

Lessons that Last: LeaderShape-related Gains in Student Leadership Capacity over Time

David M. Rosch

Assistant Professor
Agricultural Education Program
University of Illinois at Urbana-Champaign

Clinton M. Stephens

Lecturer for Leadership Education
Leadership Studies Program
Iowa State University

Jasmine D. Collins

Doctoral Research Assistant
College of Education
University of Illinois at Urbana-Champaign

Abstract

The LeaderShape Institute is a popular immersion-based leadership program that is hosted on dozens of university campuses and conducted nationally each year. As part of a comprehensive research effort, a sample of 1,279 students at 21 participating institutions completed a pre-test prior to participating, as well as a post-test immediately after and a follow-up test three to four months later. Tests included measures of leadership skills, leadership-oriented self-efficacy, motivation to lead, and motivation to advocate for social issues. Results suggest students make gains in skill, confidence, and motivation to advocate for social issues, but that not all gains emerged similarly across social identity groups. Several differences emerged when comparing gains measured from pre-test to post-test and gains that included follow-up tests. This study provides support for the effectiveness of the LeaderShape Institute, and possesses several implications for the methods used in assessing the development of leadership capacity.

Introduction

An estimated \$60 Billion (U.S.) is spent annually just in the United States on leadership development programs (Zenger, 2012) and is growing (O'Leonard, 2014). Similar increases in the number of programs and students they serve have arisen on college campuses (Owen, 2012). Meanwhile, confidence in current and future leaders in the United States has never been lower (Rosenthal, 2012), while others draw similar conclusions on a global scale (Kellerman, 2012). The perceived need for high-quality leadership development programs in higher education, which exists in part to train future societal leaders (Boyer, 1987), has never been higher. However, assessment efforts to measure and validate the success of these programs have

curiously lagged behind (Dugan, 2011). While efforts have been made to assess the comprehensive nature of leadership education (see Brungardt, Greenleaf, Brungardt, & Arensdorf, 2006; Owen, 2012), these efforts have been focused more on describing services rather than assessing outcomes. Moreover, those studies that do emphasize outcomes assessment of educational interventions tend to focus on single-instance or single-institution gains (e.g. Rosch & Caza, 2011; Priest & Donley, 2014). We designed this research to assess the long-term effects of participating in a popular international co-curricular leadership development program for college students – the LeaderShape Institute.

National Efforts to Assess the Effectiveness of Leadership Education. Prior to a decade ago, most national research efforts focused on student leadership development—while helpful in showing the benefits of participation in leadership-oriented activities—did not possess strong theoretical foundations in how leadership was conceptualized for those who participated within these studies (see Astin, 1993; Cress, Astin, Zimmerman-Oster, & Burkhardt, 2001). The overall results of these studies suggested that students who were more involved on campus and who participated in leadership-oriented activities (such as holding a position with a student organization) were more likely to identify themselves as leaders and more likely to self-report elevated levels of leadership skills. However, the skills that comprised “leadership” were left to the participants completing the data collection instrument.

The Multi-Institutional Study of Leadership (MSL), begun in 2006, ushered in a wave of research (see Dugan & Komives, 2007; 2010) based on a concept of leadership practices aligned with the Social Change Model of Leadership Development (Higher Education Research Institute, 1996). These research efforts suggested that students who participated in structured leadership education initiatives, within both the curriculum and co-curriculum, also score higher on self-reported measures of leadership confidence and Social Change Model-defined skill sets. While comprehensive in the scope of its data collection, a limitation of the MSL is its cross-sequential structure; while the MSL could show that more involved students score higher on measures of socially responsible leadership, mapping the developmental trajectories of students over time was impossible. Therefore, while a growing body of evidence exists that suggests students make gains due to involvement in leadership education initiatives, national longitudinal data that indicate growth over time as a result of programmatic participation, and based on a theoretically defined model of leadership capacity, is currently severely limited. We designed this study to begin filling this important gap in our understanding of the effectiveness of leadership education.

Theoretical Model of Leadership Capacity. Many university-based leadership development programs in the United States espouse a model of leadership capacity that can be described as “post-industrial” in nature (Rost, 1993). In this framework, leaders possess ethical values, work in collaborative ways with teammates, align others around a common purpose, and engage in socially responsible actions and strategies. A recent study showed that over 80% of co-curricular leadership programs utilize models with these foundations (Owen, 2012). In our research, we utilized a tripartite “Ready, Willing, and Able” model of leadership capacity (Keating, Rosch, & Burgoon, 2014) that includes the confidence that one’s leadership actions would lead to success if enacted (being “Ready” to lead); the motivation to engage in the work and behaviors of leadership (being “Willing” to serve), and the post-industrial leadership skills required for success (being “Able” to act as a leader). As emerging leaders develop capacity

within each of these competencies – skills, confidence, and motivation – each competency helps reinforce the others. See Figure 1. For example, with increases in skill, confidence in utilizing these skills often increases. Inversely, the lack of one competency can hinder the others. When emerging leaders lack the motivation to engage in the work of leadership, they may not apply their skills.

