

Journal of Applied Juvenile Justice Services

An Evaluation of “Transforming Youth Justice: A Leadership Development Program”

Kevin J. Earl, MS
David L. Myers, PhD
University of New Haven
Henry C. Lee College of Criminal Justice and Forensic Sciences
Department of Criminal Justice

As part of the modern evidence-based movement in juvenile justice, effective leadership is necessary for planning, implementing, and evaluating policies, programs, and practices, along with guiding innovative reform efforts. Leadership development programs hold promise for advancing the field of juvenile justice, by enhancing participant leadership skills and abilities, as well as increasing opportunities for networking and further professional development. The purpose of the current study was to evaluate a juvenile justice leadership program in Connecticut. Overall, the findings from the mixed-methods research were supportive. Implications for similar programs and future research also are discussed.

AN EVALUATION OF “TRANSFORMING YOUTH JUSTICE: A LEADERSHIP DEVELOPMENT PROGRAM”

During the past 30 years, juvenile courts and the field of juvenile justice experienced significant reform. Prior to the 21st century, rising juvenile crime rates, nationwide concern over youth violence, and a politically popular “get tough” movement produced a more “criminalized” juvenile justice system, focused on accountability, punishment, and enhanced public safety (Feld, 2017; Myers, 2005). Over the past 20 years, however, juvenile justice reform efforts have emphasized a more balanced approach to punishment and rehabilitation, with a return to a “kids are different” philosophy (Feld, 2017). Research on child and adolescent development, combined with greater knowledge about evidence-based programs and practices (i.e., those supported by scientifically rigorous research findings), played a large role in this movement (Howell, Lipsey, & Wilson, 2014). Greater attention also was given to collecting and utilizing data in juvenile justice agencies and organizations in order to better understand client risk and needs, match clients with empirically supported services, and assess whether specified goals, objectives, and outcomes are being achieved (Myers, 2013a).

Within this modern evidence-based movement, the importance of leadership has come to the forefront. Effective juvenile justice leaders are expected to communicate a knowledge and understanding of evidence-based policies, programs, and practices; balance strong interpersonal skills with the ability to motivate staff toward higher performance; assess and collaboratively shape organizational culture; engage in strategic planning and performance measurement; and facilitate organizational capacity-building and sustainability (Howell et al., 2014; Lipsey, Howell, Kelly, Chapman, & Carver, 2010; Myers, 2013b; Seigle, Walsh, & Weber, 2014). However, juvenile justice practitioners may not possess strong skills and abilities in these areas, making it difficult to develop and provide effective leadership in many agencies.

In the state of Connecticut, contemporary juvenile justice reform has been pronounced (Ma, Pfrommer, & Cooper, 2018). In the 1990s, juvenile detention centers and secure institutions were severely overcrowded, mental illness and suicide rates in these facilities were high, and placing juveniles in the adult criminal system was common. Following the turn of the century, mounting calls for reform led to a deinstitutionalization movement, along with “raise the age” legislation in 2007 that extended the age of juvenile court eligibility from 15 to 17 years old (i.e., 18- year-olds are prosecuted in the adult criminal system). Connecticut subsequently led the nation by reducing juvenile commitments 75% between 2001 and 2013 (TYJI Newsletter, Spring 2016). In 2014, with the support of the Tow Foundation, Connecticut established the Juvenile Justice Policy and Oversight Committee (JJPOC) to supervise continued reform, evaluate juvenile justice policies, and develop goals and recommended changes for juvenile justice laws and procedures. Between 2015 and 2017, diversion from juvenile courts increased by nearly 20%, youth incarceration declined by almost 50%, and official recidivism fell by 2% (Ma et al., 2018).

