some degree of control when examining change across the SRLS measures. Additionally, some research on leadership outcome measurement suggests the appropriateness of cross-sectional designs due to concerns related to response-shift bias (Dailey, 1979; Howard, 1980; Rohs & Langone, 1997; Howard & Rohs, 2002). These scholars argue that accuracy in longitudinal measurement is predicated on the assumption that:

A person's standard for measurement of the dimension being assessed will not change from pretest to posttest. If the standard of measurement were to change, the posttest ratings would reflect this shift in addition to the actual changes in the person's level of functioning. (Rohs & Langone, 1997, p. 51)

Given the understanding of leadership is shaped significantly by cognitive reasoning (Komives, Longenecker, Owen, Mainella, & Osteen, 2006; Day, Harrison, & Halpin, 2009), any shift in cognition could alter the interpretation of items between pretest and posttest violating the assumption of the same standard of measurement and distorting the internalized scale.

Numerous concerns arise related to the use of self-reported data (Bowman & Seifert, 2011; Herzog & Bowman, 2011; Porter, 2011). These can largely be organized into three themes: social desirability, the halo effect and clarity of measures, and item format. Through pilot tests and on-going psychometric research on the SRLS, MSL researchers have explored these concepts and continued to evolve the rigor of the SRLS when used in self-report and cross-sectional designs. This is further bolstered by studies specifically on the topic of leadership, which found self-reports to be generally accurate (Turrentine, 2001; Posner, 2012). For example, Posner (2012) used a large, international sample to examine self-report versus peer observations across five leadership behaviors (i.e., inspiring a shared vision, enabling

others to act, encouraging the heart, modeling the way, and challenging the process) consistent with socially responsible leadership and found only small effect size differences with observers rating higher than self-evaluations.

3.1.1. Social desirability

Scholars express considerable concern regarding the potential for socially desirous responding (i.e., purposefully inflating or deflating responses based on perceived favorability/non-favorability of a particular answer) in self-report research (Gonyea, 2005; Bowman & Seifert, 2011; Herzog & Bowman, 2011; Porter, 2011). Studies demonstrate the ways in which social desirability skews results (Bowman & Hill, 2011; Porter, 2011) validating Porter's (2011) concern that «social desirability on college student surveys may lead to distorted and misleading conclusions» (p. 74). Consistent with recommendations (Gonyea, 2005; Porter, 2011), Tyree (1998) embedded the *Marlowe-Crowne Scale of Social Desirability* (Crowne & Marlowe, 1960) in her original research that led to the creation of the SRLS. Results were used to remove any items that correlated highly with measures of social desirability. Given perceptions of social desirability may change over time, the MSL later included the *Marlowe-Crowne Scale* in pilot studies. Results again yielded no concerning relationships between items on the SRLS and socially desirous responding significantly alleviating concerns. Some research demonstrates that social desirability is most problematic with students early in their college experience (Bowman & Hill, 2011). Thus, for those looking for the most conservative approach to addressing social desirability when using the SRLS, analytic approaches such as the one employed in Dugan and Komives (2010) using a sample of only college seniors may be most appropriate.

3.1.2. The halo effect and clarity of measures

The halo effect is a well-documented phenomenon in self-report research in which participants' judgments on a particular item are influenced by more general perceptions or the perception of a shared relationship between concepts (Cooper, 1981; Pike, 1999; Bowman & Hill, 2011; Porter, 2011). Several scholars argue that the halo effect potentially increases based on the level of abstraction of the concepts being measured (Gonyea, 2005; Rosch & Schwartz, 2009; Herzog & Bowman, 2011). Suggestions for accommodating this involve greater attention toward the overall clarity of measures. This involves crafting survey items that offer clearly defined terms and avoid

broad or vague constructs (Gonyea, 2005; Rosch & Schwartz, 2009; Herzog & Bowman, 2011; Porter, 2011).