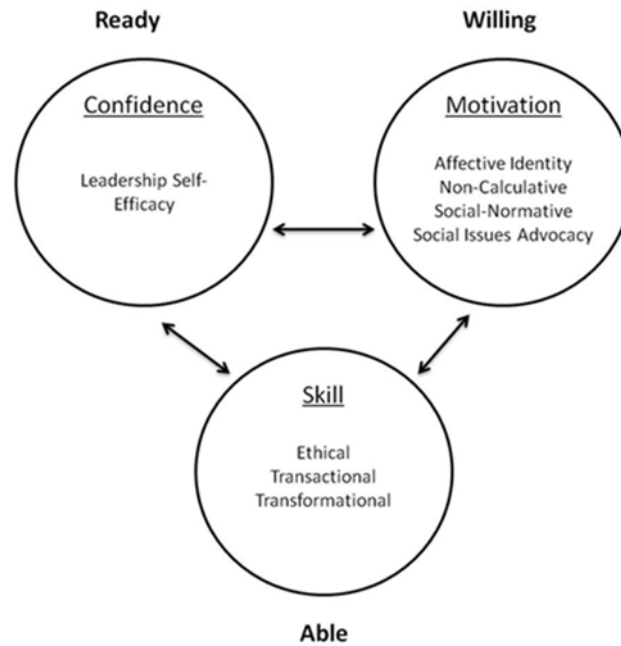


Figure 1. *The “Ready, Willing, and Able” Model of Leadership Capacity*

This study focused on examining students’ gains in their leadership capacity within the context of the model described above. The measurement of these and related gains has been problematic in the past (Uleman, 1991), while the effectiveness of current efforts that assess cross-sequential reports of single participant responses remains to be seen (Bowman, 2011). An increasing number of leadership educators (Posner, 2009) are beginning to call for the assessment of programs that are longitudinal in nature and that examine participant growth both within programs and after they have had the opportunity to implement learning within their own practice.

Research Questions. Our research focused on university students’ gains in broad-based leadership capacity that lasted beyond the end of the programmatic intervention. We were interested primarily in the gains in students’ skill, confidence, and motivation to lead. Secondly, we were interested in how university students with different social identities and leadership experiences may have differed in their gains. To these ends, our specific research questions were:

1. To what extent is a six-day immersion experience, the LeaderShape Institute, associated with long-term gains in leadership skill, confidence, and motivation to lead?
2. To what extent do differences exist in incoming levels of and long-term gains in leadership capacity by gender, race, class status, and prior leadership experience?

Methods

Population and Sample. The LeaderShape Institute is a six-day immersion retreat designed for university students to develop transformational and ethical leadership capacity. While the curriculum is designed and coordinated by LeaderShape, Inc., a not-for-profit private organization, over 100 university campuses in the United States and a small number in Canada, Mexico, and the Middle East partner with the organization to host sessions for their registered students. In addition, LeaderShape, Inc. hosts sessions during summer months that are open to any student from any university. Annually, over 4,000 students typically participate in a LeaderShape Institute session, most of whom are undergraduates.

An open call to institutions during the Fall semester of 2013 that host sessions yielded 21 universities interested in participating in this research study. These universities were diverse in terms of size, control, admissions selectivity, and faculty research output. In addition to these 21 campus-based sessions, we gathered data at four national sessions open to participants from any university and hosted directly by LeaderShape, Inc.

The sessions are hosted at disparate retreat centers. At the beginning of the session, participants were asked to complete the paper-based version of the survey before the session began. From these, 1,333 participated in the pre-test data collection, representing over 90% of all registered students at these sessions. At the end of the six-day session and before participants departed the location they were asked to again complete the paper-based version of the survey. Of the pre-test participants, 1,279 (96%) also participated in the post-test wave of data collection. The pre-test survey included additional demographic and background experience questions not included on the post-test version. Finally, approximately three months after the session, participants were contacted via email and invited to participate in an online follow-up survey. Of the pre-test participants this follow-up survey had 343 (26%) participants.

Concerned about potential bias introduced through self-selection in survey participation, we conducted extensive analysis of the participation rates in the post-test and the follow-up test. For the post-test, we employed logistic regression to estimate participation as predicted by pre-test responses, first using univariate regression and then multi-variate regression. These tested participant self-reported social identities, prior leadership training and experience levels, and responses to the eight pre-test psychometric constructs. We identified no differences between post-test participants and post-test non-participants on any of these measures. Of greater concern was the substantively lower response rate to the three-month follow-up survey invitation. We estimated the same logistic regressions to predict participation from the pre-test responses, and additionally included in the analysis the participants' changes between pre-test and post-test scores on the eight pre-test psychometric constructs—effectively estimating whether pre-test-to-post-test changes predicted follow-up test participation. Of these 32 coefficients estimated, only one variable predicted follow-up survey participation. Participants who reported higher non-calculative motivation to lead scores on the pre-test were slightly more likely to participate in the follow-up test ($t=4.10$, $p<.001$), suggesting that of all measured psychometric constructs, only participants who report leading without regard to their own personal gain would also be more likely to respond to an online survey invitation. With this singular qualification, we proceeded

with confidence that the participants in the follow-up test had no statistically significant differences from non-participants in the follow-up test.

