# Development of the Dynamic Leadership in Counseling Scale – Self-report

By: W. Bradley McKibben, L. DiAnne Borders, and Terry A. Ackerman

McKibben, W. B., Borders, L. D., & Ackerman, T. A. (2017). Development of the Dynamic Leadership in Counseling Scale – Self-report. *Journal of Counseling Leadership and Advocacy, 4*(2), 147-160. https://doi.org/10.1080/2326716X.2017.1333048

**This is an Accepted Manuscript of an article published by Taylor & Francis in *Journal of Counseling Leadership and Advocacy* on 12 June 2017, available online: http://www.tandfonline.com/10.1080/2326716X.2017.1333048.**



## Abstract:

The authors developed the Dynamic Leadership in Counseling Scale–Self-Report (DLCS-SR) and tested for evidence for validity and internal consistency with a sample of 218 participants. They found evidence for a single-factor model of global leadership behaviors among counselors in the current sample as well as evidence for convergent validity and strong internal consistency. Implications for counseling leadership research and practice are discussed in light of the findings.

**Keywords:** Counseling | counselor | leadership | measurement

## Article:

In recent years, the production of counseling leadership research has not matched the steadily increasing emphasis on leadership in counselor education, training, and practice. Paradise, Ceballos, and Hall (2010) referred to counseling leadership behaviors as neglected skills, and McKibben, Wahesh, and Webber (2017b) highlighted a lack of research-based leadership training for counselors. A primary barrier to producing counseling leadership research is the lack of a valid and reliable measure built on a conceptual framework applicable to counselors. McKibben (2016) and McKibben, Umstead, and Borders (2017a) began to address the need for a conceptual framework by clarifying how counseling leadership works and by identifying dynamics of counseling leadership. However, the need for a counseling-specific leadership measure remains unmet. In this study, we developed and tested a self-report measure of counseling leadership behaviors as a way to spur needed research on counseling leadership skills, thereby bridging the leadership research-to-practice gap.

Counselors face a variety of challenges and opportunities that highlight the need for competent leadership, including professional identity, parity, and licensure portability (e.g., American Mental Health Counselors Association, 2015; National Board for Certified Counselors

[NBCC], 2015); social justice efforts and disaster relief needs (Paradise et al., 2010); increasing centrality of school counselors in students' educational experiences (Dollarhide, 2003); and advocacy efforts in areas such as the Veterans Administration, Medicare, and TRICARE. Building on Paradise et al.'s (2010) assertion that all counselors (e.g., students, practitioners, educators, administrators, supervisors) are capable of providing leadership, McKibben (2016) noted that the complexity of issues and opportunities facing counselors necessitates a process-oriented leadership approach that allows counselors to lead flexibly depending on the context one is in. This approach, drawn from the Integrative Process Model of Leadership (IPML; Eberly, Johnson, Hernandez, & Avolio, 2013) and dynamic systems theory (e.g., Michel & Moore, 1995), allows counselors to adapt to ever-changing professional needs by drawing on a variety of skills and by attuning to how one leads rather than simply what leadership is. To further describe this leadership approach, McKibben et al. (2017a) conducted a content analysis of counseling leadership literature and found that counseling leadership was comprised of three groups of 24 themes (see Table 1). These themes contained a variety of leadership behaviors, thoughts, feelings, values, and traits. Although this study concretely described what counseling leadership looks like and how counselors might use it to lead, the groups of leadership themes were defined by expert consensus and have not been tested statistically among counseling leaders. Thus, it remains unclear how accurately the three groups of themes map onto counselors' leadership experiences.

Paralleling the need for adaptable leadership, counselors have emphasized the importance of training counselors as leaders. For example, the Council for Accreditation of Counseling and Related Education Programs (CACREP, 2016) established leadership standards for doctoral students and for some master's-level specialty areas (e.g., school counseling, college counseling and student affairs). The Association for Counselor Education and Supervision (ACES) regularly offers an emerging leaders workshop at its conference. Chi Sigma Iota (CSI) also offers regular leadership training opportunities through CSI Days workshops (CSI, 2017a), a leadership fellows and interns program (CSI, 2017b), and online webinars (CSI, 2017c). CSI established the Principles and Practices of Leadership Excellence (PPLE; CSI Academy of Leaders, 1999) in an attempt to guide counselors' understanding of effective leadership. However, with a sample of 50 CSI student chapter leaders, McKibben et al. (2017b) found that student leaders in CSI engaged inconsistently in leadership behaviors that align with the PPLE. One of their conclusions was that evidence-based, skills-focused training was needed in CSI and across the counseling profession.

Although counselors have begun implementing needed leadership training to meet the needs of the profession, counselors lack more objective methods for measuring, evaluating, and training leadership dynamics that occur in the counseling profession. For example, without survey-based information, counseling leaders have limited (if any) access to valuable feedback on their leadership efforts. Organizational leadership researchers often utilize self- and other-report scales (i.e., multirater assessment) that allow leaders to receive 360-degree feedback (i.e., feedback from multiple sources; for a review, see Day, Fleenor, Atwater, Sturm, & McKee, 2014; Fleenor, Smither, Atwater, Braddy, & Sturm, 2010). Additionally, the absence of counseling-specific leadership measurement hinders leadership research by limiting researchers' ability to test counseling leadership strategies empirically and to link strategies with outcomes (e.g., goal

achievement, term cohesion). Relatedly, counselors do not have consistent quantitative strategies to evaluate leadership education and training programs that align with a cohesive research base.

**Table 1.** Counseling leadership themes by group.

| **Leadership values and qualities** | **Sample items on DLCS-SR** |
|---|---|
| Professional identity | "Promote a unique counselor identity through professional activity that advances the profession." |
| Advocacy | "Engage in social justice efforts." |
| Vision | "Clearly communicate a vision to followers." |
| Modeling | "Serve as a role model for others." |
| Mentorship | "Build supportive relationships with mentees." |
| Service | "Seek opportunities to serve the profession." |
| Dealing with difficulty and setbacks | "Address conflict openly and directly." |
| Leadership-specific cognitive complexity | N/A |
| High standards for self and others | "Invest effort into developing personal leadership abilities." |
| Passion | N/A |
| Sense of humor | "Use humor at appropriate times." |
| Creativity | "Use creative strategies to stimulate awareness." |
| Wellness | "Attend to own personal wellness." |
| Personal and interpersonal qualities | Sample items on DLCS-SR |
| Intrinsic motivation | N/A |
| Authenticity | "Behave in a manner that is true to myself." |
| Humility | "Give credit to others for success." |
| Intentionality | "Act intentionally or strategically." |
| Dependability | N/A |
| Leadership developmental influences | N/A |
| Openness | "Gather diverse perspectives and expectations from others." |
| Principled | "Act with integrity." |
| Interpersonal skills | Sample items on DLCS-SR |
| Interpersonal influence | "Inspire individuals to make change of their own accord." |
| Assertiveness | "Set boundaries and expectations with others." |
| Role competence | "Meet professional concerns of followers." |

*Note*. N/A in sample items list indicates theme not included on the Dynamic Leadership in Counseling Scale–Self-Report (DLCS-SR).