The SRLS is a longer instrument by design as it allows for greater attention to clarity and specificity of constructs. None of the items use the word leadership, instead measuring specific attitudes, skills, and beliefs related to the varying components of socially responsible leadership. This reduces the danger of students' developmental interpretations of the term leadership introducing significant variance in responses (Dugan, 2011). It also affords a greater degree of clarity and differentiation of concepts being studied. Prior to the launch of the MSL, pilot studies were conducted to examine comprehension of items and clarity of constructs. Qualitative interviews with students representing a wide array of demographics and collegiate experiences (e.g., highly engaged to minimally engaged) confirmed comprehension of items. Participants were administered the instrument and instructed to mark questions that were confusing or for which they hesitated in their response. Debriefing explored problem questions focusing on how they could be improved for clarity, comprehension, and ease of responding accurately or whether they should be removed. Finally, similar to findings related to socially desirous responding, the halo effect appears to be more pronounced with samples of students early in their college careers (Pike, 1999; Bowman & Hill, 2011). Using the SRLS specifically with samples of college seniors may ameliorate further concerns related to the halo effect in cross-sectional studies.

3.1.3. Item format

Another major concern in using self-report data lies with question formatting. Scholars strongly urge against the use of formats that ask students to identify the degree to which the collegiate environment or a specific intervention have influenced educational gains (Bowman & Seifert, 2011; Porter, 2011). Of particular concern is the fact that «estimating one's own gains requires much greater cognitive effort and self-knowledge than making a judgment about one's current abilities» (Bowman & Seifert, 2011, p. 271). Furthermore, response options should always account for the lack of a potential influence and avoid use of response categories that presume at least a minimal impact or only offer restricted ranges (Herzog & Bowman, 2011). Finally, questions should realistically evaluate the ability of subjects to recall information both in terms of time duration (e.g., how much time has passed since the experience) and accuracy of summation (e.g., what degree of specificity is required in counting quantity; Gonyea, 2005; Herzog & Bowman, 2011; Porter, 2011).

First and foremost, the SRLS items do not rely on perceived growth prompts alleviating the concerns that come with this approach. Instead, the

questions ask respondents to assess their abilities at the point in time in which the instrument is completed. Response options are attitudinal ranging from *strongly disagree* (1) to *strongly agree* (5) across specific points of knowledge, skills, and values associated with socially responsible leadership. This avoids the presumption of impact. The SRLS is used with a cross-sectional design in MSL, which does introduce concerns about recall. However, qualitative interviews conducted during pilot studies did not indicate issues associated with the ability to retrieve necessary information when students were asked to reflect back on their abilities prior to entry into college.

3.2. *Considerations of validity*

Pike (2011) argued that conceptualizations of validity are largely flawed and would be better served by an analysis of three component parts: content (i.e., the degree to which the total construct being measured is adequately and accurately represented), construct (i.e., degree of alignment between empirical and theoretical structures), and criterion (i.e., theoretically appropriate relationships with other constructs). Content validity for the SRLS was largely established through Tyree's (1998) extensive research. However, the instrument was not examined at all for structural or criterion validity. The sections that follow offer insights as derived from MSL research.

3.2.1. Content validity

Content validity was re-visited through an expert review process to assess the degree to which the SRLS was measuring all facets of the socially responsible leadership construct. This was particularly important given the rapidly evolving differentiation of domains within the leadership development literature. It was clarified that the SRLS measures dimensions of leadership capacity or an individual's knowledge, attitudes, and skills representing an overall ability (Dugan, 2012). Leadership capacity may or may not be enacted suggesting a differentiation from enactments (i.e., leadership behaviors). Leadership capacity is also informed by leadership efficacy (i.e., one's internal belief in the likelihood of success when engaging in leadership) and leadership motivations (i.e., one's underlying desire to engage in leadership) as well as numerous other developmental constructs (Chan & Drasgow, 2001; Kark & Van Dijk, 2007; Hannah, Avolio, Luthans, & Harms, 2008; Day *et al.*, 2009). The review established that particularly through using items reflecting know, being, and doing orientations the SRLS was in alignment with the measurement of leadership capacity and distinct from efficacy, motivation, and behaviors.

The examination of content validity also explored the degree to which the SRLS represented the various aspects of the social change model from which it was derived. This led to the removal of the *Change Scale* when using the SRLS in the MSL as the measure was sound, but the latent construct being measured did not adequately align with the theoretical conceptualization. In the social change model, the original authors described change as «the value ‘hub’ which gives meaning and purpose to the 7 C’s. Change, in other words, is the ultimate goal of the creative process of leadership – to make a better world and a better society for self and others» (HERI, 1996, p. 21). However, the items comprising the *Change Scale* in the SRLS largely capture one’s skills and comfort with transition in lieu of one’s overarching ability to engage in social change work. Sample items included: «I work well in changing environments» and «Transition makes me uncomfortable». The expert review indicated that a total score across all measures of the SRLS better reflected the content described by the original authors of the social change model than the existing *Change Scale*.