In the pre-test, participants were asked to respond to questions regarding their social identities. With regard to gender, 58% identified as female, 32% as men, 1% reported a transgender identity, and 9% did not report. Approximately 48% of participants identified as Caucasian, 17% as African-American, 12% as Asian-American, 6% as Latino/a, 5% as multi-racial, 1% as Middle Eastern, and 11% did not report their racial identity. Freshmen comprised 27% of our sample, sophomores comprised 29%, juniors 25%, seniors 8%, and graduate students 2%, while 9% did not report a class standing. Lastly, 7% of the overall sample identified as an international student.

Also in the pre-test, participants were asked about their prior levels of leadership training and leadership roles. Approximately 26% had “never” received prior leadership training; 13% had “once;” 27% had “sometimes;” 17% had “often;” 8% had “almost always;” while 9% did not respond when asked about their prior leadership training.

Instrumentation. Our study was designed to measure broad-based leadership capacity through assessing growth in students’ post-industrial leadership skills, confidence in leading, and motivation to engage in leadership behaviors. To answer our research questions we assessed student leadership capacity using three different scales with a total of eight sub-scales. This approach enabled us to address research question one and its focus on gains in leadership skill, confidence, and motivation to lead.

Leadership Skill. We measured leadership skill through three sub-scales: The Leader Behavior Scale (Podsakoff, MacKenzie, Moorman, & Fetter, 1990), a popular and non-copyrighted 28-item instrument includes two sub-scales respectively focused on transformational and transactional leadership. A sample item for transformational leadership was, “I help other group members develop a team attitude and spirit among ourselves.” A sample item for transactional leadership was, “I always give positive feedback when other group members perform well.” Item responses have a 5-point Likert scale, ranging from “strongly agree” to “strongly disagree.” We chose the LBS due to its use, for many years, as a broad measure of transformational leadership unassociated with the narrower Full-range Transformational Leadership Model and its respective Multifactor Leadership Questionnaire (Bass & Avolio, 1997). The LBS has been in use for over 20 years as a psychometric tool (Yukl, 2010) in both business and education settings, with Cronbach reliabilities ranging from .71 to .89 (Yukl, 2010).

We also utilized the Ethical Leadership Scale, a 10-item measure designed to measure the degree to which participants incorporate ethical behavior into their leadership values and planned behaviors (Brown, Trevino, & Harrison, 2005). The ELS is correlated with the Idealized Influence scale within the popular Multi-Factor Leadership Questionnaire (Bass & Avolio, 1997), measuring aspects of transformational leadership. However, as ethical behavior and transformational leadership theoretical are related but distinct concepts (Bass, 1998), confirmatory factor analysis shows that the ELS measures a distinct leadership capacity but is related to transformational leadership (Brown, et al.) Moreover, initial psychometric

examination of the ELS showed no significant relationship with social desirability measures or social identity demographic factors (Brown, et al).

Leadership Confidence. To assess students' confidence in leading, we utilized the Self-Efficacy for Leadership (SEL) scale, an 8-item measure of a person's confidence in engaging in leadership behaviors (Murphy, 1992). The SEL has been in use for 20 years in professional and educational environments and, like the LBS, has undergone extensive psychometric examination (Hoyt, 2005). A sample item is, "I know how to encourage good group performance." Item responses have a 5-point Likert-scale, ranging from "strongly disagree" to "strongly agree." Research has shown that internal reliability is good, i.e., above .76 (Murphy & Ensher, 1999), and the scale has been shown to possess convergent and discriminant validity when used with measures of self-esteem and leadership experiences (Hoyt, 2005).

Leadership Motivation. Motivation to engage in leadership behaviors was measured using the Motivation to Lead (MTL) scale (Chan & Drasgow, 2001), a 27-item measure equally divided across measures of affective-identity (AI), non-calculative (NC), and social-normative (SN) motivations to lead (Chan & Drasgow, 2001). The MTL assesses the degree to which people feel "called" to lead as well as the energy they feel to engage in leadership behaviors. The AI scale concerns the degree to which an individual is personally drawn to leadership roles and includes items such as, "Most of the time, I prefer being a leader rather than a follower when working in a group." The NC scale concerns the degree to which a person avoids rationally calculating the individual costs and benefits of holding a leadership position and includes items such as, "I never expect to get more privileges if I agree to lead a group." The SN scale is used to determine the degree to which a person leads due to a sense of duty or responsibility to others and includes items such as, "People should volunteer to lead rather than wait for others to ask or vote for them." Responses use a 5-point Likert scale, ranging from "strongly disagree" to "strongly agree." The scale has been used primarily in professional and public organizations and has alpha reliabilities ranging from a low of .65 for the NC scale to a high of .91 for the AI scale. The MTL has recently been extended to the field of higher education as a tool to measure student leadership development (Rosch, 2014).