These barriers are reflected in the lack of outcome studies available on counseling leadership. In one study, Luke and Goodrich (2010) took a qualitative approach and found that leadership in CSI provided opportunities for professional identity development. In another, Mason (2010) used an externally developed leadership measure, the Leadership Practices Inventory (LPI; Kouzes & Posner, 2003), and found that leadership practices among school counselors were linked to successful school counseling program implementation. Mason's (2010) study reflects the type of research needed in counseling leadership, yet the results of this particular study were limited because the leadership behaviors measured by the LPI may not fully capture the range of counselor leadership behaviors. A counseling-specific leadership measure could serve as a catalyst to further research and to further leadership training and practice that is based on rigorous research within the profession.

Given the gaps in counseling leadership research underscored by a need for a valid and reliable measure, the purpose of this study was to develop the Dynamic Leadership in Counseling Scale–Self-Report (DLCS-SR) and to investigate for evidence of validity and reliability among a sample of counselors. In this study, we aimed to begin the process of counseling leadership measurement in hopes that the profession will evolve toward multirater assessment for leadership behavior and skills. Further, as a behavioral measure, the DLCS-SR's name embodies the flexibility of skills needed to meet constantly shifting leadership needs and goals in the counseling profession. In this study, we investigated the following research questions: (a) To what extent is there evidence of construct validity for the DLCS-SR? (b) What is the internal consistency among the subtests used to specify the factors of the DLCS-SR? (c) To what extent is there evidence of convergent validity for the DLCS-SR? (d) To what extent are participants responding in a socially desirable manner?

## Method

### *Instrument development process*

To develop the DLCS-SR, we followed guidelines from DeVellis (2017) and Lee and Lim (2008): (a) determine what to measure; (b) generate an item pool; (c) determine the measurement format; (d) submit to expert review; (e) include validation items; (f) administer items to a development sample and evaluate; and (g) optimize scale length. In *determining what to measure*, we designed the DLCS-SR to measure counseling leadership behaviors identified by McKibben et al. (2017a). In their study of counseling leadership dynamics, McKibben et al. (2017a) classified counseling leadership dynamics into three groups of 24 themes (see Table 1), and they also noted how the dynamics are observed (e.g., cognitions, emotions, behaviors, traits, values). Leadership behaviors are more objectively observable and trainable, which allows for greater triangulation and integration of information; thus, we designed the DLCS-SR to measure behaviors identified in McKibben et al. (2017a). Notably, 19 of the 24 themes from McKibben et al. (2017a) contained leadership behaviors. *Intrinsic motivation, leadership-specific cognitive complexity, passion*, and *dependability* were not described behaviorally and thus were not included on the DLCS-SR. In addition, *leadership developmental influences* was described behaviorally, but such behaviors occurred prior to becoming a counselor and thus could not be evaluated in the present by the DLCS-SR. The DLCS-SR was designed to measure behaviors among the three groups of 19 behaviorally specified leadership themes.

Next, we *generated an item pool* with at least three items per leadership theme to accurately test the factor structure (Kline, 2011). Seventy-five initial substantive items were generated across the 19 themes. We closely followed Kline's (2005) recommendations to deal with one thought at a time; be brief and precise; avoid awkward wording, irrelevant information, double negatives, all-or-none language, and indeterminate terms (e.g., "frequently"); and present items in positive language.

We then *determined the format for measurement* by adapting the Too Little/Too Much scale (TLTM scale; Kaiser & Kaplan, 2005) for the DLCS-SR. Items on this measure are scored bidirectionally from −4 (*much too little*) to +4 (*much too much*), with a score of 0 (*the right amount*) as a midpoint in the scale rating leadership behavior. Scores were computed by

calculating the absolute value of an item response (e.g., scores of −3 and +3 are both 3). Thus, lower absolute scores (closer to 0) reflected using a leadership behavior closer to the right amount, whereas higher absolute scores (farther from 0) reflected more lopsidedness (e.g., over- or underutilizing a behavior). We selected this response format because traditional Likert scales on leadership measures (e.g., 1–5 scales) may contain blind spots in assessing the extent to which a leader over- or underutilizes a given leadership approach (Kaiser & Kaplan, 2005). Hollenbeck, McCall, and Silzer (2006) noted that even leadership strengths can turn into weaknesses if used too often or not enough. A TLTM scale may clarify options for respondents, allow researchers to draw clearer inferences from data, and provide specific feedback to leaders on measurement scores.

We then submitted the initial item pool for *expert review*. The items, survey instructions, and TLTM scale were sent to two counselors with at least 10 years of counseling leadership experience with a request for feedback on wording and clarity. These counselors were selected based on their leadership experience as educators as well as prolific service in counseling organizations, publications in journals, and authorship of books (i.e., Black & Magnuson, 2005). Based on their feedback, several items were reworded to improve clarity, to remove double-barreled questions, and to more clearly link items to themes. In an effort to triangulate feedback on items, we also sought feedback from counselors who were newer to leadership efforts. To do this, we reached out to five doctoral students in a CACREP-accredited counselor education program who represented varying professional counseling backgrounds (one career counseling, one rehabilitation counseling, three clinical mental health counseling). These students were selected based on the early career status of their leadership efforts and their expressed interest in leadership. The students completed a sorting task by placing each item in the leadership theme to which they believed the item belonged. Thirteen of 75 items were reworded based on variability in sorting.

Finally, we included *validation items* to detect socially desirable responding. Leaders who are blind to their own weaknesses or who fall into more is better thinking may provide artificially inflated scores on a self-report measure, and inflated scores stemming from social desirability may increase risk of a Type I error (McKibben & Silvia, 2016). A social desirability measure (described below) was included at the end of the online survey to detect social desirability.