3.2.2. Structural validity

According to Pike (2011) «theories of the constructs that educational measures are intended to represent help to define the criteria used to evaluate validity, establish expectations regarding relationships among measures, and assist in interpreting results» (p. 54). Given the firm grounding in the social change model of leadership development, the SRLS was tested for structural validity as a measure of seven values operating across three domains all of which contribute to one’s overall capacity for socially responsible leadership. The goals of analyses were to reduce the total number of items that comprise the SRLS enhancing usability as well as advancing the psychometric rigor of its measurement. Three major results emerged: removal of the common purpose scale, reduction in total items, and validation of the original conceptual model.

A random sample of 1,430 cases from the MSL 2009 full sample was selected and reflected the diversity present in the full sample while also offering strong statistical power. CFA using *Structural Equation Modeling* (SEM) examined the underlying factor structure of the 71-item version of the SRLS, but omitted the *Change Scale* for the above stated reasons. Bootstrapping was employed to estimate the mean and standard errors of loadings, factor correlations, unique error terms, and squared multiple correlations for the final set of items.

Removal of common purpose. Initial analyses revealed that the *Collaboration and Common Purpose Scales* were for the most part conceptually redundant (i.e., respondents did not use the two sets of items to distinguish

separate constructs). When measurement error was partialled out, the two factors were highly correlated ($r = .95$). The presence of unreliability (i.e., measurement error) in the constituent items for each factor watered down the apparent strength of the interrelationship, thus making *Collaboration and Common Purpose* seem more independent when analyzing composite indices for each scale than when examining latent variables. To address this issue, a subset of five items for each factor was identified that provided strong reliability. However, re-running the CFA using this reduced subset of indicators for each of the two factors only lowered the intercorrelation to .93, which was still above desired thresholds.

An examination of items comprising the latent constructs of collaboration and common purpose demonstrated some overlap adding further confusion as to whether the issue was based on measurement (i.e., the original design of the scales for the SRLS did not accurately capture the constructs as theoretically defined) or conceptualization (i.e., the original concepts as theorized could not be operationalized in practice). Testing with alternative measures representing similar concepts (e.g., inspiring a shared vision, inspirational motivation) did not improve models demonstrating similar levels of correlation with the *Collaboration Scale*. This suggested that perhaps the issue was more conceptual. An examination of the definitional parameters for common purpose and collaboration in the original social change model publication seems to support the statistical finding as common purpose is defined as a *function of collaboration and characteristic of the group*, not necessarily a capacity of individuals within the group or even of the group itself (HERI, 1996). When the social change model is used as a process model instead of a developmental one, common purpose «provides the form and direction for the collaborative effort» (HERI, 1996, p. 57). Conceptually, this entangles the constructs confounding individual leadership capacity with group functioning. As such, a recommendation was made to remove the *Common Purpose Scale* from the SRLS in future MSL research.

Reduction of total items. To increase the utility of the SRLS further, item reduction was desired. The goal was to keep the full instrument as short as possible while preserving the content validity and reliability of the factors and avoiding excessive increases in the correlations among factors. Statistical analyses suggested that negative response items routinely demonstrated lower factor loadings than other items. This is likely because respondents often find it easier to endorse items that are true rather than rejecting items that are false when evaluating themselves. Thus, removing these items reduced the overall cognitive burden of the instrument as well as shortening its length. This resulted in a 34-item instrument comprised of six total scales (i.e., consciousness of self, congruence, commitment, collaboration, controversy with

civility, and citizenship) that: (a) fit the data well by all the targeted measures of goodness-of-fit (i.e., RMSEA, SRMR, CFI, NNFI); (b) had either 5 or 6 items for each factor and demonstrated good reliability (i.e., all alphas \geq .80); and (c) showed reasonable factor inter-correlations.