At the same time as the establishment of JJPOC, a collaboration between the State of Connecticut, the University of New Haven, and the Tow Foundation resulted in the creation of the Tow Youth Justice Institute (TYJI; Ma et al., 2018). Initiated in October 2014, among the primary goals of TYJI was to develop and offer a transformational leadership development program, designed for mid-level leaders from public and private agencies and organizations, to create a network of individuals to sustain youth justice reform efforts. “Transforming Youth Justice: A Leadership Development Program” was initiated in April 2016, with a cohort of 16 participants (TYJI Newsletter, Fall 2016). Subsequent cohorts began in April 2017 and April 2018. Over the course of this 9-month program, participants meet monthly at the University of New Haven to develop collaborative leadership skills, enhance knowledge and understanding of juvenile justice reform and evidence-based practices, and acquire core competencies in utilizing empirical research, data-driven decision-making, and restorative justice approaches (TYJI Newsletter, March 2017).

The purpose of the current study was to evaluate the leadership development program provided by TYJI. Following a brief review of the relevant literature, TYJI and its leadership development program are described in more detail. Next, basic research questions guiding the evaluation are presented, along with the research design and mixed-methods employed.

Based on the use of multiple data sources, the analysis and findings are discussed, followed by conclusions and recommendations. In conducting and reporting on this research, we hope to inform practitioners in the juvenile justice community about what is happening in the area of leadership development and potentially help guide similar efforts nationwide.

LEADERSHIP DEVELOPMENT PROGRAMS

In 1983, the W. K. Kellogg Foundation funded the development of the first organized statewide leadership development program in the nation (Black & Earnest, 2009; W. K. Kellogg Foundation, 2001). There were two primary goals of this and other subsequent programs: 1) develop leadership skills of participants, and 2) improve participants' knowledge of specified topics (Black & Earnest, 2009; Carter & Rudd, 2000). Black and Earnest (2009) and the W.K. Kellogg Foundation (2001) assert that in order to meet these goals, leadership development programs should focus on the personal, professional, policy, and practice levels of growth. Personal growth can be developed by broadening perspectives, increasing self-confidence, and helping participants gain a clearer sense of self-purpose or self-efficacy. Professional growth can be enhanced through teaching innovative approaches to management and business, as well as increasing industry representation and participation in leadership roles. Policy and practice improvements, including strengthening resource networks, also are promoted among participants in these programs.

During the past 35 years, implementation and evaluation of leadership development programs expanded rapidly, and researchers have produced various quantitative and qualitative findings across multiple disciplines. Findings generally depend upon the methods used to study each program's design and implementation. Overall, the evidence is in support of leadership development programs achieving their goals, with experiential approaches, structured behavioral modeling, and a focus on building both human and social capital being associated with measurable program success (Avolio, Hannah, Reichard, Chan, & Walumbwa, 2009; Brungardt, 1996; Day, 2001; Day, Fleenor, Atwater, Sturm, & McKee, 2014). However, critics have questioned the rigor of much of this research, as well as the actual impact of leadership development programs on organizational and behavioral outcomes (Black & Earnest, 2009; Edwards & Turnbull, 2013; Packard & Jones, 2015).

One early evaluation focused on a yearlong program designed for school superintendents in Florida (McCauley & Hughes-James, 1994). The Chief Executive Officer Leadership Development Program (CEOLDP), developed by the Center for Creative Learning, was found to produce beneficial results. This program started with a 6-day classroom component, intended to make the participants aware of their strengths and weaknesses as leaders. During the classroom experience, participants set their self-development goals and action plans. Throughout the year, participants worked with an "executive facilitator," which was another superintendent serving as a guide, supporter, and advisor, to work toward their goals and implement projects to improve their school systems. The participants also kept a journal. CEOLDP participants exhibited three outcomes, including 1) developing strategies and competencies for continuous learning, such as reflective thinking, forming

developmental relationships, and self-awareness; 2) undergoing personal changes in the way they thought about issues or problems, in their behaviors or habits, or in the way they felt about leadership; and 3) accomplishing projects more successfully than if they did not participate in the CEOLDP.

McCauley and Hughes-James (1994) concluded that the major components of the program (classroom experience, executive facilitators, journal writing, and learning projects) contributed to the success of the program in different ways. Classroom experience contributed to raising self-awareness and learning about management and leadership content. The classroom also facilitated and fostered relationships among the participants. Executive facilitators were able to provide expertise and advice, thereby helping participants construct knowledge, by asking stimulating questions, providing feedback, and acting as a sounding board. The facilitators also helped keep the participants on track, through supporting and encouraging them along the way. Reflective writing through journal entries allowed the participants to review and learn from their experiences, consider ways of dealing with issues in the future, explore their feelings, and maintain focus on their goals. Finally, the learning projects stimulated changes within the superintendents themselves, while enhancing their ability to involve others and increasing reflective thinking.