Social Issues Advocacy. Due to the LeaderShape curriculum focus, in part, on the application of social justice behaviors, we also included an adapted version of the Social Issues Advocacy Scale (SIAS; Nilsson, Marszalek, Linnemeyer, Bahner, & Misialek, 2011), recently designed to assess a person's motivation to publicly advocate for social issues considered both socially just and important to the person completing the measure. The original scale was developed by its authors to also measure aspects of political advocacy (i.e. motivation to vote, lobby, or campaign for people or political issues); these items were not included in the current study. Language within each item was also adapted for an educational environment, where we changed phrases such as "I am professionally responsible to..." to "I am personally responsible to..." Sample items within the SIAS include, "I am personally responsible to confront friends and colleagues who display signs of discrimination," and "I use social media to advocate for social issues that are important to me." While the scale has been in use for less time than the others, and therefore possesses fewer psychometric markers of strength, the original research conducted to create and validate the scale indicates an acceptable degree of convergent validity

with measures of multicultural empathy, as well as discriminant validity when measured with items on self-esteem and life satisfaction.

Data Collection. We assessed participants in three waves of data collection. A pre-test was completed in person prior to any participation at the LeaderShape Institute, often at an orientation session weeks prior to the session, but in some cases immediately prior to the session beginning. A post-test was administered immediately at the conclusion of the session, often in person and before students departed from the retreat center, but after all curriculum had been delivered. Approximately three months after their LeaderShape Institute session, we invited participants via email to complete a follow-up test through an online survey via Qualtrics survey software. The specific time frames for each LeaderShape Institute's follow-up test data collection period was determined based on their institution's unique academic calendar.

Data Analysis. The bulk of our initial analysis consisted of a series of matched-sample t-tests that compared students' pre-test scores with their post-test scores and their follow-up test scores. Participant responses to questions on prior leadership training and leadership roles were categorized into high or low groups, based on their response to the 1-5 scales. These groups were constructed by responses above or below the sample's median score, respectively, while excluding responses in the median score. This created two distinct groups separated with a two-step gap on the scale for further analysis.

Effects Coding. To examine differences in participants with regard to their categorical social identities (gender, race, sexual orientation, and class year), we utilized effects coding (Mayhew & Simonoff, 2011) to examine each category relative to other participants. Traditional quantitative practices involve either reducing a categorical variable down to an indicator variable or choosing a reference group to exclude. Both practices privilege one group over the other groups for analytical and non-theoretical reasons. This results in interpretations, for example, of racial demographic variables such as "white versus non-white" or "when compared to white participants." To avoid this false dichotomy, we employed effects coding. A relatively new technique in educational research (Mayhew & Simonoff, 2011), effects coding eliminates the need for a reference group with categorical variables. Instead, all groups are reported in the results and their coefficients are interpreted as a comparison to other members in the sample. This technique estimates statistics for every sub-group and therefore enables interpretation of every sub-group. In this study, the categorical variables with regression estimates calculated using effects coding were: race (eight groups: African-American, Asian-American, Caucasian, Hispanic, Middle Eastern, Native American, multi-racial, and no response), gender (four groups: male, female, trans, and no response), and sexual orientation (five groups: straight, gay/lesbian/bisexual, rather not say, questioning, no response).

Results

Aggregated Gains in Leadership Capacity. The means and dispersion statistics for each of the leadership scales are displayed within Table 1, while differences between pre-test,

post-test, and follow-up test scores are also shown. Difference statistics listed under follow-up test columns denote a t-test utilizing these and their respective pre-test scores.

Overall, immediately after their session concluded, participants reported increases in all eight of the leadership capacities measured except for Affective Identity motivation to lead. Three to four months later, students' scores remained elevated in regards to their transformational, transactional, and ethical leadership skill, their leadership self-efficacy, and the commitment to advocate for social issues. While their Social Normative motivation to lead scores showed a significant increase after their session, these scores fell below incoming levels when measured months later.

Leadership Skill. Measured through three scales—transformational skills, transactional skills, and ethical behavior—participants reported statistically significant higher scores on all three scales in the post-test when compared to their pre-test. Transformational leadership scale increased from a pre-test mean of 3.98 to a post-test mean of 4.29, a substantive increase ($t=21.48, p<.001$). Scores on the transactional leadership scale increased from 4.20 on the pre-test to 4.40 on the post-test, a statistically significant ($t=9.34, p<.001$). The ethical leadership scale scores moved from 4.12 to 4.35, a substantial increase ($t=14.86, p<.001$).