Validation of measurement model. The social change model asserts that leadership capacity functions at three separate levels: six unique values representing specific capacities (i.e., consciousness of self, congruence, commitment, collaboration, controversy with civility, and citizenship), three domains that cluster the values (i.e., individual, group, and societal), and at an omnibus level reflecting one's overall capacity for socially responsible leadership (HERI, 1996). Four relevant measurement models were calculated along with goodness-of-fit statistics (see Table 2). The one-factor model had an RMSEA above the criterion of acceptability, but the three-factor, six-factor, and single second-order model with six first-order factors all showed acceptable goodness-of-fit to the data by all four fit criteria. In addition, the three-factor model (difference $\chi^2 = 2196.56$, difference $df = 3$, $p < .00001$) and six-factor model (difference $\chi^2 = 3260.153$, difference $df = 3$, $p < .00001$) both fit the data significantly better than the one-factor model. Additionally, the six-factor model fit the data significantly better than the three-factor model (difference $\chi^2 = 1063.593$, difference $df = 12$, $p < .00001$). Note that although the three-factor model does not fit the data as well as the six-factor model, both of these models provide acceptable goodness-of-fit statistics.

Bentler and Mooijart (1989) suggested that if more than one model provides acceptable fit to the data, then one should choose the model that is most parsimonious. Theoretically, however, the social change model suggests that these constructs operate at all three levels simultaneously (i.e., value, domain, overall capacity). The results of these analyses confirm that measurement at all three levels is possible as each meets statistical requirements.

Table 2. – Measurement models for socially responsible leadership.

MODEL	CHI-SQUARE	DF	RMSEA	SRMR	CFI	NNFI
One Factor	6,626.23	527	.107	.068	.917	.912
Three Factors	4,429.67	524	.079	.056	.955	.952
Six Factors	3,366.08	512	.065	.056	.970	.967
One Second-Order and Six First-Order Factors	3,529.21	521	.066	.055	.969	.966

Further bootstrapping analyses were calculated for the three measurement models (i.e., three-factor, six-factor, and second-order CFA models) using the 34-item SRLS. For each model, 5,000 random samples were drawn with replacement of size 997 (1% of total sample, which provides 12 cases for each estimated parameter in the largest model) from the total sample of 96,642 cases from MSL 2009. Each model was calculated using each of its 5,000 data sets one at a time storing the model fit statistics for each of the 5,000 CFA solutions for each model. Each of these three stored, bootstrapped data sets of fit statistics were analyzed to compute mean fit statistics and their standard errors as well as to compute a 95% confidence interval for each fit statistic for each model. Each of the three models was highly stable across the 5,000 replications demonstrating strong generalizability across the full sample. Additionally, this led to more consistent reliabilities across the measures ($\alpha = .82-.90$). This furthers the structural validity of the 34-item SRLS in effectively measuring socially responsible leadership and the social change model that informs it.

3.2.3. Criterion validity

Criterion validity is focused on demonstrating theoretically appropriate relationships between and among constructs. This was achieved in a variety of ways for the SRLS through MSL research. Validation studies for the MSL embedded a number of other theoretical constructs associated with leadership. Specifically, the factors associated with transformational leadership and extensively validated using the *Multifactor Leadership Questionnaire* (MLQ) were examined for relationships with the overall measurement of socially responsible leadership. Transformational leadership is identified as part of the post-industrial leadership paradigm (i.e., emphasis on shared processes, ethics, and the common good) as is socially responsible leadership (Komives & Dugan, 2010). The MLQ measures four factors associated with transformational leadership (i.e. inspirational motivation, intellectual stimulation, idealized influence, and individualized consideration) along with two factors negatively related to transformational leadership (i.e., management by exception and laissez-faire; Avolio & Bass, 2004). Correlational analyses demonstrated convergent validity with the four factors associated with transformational leadership ($r \sim .63$). These results also indicated that socially responsible leadership was related, but not identical, to transformational leadership suggesting they are both distinct but part of the same paradigmatic understanding of leadership. Discriminant validity was established with the negatively associated factors. Socially responsible leadership capacity was inversely related to management by exception ($r = -.33$) as well as the laissez-faire approach ($r = -.42$). Additionally, as previously stated the SRLS

measures leadership capacity, which is just one domain informing leadership development (Dugan, 2012). MSL data have been used to establish the distinctiveness of socially responsible leadership capacity from leadership efficacy, leadership behaviors, and a variety of other leadership-related constructs (Dugan, Kodama, Correia, & Associates, 2013).