*Procedures*

There is no consistent a priori sample size recommendation for statistical testing in instrument development. Mvududu and Sink (2013) pointed to a common participant-to-variable ratio for factor analysis modeling, recommending that researchers seek about 10 participants for every one parameter estimated in a model. Crockett (2012) reviewed sample size determinants for structural equation modeling (SEM) and she noted that one approach, the critical *N* statistic (Hoelter, 1983), typically supports that at least 200 participants are needed for enough power. Crockett (2012) also noted that a 200-participant minimum in SEM appears to be a gold standard, which was echoed by Mvududu and Sink (2013) as a general recommendation for factor analysis. To maximize true score variance in the current study, we sought to recruit at least 200 participants across a diverse population of students, counselor educators, and practitioners. *Counseling student* was defined operationally as current enrollment as a master's

or doctoral student in a CACREP-accredited counselor education program. We sampled students in CACREP-accredited programs to ensure relative consistency in training. *Counselor educator* was defined as holding a PhD in counselor education and currently working in a counselor education program. *Practitioner* was defined as either fully licensed as a professional counselor or provisionally licensed and seeking full licensure under supervision.

On approval from the Institutional Review Board, we e-mailed counselor educators in CACREP-accredited programs and asked them to participate; we also asked them to share a link to the online study with their students. To further reach faculty not teaching in CACREP-accredited programs and practitioners, we consulted leadership directories of CSI, ACA (and each of its divisions), ACES (and each of its regions), and NBCC and contacted those whose e-mails were listed publicly to request participation. Finally, we employed snowball sampling by presenting participants with the participation criteria at the end of the survey and asking them to forward the study to colleagues who met participation criteria.

*Participants*

Of the 305 participants who began the study, 85 did not complete it (72% completion rate). Data from these 85 participants were removed prior to analysis because the DLCS-SR was not completed. Two additional participants were removed from the dataset prior to analysis because they indicated that they were students not enrolled in a CACREP-accredited program. This left data from 218 participants for analysis.

Of the 218 participants, 192 were Caucasian (88%), 11 were African American (5%), eight were Asian American (3.7%), five were American Indian or Native Alaskan (2.3%), one was Native Hawaiian or Pacific Islander (0.5%), eight preferred not to state their racial background (3.7%), and one did not respond to this item (0.5%). Nine participants were Hispanic or Latino/Latina (4.1%), 201 were not Hispanic or Latino/Latina (92.2%), six preferred not to state their ethnicity (2.8%), and two did not respond to this item (0.9%). Participants were allowed to choose more than one racial/ethnic category and to decline to select any. Fifty-five participants identified themselves as male (25.2%), and 163 identified themselves as female (74.8%). Participants ranged in age from 22 to 73 years ($M = 37$; $SD = 11.98$); 15 participants (6.9%) did not indicate their age. There were 85 students enrolled in CACREP-accredited counselor education programs (40%), 69 counselor educators (31.7%), 57 counseling practitioners (26.1%), and seven other (3.2%).

Among the 85 counseling students, 56 were pursuing a master's degree (65.9%), one an educational specialist degree (1.2%), and 28 a doctoral degree (32.9%). Credit hour completion ranged from 0 (first semester) to 130 hours ($M = 38.43$; $SD = 27.16$); four did not indicate. All students were currently enrolled in a CACREP-accredited counseling program. Twelve students were fully licensed as a counselor in their state (14.1%), 15 were provisionally licensed (17.6%), and 58 indicated that this was nonapplicable (68.2%). Current specialties were as follows: 30 clinical mental health counseling (35.3%); seven marriage, couple, and family counseling (8.2%); 17 school counseling (20%); three student affairs and college counseling (3.5%); 24 counselor education (28.2%); and four other (4.7%). Those who indicated "other" reported concentrations in community health, play therapy, research, and dual mental health/school.

Among the 69 counselor educators, there were 24 assistant professors (34.7%), 18 associate professors (26.1%), 18 full professors (26.1%), 12 tenure-track faculty (17.4%), four non-tenure-track faculty (e.g., clinical professor; 5.8%), two visiting professors (2.9%), four adjunct professors (5.8%), and three other (4.3%). Participants were allowed to select more than one option. Those who indicated "other" reported educator roles such as department chair, doctoral candidate, and tenured. Years of experience as a counselor educator ranged from one to 46 years ($M$ = 10.41; $SD$ = 9.80). Fifty-five taught in a CACREP-accredited counseling program (79.7%) and 14 (20.3%) did not teach in a CACREP-accredited counseling program. Fifty-nine graduated from a CACREP-accredited counseling program (85.5%) and 10 did not graduate from a CACREP-accredited counseling program (14.5%). They identified with the following specialties: four career counseling (5.8%); 39 clinical mental health counseling (56.5%); two marriage, couple, and family counseling (2.9%); 18 school counseling (26.1%); three addictions counseling (4.3%); two rehabilitation counseling (2.9%); and one other (generalist/school counseling; 1.4%). Fifty-two were fully licensed as a counselor in their state (75.4%), nine were provisionally licensed (13%), and eight did not respond to this item (11.6%).

Among the 57 practitioner participants, 36 were fully licensed as a counselor in their state (63.2%), 16 were provisionally licensed (28.1%), and five did not respond to this item (8.8%). Years of experience as a practitioner ranged from zero (first year) to 31 years ($M$ = 8.48; $SD$ = 7.83). They identified with the following specialties: one career counseling (1.8%); 29 clinical mental health counseling (50.9%); 10 marriage, couple, and family counseling (17.5%); 11 school counseling (19.3%); one student affairs and college counseling (1.8%); two addictions counseling (3.5%); and three other (5.3%). Those who indicated "other" reported holding multiple professional affiliations (e.g., school and professional counselor, counselor/marriage and family therapist). Forty-four practitioners had master's degrees (77.2%), five had an educational specialist degree (8.8%), and seven had a doctoral degree (12.3%). Forty-seven practitioners graduated from a CACREP-accredited counseling program (82.5%), and nine did not graduate from a CACREP-accredited counseling program (15.8%).

Among the seven "other" participants, four were fully licensed as a counselor in their state (57.1%), two were provisionally licensed and pursuing full licensure under supervision (28.6%), and one indicated that this was nonapplicable (14.3%). Years of experience in their current role ranged from zero to 12 years ($M$ = 4, $SD$ = 3.59). They identified with the following specialties: two clinical mental health counseling (28.6%); two marriage, couple, and family counseling (28.6%); one school counseling (14.3%); one student affairs and college counseling (14.3%); and one addictions counseling (14.3%).

*Instrumentation*

DLCS-SR

The DLCS-SR contained 75 items of counseling leadership behaviors (see Table 1 for sample items) scored on an adapted TLTM scale from −4 (*much too little*) to +4 (*much too much*) with 0 (*the right amount*) as a midpoint in the scale. The measure was scored by taking the absolute value for each item response and calculating a mean score for items on each leadership theme,

yielding 19 observed variables. Scores closer to 0 reflect engaging in leadership behaviors closer to the right amount. Each item also contained an N/A option, which participants could select if they felt any given leadership behavior did not apply to them.