These gains persisted into the follow-up test on transformational skills and ethical behavior, but not significantly for transactional skills. The transformational scale scores remained elevated approximately three months after the experience, going from the pretest mean of 3.98 to the follow-up test mean of 4.13 ($t=6.14, p<.001$). The ethical behavior scores also remained higher, from 4.12 pretest mean to 4.34 in the follow-up test ($t=8.69, p<.001$). The transactional scores were still higher from the pre-test in the follow-up test. However, they were no longer statistically different from the pre-test scores. This is partially a result of greater standard deviation in the transactional scores than the standard deviations of either the transformational skills scale or the ethical behavior scale.

Leadership Confidence. Using Murphy's Self-Efficacy for Leadership scale, we assessed students' confidence in their own abilities to lead effectively. In the pre-test, the mean score for the scale was 3.89. This increased substantially, and statistically significantly, in the post-test scores to 4.27 ($t=20.14, p<.001$). Three months after the program, the students reported lasting gains in their self-confidence to lead. In the follow-up test the scale's mean remained elevated at 4.04, and statistically significant ($t=4.09, p<.001$).

Leadership Motivation. To examine participants' motivation to lead, we measured this on three subscales: Affective Identity, Social Normative, and Non-Calculative as discussed earlier. These saw fewer statistically significant changes than the measures previously discussed. For Affective Identity, there were no statistically significant shifts in the mean scores of participants between the pre-test, post-test, or follow-up test. We tested for both directions, so while scores did not appreciably increase, they also did not decrease significantly.

For the Social Normative subscale, all participants' pre-test mean score was 3.71. This increased to 3.78 ($t=4.21, p<.001$) in the post-test. Importantly, in the follow-up test the mean

score shifted in the opposite direction by nearly the same magnitude to 3.57 ($t=-4.82$, $p<.001$) and the standard deviations across all three tests were consist (0.43, 0.49, and 0.46 respectively).

The mean score for the Non-Calculative subscale was 3.71 and increased considerably in the post-test to 4.24 ($t=14.42$, $p<.001$). However in the follow-up test the mean score was no longer statistically significant at 3.99 ($t=1.27$, $p<.001$).

Social Issues Advocacy. We used a newer instrument to examine participants' motivations to speak out publicly for socially just issues. In the pre-test survey participants' responses had a mean score of 3.98, which increased remarkably in the post-test to 4.15 ($t=17.58$, $p<.001$). This remained elevated and statistically significant in the follow-up test at 4.00 ($t=6.16$, $p<.001$).

Table 1. *Leadership Scales Descriptive and Difference Statistics*

	Pre-Test μ (σ)	Post-Test μ (σ)	t	Follow-up μ (σ)	t
Transformational Leadership	3.98 (.36)	4.29 (.38)	21.48***	4.13 (.42)	6.14***
Transactional Leadership	4.20 (.50)	4.40 (.54)	9.34***	4.28 (.55)	2.25
Ethical Leadership	4.12 (.38)	4.35 (.40)	14.86***	4.34 (.41)	8.69***
Self-Efficacy	3.89 (.50)	4.27 (.45)	20.14***	4.04 (.61)	4.09***
Social-Normative Motivation	3.71 (.43)	3.78 (.49)	4.21***	3.57 (.46)	- 4.82***
Non-Calculative Motivation	3.71 (.43)	4.24 (.55)	14.42***	3.99 (.64)	1.27
Social Issue Advocacy	3.98 (.51)	4.15 (.52)	17.58***	4.00 (.54)	6.16***

Notes: * $p<.05$, ** $p<.01$, *** $p<.001$

Disaggregated Gains in Leadership Capacity. Further, beyond the changes across the pre-test, post-test, and follow-up tests of all participants, we also examined differential gains with regard to each measure of leadership between men and women, and across racial

categorizations, class standing, and degree of prior leadership experience. Here we highlight emergent significant results within each grouped analysis.

Gender. Prior to participating in the LeaderShape Institute, women reported lower degrees of transformational leadership skill and confidence compared to men ($t=4.21, p<.001$). By the conclusion of the LeaderShape Institute, while both women and men reported significant gains (women: $t=16.84, p<.001$; men: $t=10.85, p<.001$), women displayed even higher scores than men ($t=2.68, p<.01$) in the post-test. Approximately three months later, both groups' scores remain elevated, while the disparity in scores disappeared.

Race. Prior to participating in the LeaderShape Institute, students reported significant differences by race in their motivation to advocate for issues of social responsibility and justice. African-American students ($t=2.43, p<.01$) and Asian-American students ($t=2.85, p<.001$) reported slightly higher levels of motivation compared to all other participants on the pre-test. However, these gaps disappeared by the conclusion of the program, and remained non-existent three to four months later. However, students identifying as Caucasian were the only racial group to report elevated levels of social issues advocacy motivation months later.