4. IMPLICATIONS AND FUTURE DIRECTIONS

Psychometric work with the SRLS is critically important as it advances a theoretical grounding of leadership as an educational outcome. Unfortunately, too much of the leadership and college impact literature is atheoretical limiting the opportunities to deeply examine validity (Dugan, 2011). This paper provides previously unreported information regarding the measurement of socially responsible leadership using the SRLS. It offers insights into the justification of its use as well as support for the validity of a 34-item version of the instrument. Given the broad support for the social change model of leadership and its frequent usage in educational programs targeting college students (Kezar *et al.*, 2006; Owen, 2012), this information is critical in advancing evidence-based practice. There remain, however, a number areas for further exploration.

A major finding advancing the use of the SRLS is the establishment of its structural validity. The original authors of the social change model should be commended for their sophisticated conceptualization. Similarly, Tyree (1998) should be commended for her operationalization of these concepts. There is elegance to the conceptualization of the social change model's three levels (i.e., values, domains, overall capacity) and the ability to justify measurement at each of these levels. Further research is needed that examines the parts (i.e., values, domains) as well as the whole (i.e., overall capacity) in the process of leadership development. The reduction of the instrument from 104-items in its original format to just 34-items while retaining quality of measurement should support the use of the tool more widely. Those interested in conducting research using the SRLS are encouraged to contact the National Clearinghouse of Leadership Programs (<http://www.nclp.umd.edu>) to obtain necessary permissions.

Conversely, validity studies also pointed to the need for refinement of the SRLS. First, the measure of change did not accurately capture the definitional parameters provided by the original authors. The overall measure of socially responsible leadership capacity offers a much more theoretically congruent as well as effective approach. Second, conflation in measurement of collaboration and common purpose was problematic. When the social change model is used as a process model for groups engaging in leadership,

common purpose offers valuable insights. When the social change model is used as a developmental model focused on building individual capacities, however, common purpose could not be uniquely measured apart from collaboration. Similar overlap when related constructs (i.e., inspiring a shared vision, inspirational motivation) were substituted seems to support the conflation. Pike (2011) asserted that the inability to replicate the theoretical factor structure of constructs is hardly reason for wholesale abandonment, but a useful tool in the process of refining both theory and measurement suggesting the need for increased definitional clarity and refinement. The original authors of the social change model would appear to agree. They attest that the social change model:

Is presented as a working framework that is subject to regular revision and refinement based on the experience of those who use it. Practitioners and students may well find certain elements in the model to be more applicable or relevant than others (HERI, 1996, p. 18).

Thus, the social change model and socially responsible leadership represent «living» concepts best when allowed to evolve to reflect emerging empirical evidence. Results from this research would suggest the importance of further theorizing regarding the role of common purpose as an individual leadership capacity. Until this is better understood, those using the SRLS are encouraged to omit the *Common Purpose Scale* to avoid over inflation of the weight of common purpose and collaboration in studies measuring socially responsible leadership at the domain or overall capacity levels.

Understanding directionality in the social change model is a further area of study that could extend validity work as well as deepen the understanding of leadership development as a process. The original authors of the model hypothesized that each of the domains (i.e., individual, group, societal) interacted with one another contributing to «a continuous feedback loop» (HERI, 1996, p. 19). Thus, bidirectional arrows appear between each of the domains. However, when MSL researchers employed path analytic techniques using SEM, they found that the relationships between and among domains in the model were much more sequential than hypothesized (Dugan *et al.*, 2013; Dugan, Bohle, Woelker, & Cooney, 2014). These studies have not, for example, found direct pathways between the individual domain and the societal domain. Rather, the effect of the individual domain on the societal domain is indirect with the group domain serving as a mediator. This could be due to the cross-sectional nature of the MSL research design, which does not allow for recursive analyses. However, MSL researchers also asserted that this might reflect another example of confounding the process versus development orientations that characterize the social change model (Dugan *et al.*, 2014). Additional research, particularly longitudinal designs, is needed

to help clarify the interactions between and among the constructs comprising the social change model.