Global transformational leadership scale (GTL; Carless, Wearing, & Mann, 2000)

The GTL, included as a test for convergent validity, is a 7-item measure of transformational leadership scored on a Likert scale ranging from 1 (*rarely or never*) to 5 (*very frequently, if not always*). The seven items measure a global construct of transformational leadership along the following components: vision, staff development, supportive leadership, empowerment, innovative thinking, lead by example, and charisma. Higher GTL scores reflect higher levels of transformational leadership skills. Carless et al. (2000) reported a Cronbach's alpha ($\alpha$) of .93 in their initial validation of the GTL. In the current study, $\alpha$ was .84.

Balanced inventory of desirable responding–short form (BIDR-SF; Steenkamp, de Jong, & Baumgartner, 2010)

To detect socially desirable responding, we included the BIDR-SF, a 20-item measure scored on a Likert scale ranging from 1 (*not true*) to 7 (*very true*). The BIDR-SF contains a 10-item impression management scale that detects conscious intentional attempts to present one's self favorably (e.g., "I have said something bad about a friend behind his or her back") and a 10-item self-enhancement scale that detects unconscious overly positive self-views that are projected onto a survey (e.g., "My first impressions of people usually turn out to be right"). Higher scale and composite scores reflect higher social desirability. In a multinational study with over 12,000 participants, Steenkamp et al. (2010) reported $\alpha$s for the self-enhancement and impression management scales of .67 and .73, respectively. In the current study, self-enhancement and impression management $\alpha$s were .68 and .80, respectively.

*Data analysis*

We first examined the dataset for any incidence of missing data and found no incidences of missing data among the 218 participants retained in the sample. We also screened for how often participants selected N/A on the scale to determine if any items might not be applicable to counseling leaders. Collectively, counseling students selected N/A 372 times throughout the survey, making them more likely to select N/A than counselor educators (31 times), practitioners (97 times), and others (1 time). Students, counselor educators, and practitioners all selected N/A more often for items measuring role competence compared to their respective N/A responses to other items. Students and practitioners tended to select N/A more often for items measuring vision and mentorship. An item measuring advocacy, "Shape the intellectual capital that advances the counseling profession in counseling journals by reviewing manuscripts," was marked N/A 28 times by students and 20 times by practitioners, but only two times by counselor educators.

Next, we checked assumptions of normality via item skew (>3.00) and kurtosis (>10.00), and item-total correlations, which also allowed us to determine if any statistically weak items should be dropped prior to testing the factor structure. Item-total correlations below .2 were flagged for

removal (Everit, 2002). Item means were around 1.0 with standard deviations less than 2.0, indicating good variability around the mean. No items demonstrated high skew (>3.00) or kurtosis (>10.00), and no item-total correlations were below the .2 cutoff for removal. Thus, all items were retained for the exploratory factor analysis (EFA). We aggregated item mean scores together for each leadership theme to yield 19 composite observed variables.

For research question one, we used EFA in SPSS (IBM Corporation, 2016) to test the underlying structure for the DLCS-SR. Assumptions of normality and sampling adequacy held (KMO coefficient = .934, significant Bartlett's test of sphericity [$\chi^2 = 1619.92$, $df = 171$, $p = .000$]), indicating the extracted model was an acceptable fit to the data. However, the correlation matrix determinant was less than .0001, indicating possibly high multicollinearity among extracted factors. Because the data appeared normally distributed, we used a maximum likelihood (ML) EFA approach given the range of fit indices available (Fabrigar, Wegener, MacCallum, & Strahan, 1999). Given the probable multicollinearity among the extracted factors, we used promax rotation to further interpret the EFA. As an oblique rotation method, promax rotation allows for factor interpretation when latent factors are correlated (Fabrigar et al., 1999). To evaluate the factor structure of the DLCS-SR with the current sample, we used the following criteria: (a) a priori hypotheses that DLCS-SR might consist of one or three factors; (b) eigenvalues greater than one; (c) scree test (e.g., Rencher & Christensen, 2012); (d) minimum factor coefficient of .4 for each observed variable (Guadagnoli & Velicer, 1988); and (e) examination of factor cross-loadings.

For research question two, we examined internal consistency with $\alpha$ coefficients: .70 – .80 were considered acceptable, .80–.90 were considered very good, and above .90 were considered excellent (DeVellis, 2017). For research question three, we tested for evidence of convergent validity by correlating the DLCS-SR with the GTL. DLCS-SR scores closer to zero should correlate with higher GTL scores (i.e., negative, significant correlations). For research question four, we tested for social desirability by correlating the DLCS-SR with the two scales on the BIDR-SF.

## Results

### *Construct validity*

The EFA extracted three factors with an eigenvalue greater than one that accounted for 53.01% of the variance (see Table 2). The first factor accounted for the most variance (41.68%), and the other two factors each accounted for about 5% of additional variance. Most of the 19 observed variables loaded moderately (>.4) to strongly (>.7) onto one of the extracted factors (see pattern matrix, Table 2). Observed variables for dealing with difficulties and setbacks, humor, creativity, and intentionality did not load onto a factor above the .4 cut score. Interpersonal influence cross-loaded nearly equally onto two factors.

In addition to the second and third factors not accounting for much variance, the three extracted factors were not clearly defined as most of the observed variables correlated moderately to strongly with all three factors (see structure matrix, Table 2). All three extracted factors also correlated with one another strongly (>.6, see Table 2). These correlations among observed

variables and factors, and among the factors themselves, are likely the source of multicollinearity detected when testing EFA assumptions. The scree plot (see Figure 1) drops significantly between the first and second factors, which Rencher and Christensen (2012) refer to as the elbow rule and recommend not interpreting factors after the drop off. Collectively, this evidence suggested that a single-factor model may provide a more parsimonious fit to the data with the current sample compared to a three-factor model.