Leadership development evaluations quickly grew in number as the popularity of programs spread across disciplines. An early meta-analysis of 83 leadership development program studies by Collins and Holton (2004) revealed many positive effects, leading the authors to conclude that “practitioners can attain substantial improvements in both knowledge and skills if sufficient front-end analysis is conducted to assure that the right development is offered to the right leaders” (p. 217). Other reviews and individual studies revealed that leadership development actually starts at a young age and involves multiple interactions that persist over time (Brungardt, 1996; Day et al., 2014), while the benefits of later leadership development programs can be enhanced through ongoing coaching, mentoring, networking, and active learning (Day, 2001). Although the combined findings of several decades of research indicate positive results from leadership development programs (Avolio et al., 2009; Day et al., 2014), important research limitations include organizational and performance outcomes rarely being measured, and relatively few studies utilizing multiple sources of data (Black & Earnest, 2009; Edwards & Turnbull, 2013; Packard & Jones, 2015).

In sum, the available literature suggests leadership development programs can produce beneficial results for participants, although studies vary in quality, and much still remains to be learned about program content, implementation, and outcomes. In addition, we are not aware of any previously published evaluation of a leadership development program specifically focused on juvenile justice professionals. Therefore, to add to the existing research, the current evaluation sought to examine the implementation and impact of a relatively new leadership development program uniquely focused on youth justice reform in the field of juvenile justice.

CURRENT EVALUATION

TYJI was founded in the Fall of 2014 to support and sustain major youth justice reform efforts in the State of Connecticut and across the northeast region, and to increase the use of evidence-based practices in the field (Tow Youth Justice Institute, March 2016). The institute is unique in its multidisciplinary research-driven structure, its diverse partnerships and its broad scope. The work of TYJI spans academic research; partnerships with public agencies, nonprofits, and community members; curricula development for future generations of youth justice advocates; advice to policymakers; and leadership development. TYJI serves as a research partner to the Connecticut JJPOC, but it aspires to serve as a national model for expanding and sustaining progressive reform and demonstrating how alternative approaches can become institutionalized. At its founding, the institute established four primary objectives:

1. Develop a youth justice leadership training program targeting law enforcement and key staff in public and private agencies.
2. Develop a new youth justice curriculum to broaden course offerings and support the formation of strong, innovative youth justice undergraduate and graduate programs, and to provide a national model for dissemination.
3. Increase opportunities for students to participate in youth justice internships and cooperative work experiences in collaboration with state, local, and private agencies.
4. Provide research and evaluation services, as well as evidence-based program development assistance to public and private agencies in the youth justice system.

Concerning Objective #1, “Transforming Youth Justice: A Leadership Development Program” was initiated as an intensive, 9-month program for mid-level managers in the juvenile justice system, along with community stakeholders and advocates. Three goals of the leadership program include:

1. Building the capacity of present and future leaders as agents of change, thereby transforming youth justice in a sustainable community response paradigm.
2. Advancing leadership development skills and knowledge of best practices in reform, organizational and community change, and self-awareness.
3. Becoming a resource to the organizations, communities, and systems serving youth, through a network of dedicated leaders.

Program applicants must go through an intensive application process that includes the submission of a personal statement and a letter of recommendation. They must also fill out a form in which they describe their experiences, both professionally in their career and in other leadership development programs, as well as in community service and volunteer work, as it pertains to youth justice.

Selected program participants meet monthly to complete the program, which begins with a 2-day process with a half-day orientation. Subsequent sessions are held on a single day. In general, each session emphasizes youth justice reform, best practice strategies, and leadership competencies. Facilitator outlines and formal presentations are developed for all

sessions, along with corresponding local resources and reference materials. Delivery of the program includes guest speaker presentations, as well as a participant “circle process” based on restorative justice practices for storytelling and personal reflections. Throughout the 9-month program, cohort members experience many opportunities for individual assessment and critical reflection, small group discussions, and collective feedback, along with learning from subject matter experts, experienced practitioners from the field, and university faculty.