Finally, the significant validation work reported here supports the use of the SRLS in the United States. Additional studies are necessary to determine the degree of validity in its application outside the U.S. context. The National Clearinghouse for Leadership Programs currently coordinates permissions for use of the SRLS. These permissions have been granted to a wide range of researchers across the globe including in Kenya, Colombia, China, Japan, Italy, and Lithuania just to name a few. Problematically, however, usage of the SRLS across global contexts is not regulated in terms of the process for establishing cultural and language translations nor for re-assessing the validity of the instrument. Furthermore, it is often unclear which version of the SRLS the researchers employed with few following recommendations to use the 34-item version of the instrument detailed here. Only in rare cases do the researchers publish findings in peer review journals related to the cross-cultural adaptation and measurement issues that may arise (Humphreys, 2011; du Mérac, 2014).

MSL researchers have collected data using the SRLS in Canada, Jamaica, and Mexico following a standardized protocol for cross-cultural adaptations. This protocol involves first determining the need for a non-English version of the instrument. In cases where this is appropriate, translation standards are employed with initial translation by native speakers, back-translation by non-native speakers, and cross-referencing these documents for clarity of meaning (Wild *et al.*, 2005). Cultural translations are then conducted to determine any basic language changes that may be necessary (e.g., terminology differences between college and university in question prompts) as well as alter any items that may not hold the same meaning across cultural contexts. Finally, both structural and criterion validity work is conducted similar to that which was previously reported in the above sections. An example of the cross-cultural validation work conducted by the MSL can be seen in part in a publication by Dugan, Rossetti-Morosini and Beazley (2011) examining the use of the SRLS in Mexico. However, much more extensive work is necessary to determine the broad transferability of the SRLS across national boundaries and scholars are encouraged to publish these results.

5. CONCLUSION

As use of the social change model of leadership development continues to expand both across educational contexts and globally, so to does the need to understand how best to advance the theory, research, and practice that

inform it. The primary vehicle for this is the SRLS. This paper offered insights into the nature of the measurement tool, considerations for its use, and results from validation studies attempting to expand the understanding of how it best functions. These insights have direct implications for research and assessment as well as future theory building.

REFERENCES

- Association of American Colleges & Universities (2007). *College learning for the new global century*. Washington, DC: Author.
- Avolio, B. J., & Bass, B. M. (2004). *Multifactor Leadership Questionnaire: Manual and sampler set* (3rd ed.). Menlo Park, CA: Mindgarden.
- Bentler, P. M., & Mooijart, A. (1989). Choice of structural model via parsimony: A rationale based on precision. *Psychological Bulletin*, *106*, 315-317.
- Bowman, N. A., & Hill, P. L. (2011). Measuring how college affects students: Social desirability and other potential biases in college student self-reported gains. In N. A. Bowman, & S. Herzog (Eds.), *Validity and limitations of student self-report data* («New Directions in Institutional Research», no. 150, pp. 73-85). San Francisco: Jossey-Bass.
- Bowman, N. A., & Seifert, T. A. (2011). Can college students accurately assess what affects their learning and development? *Journal of College Student Development*, *52*, 270-290.
- Chan, K., & Drasgow, F. (2001). Toward a theory of individual differences and leadership: Understanding the motivation to lead. *Journal of Applied Psychology*, *86*, 481-498.
- Cooper, W. (1981). Ubiquitous halo. *Psychological Bulletin*, *90*, 218-244.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Counseling Psychology*, *24*, 349-354.
- Day, D. V., Harrison, M. M., & Halpin, S. M. (2009). *An integrative approach to leader development*. New York: Taylor & Francis.
- DeVellis, R. F. (2012). *Scale development* (3rd ed.). Thousand Oaks, CA: Sage.
- Dugan, J. P. (2006a). Explorations using the social change model: Leadership development among college men and women. *Journal of College Student Development*, *47*, 217-225.
- Dugan, J. P. (2006b). Involvement and leadership: A descriptive analysis of socially responsible leadership. *Journal of College Student Development*, *47*, 335-343.
- Dugan, J. P. (2011). Research on college student leadership. In S. R. Komives, J. P. Dugan, J. E. Owen, W. Wagner, C. Slack, & Associates, *Handbook for student leadership development* (pp. 59-84). San Francisco: Jossey-Bass.