**Table 2.** Exploratory factor analysis (EFA) factor matrices.

| | **Pattern matrix** | | | **Structure matrix** | | |
|---|---|---|---|---|---|---|
| | **Factor** | | | **Factor** | | |
| | **1** | **2** | **3** | **1** | **2** | **3** |
| Professional identity | .022 | .163 | **.508** | .476 | .505 | .627 |
| Advocacy | .132 | .087 | **.457** | .494 | .481 | .598 |
| Vision | **.708** | −.122 | .159 | .718 | .524 | .539 |
| Modeling | .104 | **.584** | .048 | .584 | .695 | .489 |
| Mentorship | −.403 | **.954** | .118 | .406 | .720 | .468 |
| Service | −.031 | −.053 | **.830** | .465 | .454 | .776 |
| Difficulties/setbacks | .285 | .361 | −.061 | .523 | .541 | .354 |
| Humor | .120 | .271 | .048 | .359 | .394 | .299 |
| Creativity | .384 | .122 | .154 | .577 | .515 | .480 |
| High standards | **.422** | .113 | .214 | .647 | .574 | .559 |
| Wellness | **.635** | −.092 | −.064 | .523 | .354 | .287 |
| Authenticity | **.429** | .140 | −.035 | .514 | .447 | .332 |
| Humility | .308 | **.409** | .034 | .644 | .667 | .495 |
| Intentionality | .263 | .346 | .074 | .576 | .594 | .465 |
| Openness | .347 | **.442** | −.072 | .640 | .662 | .435 |
| Principled | .293 | **.442** | −.123 | .553 | .589 | .349 |
| Interpersonal influence | **.507** | **.516** | −.107 | .834 | .837 | .551 |
| Role competence | .313 | **.407** | .198 | .754 | .774 | .661 |
| Assertiveness | **.672** | −.123 | .091 | .637 | .451 | .447 |
| Eigenvalues | 7.919 | 1.137 | 1.016 | | | |
| % of Variance | 41.680 | 5.986 | 5.347 | | | |
| Cumulative % | 41.680 | 47.666 | 53.013 | | | |
| Factor 1 | | | | 1 | .768 | .647 |
| Factor 2 | | | | | 1 | .640 |
| Factor 3 | | | | | | 1 |

*Note*. Pattern matrix reflects rotated factor loadings. Boldface indicates factor loadings above .4 cut score. Structure matrix reflects correlations between the observed variables and each of the three extracted factors. Both matrices were obtained from the promax rotated solution.

To test the fit of a single-factor model more directly, we loaded the 19 observed variables from the same dataset onto a single factor (i.e., global leadership) and tested the model using confirmatory factor analysis (CFA) in Mplus (Muthén & Muthén, 1998–2015). Whereas EFA attempts to extract an underlying factor structure from data, CFA tests a specified model's fit to the data. Evidence from the EFA suggested a single factor might best fit the data with our sample, and the CFA allowed us to observe this more directly by testing only a single-factor model. The global fit indices indicated that the single-factor model was a good fit to the data with this sample (i.e., root mean square error of approximation = .056, 90% CI; comparative fit

index = .927; standardized root mean square residual = .05). All observed variables, with the exception of humor, loaded onto the single factor above the .4 cut score (see Figure 2), indicating that the observed variables likely better specified a single-factor model with the current sample.

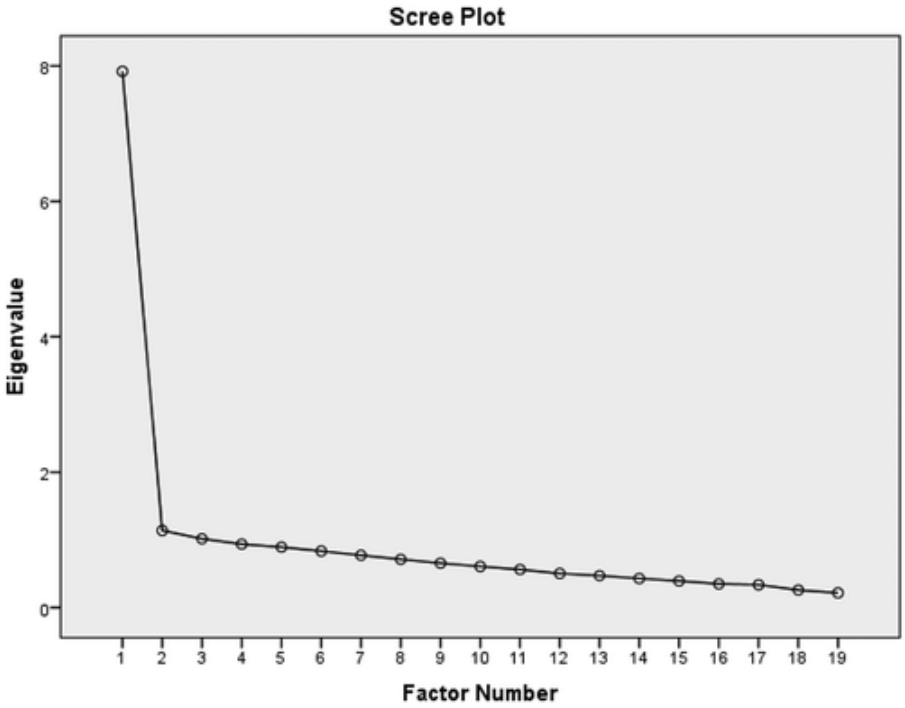


**Figure 1.** Exploratory factor analysis (EFA) scree plot.

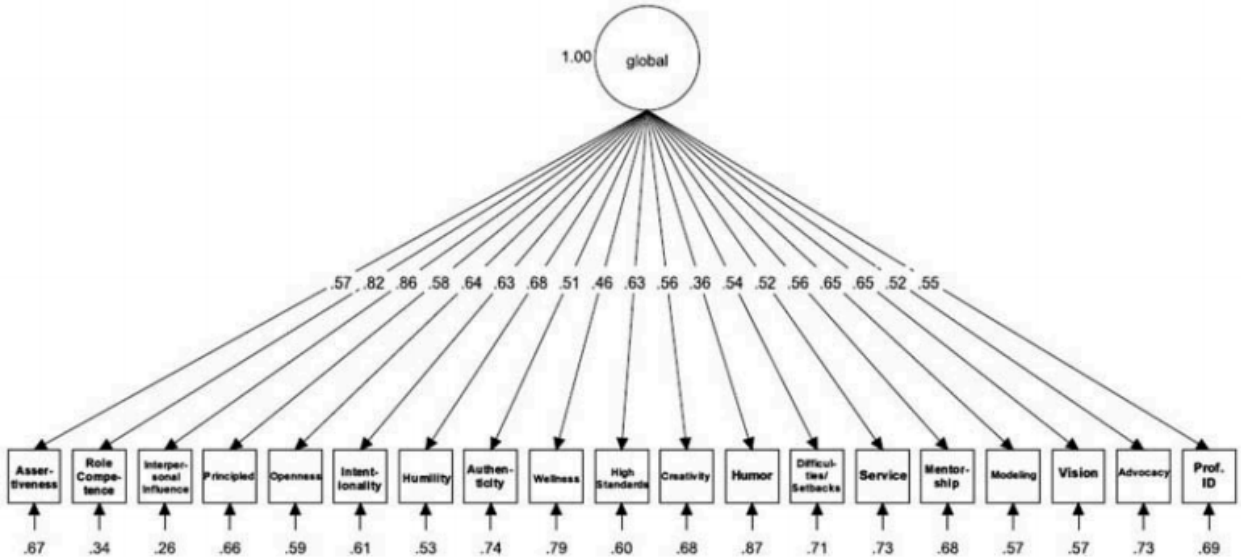


**Figure 2.** Single-factor confirmatory factor analysis (CFA) model. Factor loadings shown represent standardized factor loadings. All factor loadings significant at $p < .0001$.