Class Standing. No differences existed in students' confidence in leading peers when compared by class standing prior to participating in the LeaderShape Institute. While all class years reported significant gains immediately after the program, only freshmen and seniors' gains persisted after three to four months (freshmen: $t=3.23, p<.01$; seniors: $t=2.98, p<.01$). Moreover, when compared to other class years, freshmen were the only class that reported durable gains in non-calculative motivation to lead ($t=2.35, p<.05$) and transactional leadership skill ($t=2.31, p<.05$) approximately three months later.

Prior Leadership Experience. While all groupings of students, reported significant gains in their confidence in leading others upon conclusion of the LeaderShape Institute, these gains lasted only for students who reported low degrees of experience or no experience in leading others prior to participating ($t=4.22, p<.001$). Similarly, while all groupings of students gain in transformational skill after participating, students who reported a low degree of prior experience in leadership training events or courses report larger gains approximately three months later when compared with those students with a high degree of prior experience ($t=4.55, p<.001$ vs $t=2.21, p<.05$). Even in areas where, in the aggregate, students did not make significant gains (such as in their Affective Identity motivation to lead, for example), the gap between students who entered with a high degree of prior leadership experience and those reporting an initial low degree of experience disappeared three to four months after their LeaderShape Institute experience.

Discussion

We conducted this research study to examine the lasting gains in broad-based leadership capacity within students after participating in a six-day leadership immersion program. Our results showed that students made gains in their skills, confidence, and motivation to lead and that these gains largely lasted at least three months beyond participating in the program. The most substantial gains observed were in the areas of transformational leadership and ethical

leadership skill, leadership self-efficacy, and motivation to advocate for issues related to social justice.

In leadership skill, the results showed that the program had a substantial impact on participants, and this impact persisted through the follow-up test for two of the three scales measured: transformational leadership and ethical leadership. In the follow-up test scores on the transactional leadership scale returned to slightly above the pre-test levels used as a baseline.

We observed substantial increased in participants' confidence as well. This was much higher in the post-test, and continued to remain elevated in the follow-up test three months later. This is evidence that the program contributed to participants' leadership self-efficacy in meaningful and lasting ways.

For the measures of leadership motivation, more varied results are observed. We measured motivation to lead using subscale measures related to affective-identity, social-normative, and non-calculative sources of motivation. Scores increased significantly from the pre-test to the post-test in social-normative and non-calculative motivation to lead. Affective-identity scores did not move significantly from the pre-test in either the post-test or the follow-up test. This lack of movement indicates that the program does not shift participants' motivation to lead based on their self-identity as a leader. Further, none of the subscales' follow-up scores showed significant increases. The marked decrease observed in social-normative motivation from pre-test to follow-up test stands out even more. These results may indicate that participants' motivation to lead based on others' expectations decreases considerably three months later.

Across all measures of leadership capacity, three areas saw noteworthy immediate and longitudinal gains. Measurements of leadership self-efficacy from pre- to post-test saw the largest gains, which shrank considerably by the follow-up test. However, responses to the scales of Transformative Leadership, Ethical Leadership, and Social Issues Advocacy all remained elevated when assessed three months later. These findings may indicate that participants had gained and retained growth in these areas at levels that are comparable to their more visible gains in confidence.

Disaggregated Results Discussion. Disaggregating results by gender, race, class standing, and prior leadership experience resulted in differences in gains within and across these groupings. Broadly speaking, groups of students with lower incoming capacity made gains after participating within the program, and three months later, displayed levels of leadership capacity equal to students who entered the program at higher levels.

Briefly, examining class year and leadership experience, we observed important differences in a three categories. The reported gains in leadership self-efficacy were persistent three months later only for freshmen and seniors. Freshmen also were the only class year to report significantly elevated gains when compared to other class years in non-calculative motivation to lead and in transactional leadership. Participants reporting low prior leadership experience also reported high gains in leadership self-efficacy. However, the differences in prior leadership experience did not correlate with any other differences in scores three months later.

Participation in the LeaderShape Institute results in largely similar effects, regardless of the prior leadership experience level of the participant.

We observed that women entered the program reporting lower scores on transformational leadership and on leadership self-efficacy when compared to men entering the program. However, women reported higher scores in the post-test—higher than men—but no differences emerged between genders in the follow-up scores. While one could interpret this as women gaining more from the program, one may also interpret this as the program curriculum affirming skills that were under-reported in the pre-test, while the follow-up test more accurately demonstrates the parity between genders emerging three months later.

The largest differences we observed among racial groups emerged on their respective propensity to advocate publicly for social issues. Several racial groups were numerically small, and due to the post-hoc analytical techniques used, we only detected large effect sizes that were also statistically significant. In the pre-test scores on the SIAS, both African-American participants and Asian-American participants reported higher motivation to advocate for social issues as compared to all other participants. This gap disappeared in the post-test and in the follow-up test, with no significant differences between racial groups. Durable changes over time, from the pre-test to the follow-up test, emerged in Caucasian participants gaining higher motivation to advocate for social issues measured three months after the experience compared to before it started. Collectively, these results demonstrate that many participants of color enter with higher motivation to advocate for social issues and that remains unchanged over time. What did change over time was the substantial increase in Caucasian participants' motivation to advocate for social issues.