- Dugan, J. P. (2012). Exploring local to global leadership education assessment. In K. L. Guthrie & L. Osteen (Eds.), *Developing students' leadership capacity* («New Directions for Student Services», no. 140, pp. 89-101). San Francisco: Jossey-Bass.
- Dugan, J. P., Bohle, C. W., Woelker, L., & Cooney, M. A. (2014). The role of social perspective-taking in developing students' leadership capacities. *Journal of Student Affairs Research and Practice*, 51(1), 1-15.
- Dugan, J. P., Kodama, C., Correia, B., & Associates (2013). *Multi-Institutional Study of Leadership insight report: Leadership program delivery*. College Park, MD: National Clearinghouse for Leadership Programs.
- Dugan, J. P., & Komives, S. R. (2010). Influences on college students' capacity for socially responsible leadership. *Journal of College Student Development*, 51, 525-549.
- Dugan, J. P., Komives, S. R., & Segar, T. C. (2008). College student capacity for socially responsible leadership: Understanding norms and influences of race, gender, and sexual orientation. *NASPA Journal*, 45, 475-500.
- Dugan, J. P., Rossetti Morosini, A. M., & Beazley, M. R. (2011). Cultural transferability of socially responsible leadership: Findings from the United States and Mexico. *Journal of College Student Development*, 52, 456-474.
- du Mérac, E. R. (2014). The measurement of leadership attitudes of adolescents in two educational contexts: School and scouting. *Italian Journal of Educational Research*, 6(11), 95-111.
- European Commission (2012). *Citizenship education in Europe*. Bruxelles: Educational, Audiovisual, and Culture Executive Agency.
- Gonyea, R. M. (2005). Self-reported data in institutional research: Review and recommendations. In P. D. Umbach (Ed.), *Survey research* («New Directions in Institutional Research», no. 127, pp. 73-89). San Francisco: Jossey-Bass.
- Hannah, S. T., Avolio, B. J., Luthans, F., & Harms, P. D. (2008). Leadership efficacy: Review and future directions. *Leadership Quarterly*, 19, 669-692.
- Herzog, S., & Bowman, N. A. (Eds.). (2011). Validity and limitations of student self-report data. («New Directions for Institutional Research», no. 150). San Francisco: Jossey-Bass.
- HERI – Higher Education Research Institute (1996). *A social change model of leadership development*. College Park, MD: National Clearinghouse for Leadership Programs.
- Hoskins, B. L., & Mascherini, M. (2009). Measuring active citizenship through the development of a composite indicator. *Social Indicators Research*, 90, 459-488.
- Howard, G. S. (1980). Response shift bias: A problem in evaluating interventions with pre/post self-reports. *Evaluation Review*, 4, 93-106.
- Howard, G. S., & Dailey, P. R. (1979). Response-shift bias: A source of contamination in self-report measures. *Journal of Applied Psychology*, 64, 144-150.