*Convergent validity and internal consistency*

Given that we failed to reject a single-factor model fit during EFA and CFA, we tested our remaining hypotheses with the DLCS-SR as a single-factor instrument. The DLCS-SR correlated significantly with the GTL in the hypothesized direction ($r = -.562$, $p < .001$), providing evidence for convergent validity. Internal consistency among the subtests of all items on the DLCS-SR yielded an $\alpha$ of .942, indicating excellent internal consistency across the measure as a whole.

*Social desirability*

The DLCS-SR correlated significantly with BIDR-SF self-enhancement ($r = -.218$, $p < .001$) and impression management ($r = -.134$, $p < .05$) scales, indicating that DLCS-SR scores may have been influenced by socially desirable responding. Notably, although the correlations were statistically significant, the effect sizes for self-enhancement and impression management were small ($r^2 = .05$ and .02, respectively).

**Discussion**

Using behavioral indicators from McKibben et al.'s (2017a) comprehensive description of counseling leadership, we developed and tested a self-report measure of counseling leadership. In contrast to McKibben et al. (2017a), we found that the 19 behavioral leadership themes among three conceptual categories were not statistically distinct for the current sample. Rather, we found evidence that counseling leadership behaviors as measured by the DLCS-SR may best be explained by a single factor. The single-factor model demonstrated excellent internal consistency and was significantly correlated to a measure of global transformational leadership. Thus, counseling leadership behaviors on the DLCS-SR may better be conceptualized as a global skill set rather than distinct entities.

A single-factor model aligns with Eberly et al.'s (2013) and McKibben's (2016) assertions that leadership behaviors make up a nonlinear systemic interaction among people. Rather than containing distinguishable components, counseling leadership behaviors may be important ingredients in a broader social interaction. This idea is further underscored by dynamic systems theory, which states broadly that elements of any system (e.g., leadership) emerge nonlinearly based on a web of interacting components around the elements. In other words, a potential reason why a single leadership factor was found in this study could be because leadership behaviors occur rapidly in nuanced ways depending on specific environmental causes. Because behaviors occur rapidly and are contextually expressed, they may be difficult (if not impossible) to classify neatly into categories.

The current finding of a single-factor model over a three-factor model also highlight ongoing dimensionality issues in leadership measurement. Primarily, leadership instrument factor structures may look different depending on who is being evaluated and by whom. For example, Neider and Schriesheim (2011) found differing factor structures on the Authentic Leadership Inventory when participants rated U.S. Senator John McCain compared to when participants rated either U.S. President Barak Obama or their current work supervisor. Similarly, the factor

structure of the popular Multifactor Leadership Questionnaire has been both supported and refuted since its development (e.g., Tejeda, Scandura, & Pillai, 2001). These mixed findings across disciplines might also be explained by the nuanced and ever-changing ways in which leaders behave based on the situation they are in.

In this study, we asked participants to self-evaluate behaviors, and this yielded a single-factor model. The presence of social desirability may have produced scores artificially closer to 0 (*the right amount*), which could have influenced the factor structure. It is possible that asking participants to evaluate someone else could yield an alternate factor structure and allow researchers to observe statistically distinct leadership factors. Leadership researchers and theorists have not been able to determine whether dimensionality issues reflect limitations in measurement, clarity/specificity of theory, or both, but more research is needed for counseling leadership in particular.

When observed variables were allowed to load onto extracted factors in the EFA, dealing with difficulties and setbacks, humor, creativity, and intentionality did not load strongly onto a factor. However, only humor continued to perform poorly when observed variables were constrained to a single factor during CFA. Although humor has been noted as useful if used intentionally (Haight & Shaughnessy, 2006), we did not find support for humor being a behavioral descriptor of counseling leadership in this study and the humor items could be deleted from the DLCS-SR. Broadly speaking, humor has not been regularly discussed in leadership theories across disciplines, and may best be considered a unique relational approach in leadership. If used at the wrong time or in the wrong way, it is possible for humor to backfire and derail leadership efforts. Although we did not find evidence for humor at the behavioral level, it may be relevant to counseling leadership as a trait or value (McKibben et al., 2017a).

*Limitations*

There are limitations worth noting with this study. First, the DLCS-SR was designed to measure counseling leadership behaviors. The behavioral analysis of counseling leadership does not assess the full range of possible leadership dynamics (e.g., thoughts, feelings, values, traits). Relatedly, there may be statistically observable distinctions of leadership at a broader scope of measurement that was not observable with this behavioral measure. Another limitation was the interaction of participants with the TLTM scale. This scale is different from traditional Likert scales, which potentially could have influenced survey responses. We attempted to minimize confusion prior to administering the DLCS-SR by clarifying the instructions (e.g., specific wording, including item examples) and revising items to be more specific after expert review.

Sampling limitations are also noteworthy. The demographic representation of counselor educators, students, practitioners, and others in our sample may not proportionally represent the population of counseling leaders. Also, our efforts to sample difficult-to-reach groups (i.e., practitioners) via snowball sampling also impacted our sampling rigor and thus may not have yielded a completely representative sample of counseling leaders. We attempted to gain as much demographic information about our sample as possible to minimize this limitation. Finally, our sample identified mostly as White (88%) and female (75%), meaning that generalizing results of this study beyond these intersecting social identities should be done tentatively.

Our analysis of leadership among counselor educators, students, and practitioners contains an assumption that leadership experiences are relatively homogeneous across groups. During item analysis, we found that students were more likely to select N/A for leadership behaviors compared to any other participant group. Researchers should be mindful in administering the DLCS-SR to and interpreting results from students because the items on the instrument may not be as readily applicable to students as to other counseling leaders. Finally, the possibility that some participants were responding in a socially desirable way may have introduced error variance into the results. Though we were able to detect it, we were not able to remove it. Although social desirability was significantly correlated with the DLCS-SR in this study, the low effect sizes indicated that social desirability likely did not account for much variance in scores on the DLCS-SR. Nevertheless, additional research with the DLCS-SR or other leadership self-report measures in counseling should continue to control for this potential source of error.