Collectively, these disaggregated results show positive outcomes for the LeaderShape Institute's key focus on social justice issues. The gains reported by women in confidence to lead demonstrate a contribution to these participants self-efficacy in leading. Our results also show that Caucasian students increase substantially in their motivation to advocate for social issues and remain motivated three months later. These positive effects indicate that the LeaderShape Institute's key focus is indeed making a positive difference in participants' approaches to leadership.

Implications

Taken together, these aggregated and disaggregated results suggest that students experience developmental opportunities differently based on their incoming capacities, and these results imply the need for leadership educators to consider providing differential experiences for students at different levels of incoming capacity, or for different types of students. Freshmen and women seemed to make the largest gains, overall, in their leadership capacity within the context of the program, according to our results. Owen (2012) reports that many co-curricular programs are still developing their resources and structure; these programs might be in an opportune level in their own development to consider differential structures.

In addition, our results showed several differences in students' scores when measured immediately after the program compared to three months later. This highlights the need for more time-sensitive assessment practice among leadership educators, as a post-program evaluation

process may neglect to show lasting aspects of leadership development. At the end of many leadership development experiences, participants have a clear recognition of what they gained. However, if the goal of leadership education is lasting growth in students' leadership capacities, then research and assessment efforts should attend to these outcomes in longitudinal timeframes.

Further Research

Next in our research is to examine how these key leadership capacities interact during the developmental process for students. This includes understanding the pathways through which students progress on their way to strengthening their leadership skills, confidence, and motivation. The research reported here indicates that students do not grow in linear or common ways in their pathways to emerging leadership.

We saw key differences in students' scores reported in the immediate post-test survey as compared to the three-month follow-up survey. Further research is needed to determine whether the follow-up survey measured scores indicative of permanent changes in leadership capacity or whether the scores continue to fluctuate beyond our three-month timeframe. A more long-term survey would deepen our understanding of lasting effects from leadership development programs.

Lastly, while our research was national and representative in scope related to participants, it only examined one leadership development curriculum. As evidenced in this study with differing results in the post-test and follow-up test the duration of a program is a key factor. Further research focused on other workshops, courses, and conferences is necessary for a more complete picture of the student leadership developmental process.

References

- Astin, A. W. (1993). *What matters in college?* San Francisco: Jossey-Bass.
- Bass, B. M. (1998). *Transformational leadership*. Hillsdale, NJ: Erlbaum.
- Bass, B., & Avolio, B. (1997). *Revised manual for the Multi-Factor Leadership Questionnaire*. Mind Garden, Palo Alto, CA.
- Bowman, N. A., & Seifert, T. A. (2011). Can college students accurately assess what affects their learning and development? *Journal of College Student Development*, 52(3), 270-290.
- Boyer, E. L. (1987). *College: The undergraduate experience in america*. New York: Harper and Row.
- Brown, M. E., Treviño, L. K., & Harrison, D. A. (2005). Ethical leadership: A social learning perspective for construct development and testing. *Organizational Behavior and Human Decision Processes*, 97(2), 117-134.

- Brungardt, C., Greenleaf, J., Brungardt, C., & Arensdorf, J. (2006). Majoring in leadership: A review of undergraduate leadership degree programs. *Journal of Leadership Education*, 5(1), 4-25.
- Chan, K. Y., & Drasgow, F. (2001). Toward a theory of individual differences and leadership: Understanding the motivation to lead. *Journal of Applied Psychology*, 86(3), 481-498.
- Cress, C. M., Astin, H., Zimmerman-Oster, K., & Burkhardt, J. C. (2001). Developmental outcomes of college students' involvement in leadership activities. *Journal of College Student Development*, 42(1), 15-27.
- Dugan, J. P. (2011). Pervasive myths in leadership development: Unpacking constraints on leadership learning. *Journal of Leadership Studies*, 5(2), 79-84. doi: 10.1002/jls.20223
- Dugan, J. P., & Komives, S. R. (2007). *Developing leadership capacity in college students: Findings from a national study*. A report from the Multi-Institutional Study of Leadership. College Park, MD: National Clearinghouse for Leadership Programs.
- Dugan, J. P., & Komives, S. R. (2010). Influences on college students' capacities for socially responsible leadership. *Journal of College Student Development*, 51(5), 525-549.
- Higher Education Research Institute. (1996). *A social change model of leadership development (Version III)*. Los Angeles: University of California Higher Education Research Institute.
- Hoyt, C. L. (2005). The role of leadership efficacy and stereotype activation in women's identification with leadership. *Journal of Leadership and Organizational Studies*, 11(4), 2-14.
- Kellerman, B. (2012). *The end of leadership*. New York, NY: HarperCollins.
- Kezar, A. J., & Moriarty, D. (2000). Expanding our understanding of student leadership development: A study exploring gender and ethnic identity. *Journal of College Student Development*, 41(1), 55-69.
- Mayhew, M. J. and Simonoff, J. S. (2011). Nonwhite, no more: Effect coding as an alternative to dummy coding with implications for higher educational researchers. *Association for the Study of Higher Education 2011 Conference*, Vancouver, BC.
- Multi-Institutional Study of Leadership. www.leadershipstudy.net.
- Murphy, S. E., & Ensher, E. A. (1999). The effects of leader and subordinate characteristics in the development of Leader-Member Exchange quality. *Journal of Applied Social Psychology*, 29(7), 1371-1394.