- Humphreys, M. J. (2011). A new generational of leaders for Eastern Europe: Values and attitudes for active citizenship. *Christian Higher Education, 10*, 215-236.
- Kark, R., & Van Dijk, D. (2007). Motivation to lead, motivation to follow: The role of the self-regulatory focus in leadership processes. *Academy of Management Review, 32*, 500-528.
- Kezar, A. J., Carducci, R., & Contreras-McGavin, M. (2006). Rethinking the «L» word in higher education. *ASHE Higher Education Report, 31*(6). San Francisco: Jossey-Bass.
- Komives, S. R. (2011). Advancing leadership education. In S. R. Komives, J. P. Dugan, J. E. Owen, W. Wagner, C. Slack, & Associates, *Handbook for student leadership development* (pp. 1-34). San Francisco: Jossey-Bass.
- Komives, S. R., & Dugan, J. P. (2010). Contemporary leadership theories. In R. A. Couto (Ed.), *Political and civic leadership* (pp. 109-125). Thousand Oaks, CA: Sage.
- Komives, S. R., Longerbeam, S., Owen, J. O., Mainella, F. C., & Osteen, L. (2006). A leadership identity development model: Applications from a grounded theory. *Journal of College Student Development, 47*, 401-418.
- Komives, S. R., Wagner, W., & Associates (2009). *Leadership for a better world*. San Francisco: Jossey-Bass.
- National Association of Student Personnel Administrators & American College Personnel Association (2004). *Learning reconsidered*. Washington, DC: NASPA - ACPA.
- Northouse, P. G. (2012). *Leadership: Theory and practice* (6th ed.). Thousand Oaks, CA: Sage.
- Owen, J. E. (2012). *Findings from the Multi-Institutional Study of Leadership institutional survey*. College Park, MD: National Clearinghouse for Leadership Programs.
- Pascarella, E. T. (2001). Using student self-reported gains to estimate college impact: A cautionary tale. *Journal of College Student Development, 42*, 488-492.
- Pike, G. (1999). The constant error of the halo in educational outcomes research. *Research in Higher Education, 40*, 61-86.
- Pike, G. (2011). Using college students' self-reported learning outcomes in scholarly research. In N. A. Bowman & S. Herzog (Eds.), *Validity and limitations of student self-report data* («New Directions in Institutional Research», no. 150, pp. 41-58). San Francisco: Jossey-Bass.
- Porter, S. R. (2011). Do college student surveys have any validity? *The Review of Higher Education, 35*, 45-76.
- Posner, B. Z. (2012). Effectively measuring student leadership. *Administrative Sciences, 2*, 221-234.
- Rohs, F. R. (2002). Improving the evaluation of leadership programs: Control response shift. *Journal of Leadership Education, 1*(2), 1-12.
- Rohs, F. R., & Langone, C. A. (1997). Increased accuracy in measuring leadership impacts. *Journal of Leadership Studies, 4*(1), 150-158.

- Rosch, D. M., & Schwartz, L. M. (2009). Potential issues and pitfalls in outcomes assessment in leadership education. *Journal of Leadership Education*, 8(1), 177-194.
- Turrentine, C. G. (2001). A comparison of self-assessment and peer assessment of leadership skills. *NASPA Journal*, 38, 361-371.
- Tyree, T. M. (1998). Designing an instrument to measure socially responsible leadership using the social change model of leadership development. *Dissertation Abstracts International*, 59(6), 1945. AAT 9836493.
- Wabash National Study, see: <http://www.liberalarts.wabash.edu>
- Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S. Verjee-Lorenz, A. et al. (2005). Principals of good practice for the translation and cultural adaptation process for patient-reported outcomes measures: Report from the ISPOR task force for translation and cultural adaptation. *Value in Health*, 8, 94-104.

RIASSUNTO

Così come le istituzioni educative vedono crescere l'attenzione verso la preparazione dei cittadini che diventeranno futuri dirigenti, anche le procedure per la valutazione dell'efficacia di tali interventi dovranno migliorare e diventare più sofisticate. La leadership socialmente responsabile e il modello di cambiamento sociale, in cui si sviluppa tale leadership e da cui essa deriva, sono tra i sistemi concettuali che maggiormente ispirano il concetto di leadership in educazione. In ambito educativo i vantaggi ottenuti da questo approccio sono generalmente valutati attraverso la scala «Socially Responsible Leadership» (SRLS), uno strumento che è utilizzato ampiamente nella ricerca educativa negli Stati Uniti. Questo studio si propone di estendere la misurazione psicometrica della SRLS utilizzando i risultati dello Studio Multi-Istituzionale della Leadership (MSL): un programma di ricerca internazionale che esamina le influenze evolutive ed educative sulla leadership socialmente responsabile. Fino ad oggi attraverso l'MSL si sono raccolti dati in cinque paesi, in oltre 300 istituti di istruzione superiore con più di 300.000 partecipanti. Questo articolo intende mostrare come il rigore psicometrico della SRLS può continuare a svilupparsi, così come il modello teorico di cambiamento sociale e di leadership socialmente responsabile.

Parole chiave: Leadership, Misurazione, Responsabilità sociale, Sviluppo, Valutazione.

How to cite this Paper: Dugan, J. P. (2015). The measurement of socially responsible leadership: Considerations in establishing psychometric rigor [La misurazione della leadership socialmente responsabile: considerazioni sulla definizione del rigore psicometrico]. *Journal of Educational, Cultural and Psychological Studies*, 12, 23-42. doi: 10.7358/ecps-2015-012-duga