*Implications for research and practice*

Despite the noted limitations, the DLCS-SR is a promising research tool to begin measuring counseling leadership behaviors, and the TLTM scale promotes useful concrete feedback for counseling leaders. The initial tests for validity and internal consistency in this study highlight that additional work is needed to clarify which items provide useful information and which items can be removed to potentially shorten the scale. Differential item functioning (DIF) analysis may provide researchers with a better understanding of which items match participants' ability and most closely capture what is being assessed. Similarly, more research is needed with larger more diverse samples to further test the factor structure of the DLCS-SR and to clarify our understanding of generalizability of leadership behaviors across cultures and social identities.

Future researchers might triangulate data from multiple sources (e.g., leader, collaborators) by developing an other-report version of the DLCS-SR. Multirater assessments are common in leadership research, and this approach may allow researchers to draw more valid and reliable conclusions from data (Conway & Huffcut, 1997). Multirater assessment also may provide more robust feedback to counseling leaders. The social desirability detected in this study may be indicative of leaders' inability to see, or resistance to admitting, their over- orunderuse of leadership behaviors, and an other-report measure may allow researchers to see beyond this blind spot in self-reporting by triangulating information for multiple sources.

With the DLCS-SR providing opportunities for more detailed investigations into counseling leadership, important next steps for researchers include descriptive cross-sectional designs to investigate leadership behaviors among various groups of counselors (e.g., students, educators, practitioners), developmental phases (e.g., McKibben et al., 2017b), and leadership settings and contexts. Similarly, longitudinal research is needed to investigate changes in leadership behaviors over time. This approach may be useful to better understand how students learn and implement behaviors developmentally. Finally, researchers might employ qualitative research that builds off the behaviors on the DLCS-SR to investigate how leadership behaviors are utilized in various contexts, which could shed light on behaviors that are more applicable across groups of counselors and across contexts. Similar to multirater assessment, these research

approaches may allow researchers to triangulate sources of information on counseling leadership behaviors, thus gaining a clearer understanding of how counseling leadership works.

The DLCS-SR also provides opportunities for leadership researchers to engage in needed outcome research. Researchers across disciplines have investigated leadership outcomes, including follower commitment, employee performance and satisfaction, leader effectiveness and performance, motivation, and group performance (for a review, see Bass & Riggio, 2006; Hiller, DeChurch, Murase, & Doty, 2011). However, counseling leadership outcome research remains an underexplored frontier. Mason's (2010) study linking school counselor leadership practices to program implementation is an exemplary model of needed outcome research. The DLCS-SR allows for investigation of outcomes, particularly using the TLTM scale to observe how under- or overuse of behaviors impacts how desirable leadership outcomes occur.

This study was a first step in a needed, though complex, process of measuring counseling leadership behaviors. Because the DLCS-SR is new and more research is needed to ascertain its utility, we caution against using the DLCS-SR in high-stakes testing or decision making until a more solid research base is available. Nevertheless, there are important implications for practice. The DLCS-SR offers concrete and tangible leadership skills that can be taught and trained, which has been repeatedly noted as a need to advance leadership practice in counseling (McKibben et al., 2017b; Paradise et al., 2010; Wahesh & Myers, 2014). Further, CACREP standards (CACREP, 2016) detail leadership knowledge and skills in doctoral- and certain master's-level specialty areas. Counselor educators working to meet these standards now have a potentially useful instructional tool that concretely identifies leadership behaviors and promotes conversation about how students can lead based on the context of their leadership efforts. Because students were more likely to mark N/A on DLCS-SR compared to other participants, the DLCS-SR may serve as a useful tool for students in discussing what counseling leadership looks like for them and how leadership might evolve over time. Counselors might use the DLCS-SR to self-assess their leadership behaviors, particularly where they are under- or overused. Self-assessment may serve as a catalyst for self-reflection and feedback on how to improve their leadership skills, a practice consistent with CSI's PPLE (CSI Academy of Leaders, 1999). Finally, the DLCS-SR may add vitality to leadership training and consultation efforts by providing opportunities for concrete feedback.

## Acknowledgments

W. Bradley McKibben is now at the Department of Counseling, Nova Southeastern University. Terry A. Ackerman is now at American College Testing (ACT). The authors thank Kelly L. Wester, Department of Counseling and Educational Development, and Paul Silvia, Department of Psychology, University of North Carolina at Greensboro, for their contributions to this study.

## References

American Mental Health Counselors Association. (2015). *AMHCA, ACES and NBCC jointly endorse plan for licensure portability*. Retrieved from http://www.amhca.org/news/244702/AMHCA-ACES-and-NBCC-Jointly-Endorse-Plan-for-Licensure-Portability.htm

Bass, B. M., & Riggio, R. E. (2006). *Transformational leadership* (2nd ed.). Mahwah, NJ: Erlbaum.

Black, L. L., & Magnuson, S. (2005). Women of spirit: Leaders in the counseling profession. *Journal of Counseling & Development*, 83, 337–342. doi:10.1002/j.1556-6678.2005.tb00352.x

Carless, S. A., Wearing, A. J., & Mann, L. (2000). A short measure of transformational leadership. *Journal of Business and Psychology*, 14, 389–405. doi:10.1023/A:1022991115523

Chi Sigma Iota. (2017a). *CSI Days*. Retrieved from http://www.csi-net.org/?page=CSI_Day

Chi Sigma Iota. (2017b). *Leadership fellowship and internship program*. Retrieved from http://www.csi-net.org/?page=Awards_LFI

Chi Sigma Iota. (2017c). *Webinars*. Retrieved from http://www.csi-net.org/?page=Webinars

Chi Sigma Iota Academy of Leaders. (1999). *Principles and practices of leadership excellence*. Greensboro, NC: Author. Retrieved from http://www.csi-net.org/?page=Leadership_Practices

Conway, J. M., & Huffcutt, A. I. (1997). Psychometric properties of multisource performance ratings: A meta-analysis of subordinate, supervisor, peer, and self ratings. *Human Performance*, 10, 331–360. doi:10.1207/s15327043hup1004_2

Council for Accreditation of Counseling & Related Educational Programs. (2016). *2016 CACREP standards*. Retrieved from http://www.cacrep.org/wp-content/uploads/2015/05/2016-CACREP-Standards.pdf

Crockett, S. A. (2012). A five-step guide to conducting SEM analysis in counseling research. *Counseling Outcome Research & Evaluation*, 3, 20–47. doi:10.1177/2150137811434142

Day, D. V., Fleenor, J. W., Atwater, L. E., Sturm, R. E., & McKee, R. A. (2014). Advances in leader and leadership development: A review of 25 years of research and theory. *Leadership Quarterly*, 25, 63–82. doi:10.1016/j.leaqua.2013.11.004

DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). Thousand Oaks, CA: Sage.