- Nilsson, J. E., Marszalek, J. M., Linnemeyer, R. M., Bahner, A. D., & Misialek, L. H. (2011). Development and Assessment of the Social Issues Advocacy Scale. *Educational and Psychological Measurement*, 71(1), 258-275.
- Ofori, G. (2009). Ethical leadership: Examining the relationships with full range leadership model, employee outcomes, and organizational culture. *Journal of Business Ethics*, 90(4), 533-547.
- O'Leonard, K. (2014). *The corporate learning factbook 2014: Benchmarks, trends, and analysis of the U.S. training market*. Deloitte Development, LLC.
- Owen, J. (2012). *Examining the design and delivery of collegiate student leadership development programs: Findings from the Multi-Institutional Study of Leadership (MSL-IS), a national report*. Washington, DC: Council for the Advancement of Standards in Higher Education.
- Keating, K., Rosch, D. M., & Burgoon, L. (2014). Developmental readiness for leadership: The differential effects of leadership courses on creating "ready, willing, and able" leaders. *Journal of Leadership Education*, 13(3), 1-16.
- LeaderShape (n.d.) www.leadershape.org
- Murphy, S. E. (1992). *The contribution of leadership experience and self-efficacy to group performance under evaluation apprehension*. (Ph.D. 9230410), University of Washington, United States -- Washington. Retrieved from <http://search.proquest.com/docview/304005264?accountid=14553> ProQuest Dissertations & Theses (PQDT) database.
- Podsakoff, P. M., MacKenzie, S. B., Moorman, R. H., & Fetter, R. (1990). Transformational leader behaviors and their effects on followers' trust in leader, satisfaction, and organizational citizenship behaviors. *The Leadership Quarterly*, 1(2), 107-142.
- Posner, B. Z. (2009). A longitudinal study examining changes in students' leadership behavior. *Journal of College Student Development*, 50(5), 551-563.
- Priest, K. L., & Donley, S. (2014). Developing leadership for life: Outcomes from a collegiate student-alumni mentoring program *Journal of Leadership Education*, 13(3), 107-117. doi: 1012806/V13/I3/A2
- Rosch, D. M. (2014). Predicting student leadership in agricultural professional preparation organizations. *NACTA Journal*, 58(1), 7-10.
- Rosch, D. M., & Caza, A. (2011). The durable effects of short-term programs on student leadership development. *Journal of Leadership Education*, 12(1), 28-48.

Rosenthal, S. A. (2012). *National leadership index 2012: A national study of confidence in leadership*. Cambridge, MA: Center for Public Leadership, Harvard Kennedy School.

Rost, J. C. (1993). *Leadership for the 21st century*. Westport, CT: Praeger Publishers.

Uleman, J. S. (1991). Leadership ratings: Toward focusing more on specific behaviors. *Leadership Quarterly*, 2(3), 175-187.

Yukl, G. (2010). *Leadership in organizations* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Zenger, J. (2012). Does leadership development really work? *Forbes*, May 2, 2012. Retrieved from <http://www.forbes.com/sites/jackzenger/2012/05/02/does-leadership-development-really-work-2/>

Author Biographies

David M. Rosch serves as an Assistant Professor at the University of Illinois at Urbana-Champaign. His particular areas of interest include programmatic training in leadership development and the accurate assessment of leadership effectiveness in student and professional organizations. He earned a Ph.D. in Higher Postsecondary Education from Syracuse University, a M.S. in Student Affairs in Higher Education from Colorado State University, and a B.A. in Psychology from Binghamton University (NY).

Clinton M. Stephens is a lecturer for leadership education at Iowa State University. Currently, Stephens coordinates the Catt Center's leadership program and teaches classes in leadership development. Stephens studies student leadership development, specifically assessing the effectiveness of courses and workshops to develop participants' leadership skills. Stephens completed a B.S. in Business Administration at Kansas State University, a M.S. in College Student Development at Oklahoma State University and a Ph.D. in Higher Education Administration at Iowa State University.

Jasmine D. Collins is a doctoral student in the department of Education Policy, Organization and Leadership at the University of Illinois at Urbana-Champaign. Her research interests include higher education diversity policy, campus climate, access to higher education, and student leadership development. Collins holds a Master of Education (Ed.M.) in Educational Policy Studies from the University of Illinois at Urbana-Champaign and a B.A. in Communication and Culture from Indiana University.