Dollarhide, C. T. (2003). School counselors as program leaders: Applying leadership contexts to school counseling. *Professional School Counseling*, 6, 304–308.

Eberly, M. B., Johnson, M. D., Hernandez, M., & Avolio, B. J. (2013). An integrative process model of leadership: Examining loci, mechanisms, and event cycles. *American Psychologist*, 68, 427–443. doi:10.1037/a0032244

Everit, B. S. (2002). *Cambridge dictionary of statistics* (2nd ed.). New York, NY: Cambridge University Press.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4, 272–299. doi:10.1037/1082-989X.4.3.272

Fleenor, J. W., Smither, J. W., Atwater, L. E., Braddy, P. W., & Sturm, R. E. (2010). Self-other rating agreement in leadership: A review. *Leadership Quarterly*, 21, 1005–1034. doi:10.1016/j.leaqua.2010.10.006

Guadagnoli, E., & Velicer, W. F. (1988). Relation to sample size to the stability of component patterns. *Psychological Bulletin*, 103, 265–275. doi:10.1037/0033-2909.103.2.265

Haight, M. G., & Shaughnessy, M. F. (2006). An interview with Samuel T. Gladding: Thoughts on becoming a counselor. *Journal of Counseling & Development*, 84, 114–119. doi:10.1002/j.1556-6678.2006.tb00385.x

Hiller, N. J., DeChurch, L. A., Murase, T., & Doty, D. (2011). Searching for outcomes of leadership: A 25-year review. *Journal of Management.*, 37, 1137–1177. doi:10.1177/0149206310393520

Hoelter, J. W. (1983). The analysis of covariance structures: Goodness-of-fit indices. *Sociological Methods & Research*, 11, 325–344. doi:10.1177/0049124183011003003

Hollenbeck, G. P., McCall, M. W., & Silzer, R. F. (2006). Theoretical and practitioner letters: Leadership competency models. *Leadership Quarterly*, 17, 398–413. doi:10.1016/j.leaqua.2006.04.003

IBM Corporation. (2016). SPSS statistics for Windows (Version 24.0) [Computer software]. Armonk, NY: Author.

Kaiser, R. B., & Kaplan, R. E. (2005). Overlooking overkill? Beyond the 1-to-5 rating scale. *Human Resources Planning*, 28, 7–11.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford.

Kline, T. J. B. (2005). *Psychological testing: A practical approach to design and evaluation*. Thousand Oaks, CA: Sage.

Kouzes, J. M., & Posner, B. Z. (2003). *The leadership practices inventory: Self instrument* (3rd ed.). San Francisco, CA: Jossey-Bass.

Lee, D., & Lim, H. (2008). Scale construction. In P. P. Heppner, B. E. Wampold, & D. M. Kivlighan (Eds.), *Research design in counseling* (3rd ed., pp. 494–510). Belmont, CA: Brooks/Cole.

Luke, M., & Goodrich, K. M. (2010). Chi Sigma Iota leadership and professional identity development in early career counselors. *Counselor Education & Supervision*, 50, 56–78. doi:10.1002/j.1556-6978.2010.tb00108.x

Mason, E. (2010). Leadership practices of school counselors and counseling program implementation. *NASSP Bulletin*, 94, 274–285. doi:10.1177/0192636510395012

McKibben, W. B. (2016). The content and process of counseling leadership: Implications for research and practice. *Journal of Counselor Leadership & Advocacy*, 3, 147–157. doi:10.1080/2326716X.2016.1147396

McKibben, W. B., & Silvia, P. J. (2016). Inattentive and socially desirable responding: Addressing subtle threats to validity in quantitative counseling research. *Counseling Outcome Research & Evaluation*, 7, 53–64. doi:10.1177/2150137815613135

McKibben, W. B., Umstead, L. K., & Borders, L. D. (2017a). Identifying dynamics of counseling leadership: A content analysis study. *Journal of Counseling & Development*, 95, 192–202. doi:10.1002/jcad.12131

McKibben, W. B., Webber, W. B., & Wahesh, E. (2017b). Exploring CSI chapter leaders' development toward leadership excellence. *Journal of Counselor Leadership & Advocacy*, 4, 52–65. doi:10.1080/2326716X.2017.1282332

Michel, G. F., & Moore, C. L. (1995). *Developmental psychobiology: An interdisciplinary science*. Cambridge, MA: MIT Press.

Muthén, L. K., & Muthén, B. O. (1998-2015). *Mplus user's guide* (7th ed.). Los Angeles, CA: Author.

Mvududu, N. H., & Sink, C. A. (2013). Factor analysis in counseling research and practice. *Counseling Outcome Research & Evaluation*, 4, 75–98. doi:10.1177/2150137813494766

National Board for Certified Counselors. (2015). *AMHCA-ACES-NBCC portability standards for counselors*. Retrieved from http://www.nbcc.org/assets/eblast/AMHCA-ACES-NBCC_Portability_Plan.pdf

Neider, L. L., & Schriesheim, C. A. (2011). The Authentic Leadership Inventory (ALI): Development and empirical tests. *Leadership Quarterly*, 22, 1146–1164. doi:10.1016/j.leaqua.2011.09.008

Paradise, L. V., Ceballos, P. T., & Hall, S. (2010). Leadership and leader behavior in counseling: Neglected skills. *International Journal for the Advancement of Counselling*, 32, 46–55. doi:10.1007/s10447-009-9088-y

Rencher, A. C., & Christensen, W. F. (2012). *Methods of multivariate analysis* (3rd ed.). Hoboken, NJ: Wiley.

Steenkamp, J. E. M., de Jong, M. G., & Baumgartner, H. (2010). Socially desirable response tendencies in survey research. *Journal of Marketing Research*, 47, 199–214. doi:10.1509/jmkr.47.2.199

Tejeda, M. J., Scandura, T. A., & Pillai, R. (2001). The MLQ revisited: Psychometric properties and recommendations. *Leadership Quarterly*, 12, 31–52. doi:10.1016/S1048-9843(01)00063-7

Wahesh, E., & Myers, J. E. (2014). Principles and practices of leadership excellence: CSI chapter presidents' experience, perceived competence, and rankings of importance. *Journal of Counselor Leadership & Advocacy*, 1, 83–97. doi:10.1080/2326716X.2014.886977