Completion of the program requires participants to produce a capstone project, which progresses through five stages: brainstorming and issue development, planning guidelines, proposal outlines, plan development, and a final strategy with expected outcomes. Throughout the capstone development process, on multiple occasions, participants meet with each other in small group or one-on-one meetings to discuss their capstones and provide feedback. They also participate in a peer-review process, which encourages them to help each other produce a strong capstone project. Through the small-group activities, participants also are encouraged to discuss the leadership skill(s) they used in planning for developing their concepts.

Three cohorts of participants completed the leadership development program in 2016, 2017, and 2018. To evaluate this program, four research questions were formulated:

1. Is the intended target population being reached? Answering this question occurred primarily through an assessment of program application data and a comparison of accepted and declined applicants.
2. Are the intended services being provided? This question was answered through an examination of available program documentation, session-by-session participant survey data, and direct observation of monthly sessions for the third (2018) cohort only.
3. Are the intended personnel and other stakeholders in place and engaged? This question also was answered through assessing the session-by-session participant survey data and direct observation of program sessions.
4. Are the intended outcomes being achieved? Answering this question was based on the results of a participant pre-test/post-test survey and a follow-up survey of participants who completed the program in Cohort 1.

RESEARCH DESIGN AND METHODS

To answer the four research questions, five sources of data were utilized. Each is described below. We utilized a mixed-methods approach to address at least some of the limitations associated with prior research, as well as to reduce the potential for bias and enhance the validity of our findings.

Available application data and program documentation (2016, 2017, and 2018)

In order to address the first research question, whether the intended target population is being reached, available application data for all three cohorts were utilized. All application forms for the first three cohorts were used to extract data for examining demographic variables (e.g., gender, years of experience, education attained) for those applicants accepted

and not accepted into the program. Univariate and bivariate statistics, including t-tests and chi-square analysis, were used to examine the application data.

Participant session-by-session assessment (2016 and 2017)

At the end of every session, participants were provided with a survey to capture participant perceptions of that session. Each survey contained five quantitative items based on a Likert scale, three open-response questions, and an optional prompt for additional comments. The survey responses for the first two cohorts were examined using t-tests and a general linear model (GLM) with repeated measures ANOVA. With these data and analyses, participant satisfaction with each session (e.g., materials, concepts, speakers, and activities) and consistency in satisfaction from session-to-session were assessed based on data collected from the first two cohorts. Participant attendance at each session also was examined.

Participant Pre/Post Test (2016 and 2017)

Before the start of the program, all participants were given a pre-test survey to capture each participant's base or starting knowledge, engagement, involvement, and familiarity regarding juvenile justice practices, concepts, and initiatives. The survey contained 25 individual questions, with 7 follow-up questions pertaining to 5 of the 25 individual questions (i.e., there were 32 survey items in total). Five questions were presented in a dichotomous (yes/no) format, 24 questions were presented in Likert format, and there were three open-ended questions. At the end of the program, typically during the final session for the cohort, the survey was distributed again to capture the post-test scores for each survey item. In order to address the fourth research question, whether the intended outcomes are being achieved, t-tests were used to examine whether a statistically significant difference existed for each item between pre- and post-test. A change in a total combined score also was assessed.

Alumni follow-up survey (2016)

An 18-month follow-up survey was distributed to the first (2016) cohort after completion of the program. This alumni survey contained 29 items presented in dichotomous (yes/no), visual analogue, and multiple-response formats. The aim of the follow-up survey was to capture the impact the leadership development program had on its participants in a range of areas (e.g., leadership skills/practices, capstone project completion, juvenile justice policy change, and educational attainment). The follow-up survey distributed to the first cohort was used to address the fourth research question (i.e., whether the intended outcomes are being achieved).

In-class direct observation (2018)

A qualitative approach was used to add data for examining the second and third research questions (whether the intended services are being provided and whether the intended personnel and stakeholders are in place and engaged), as well as to provide insight as to what the day-to-day functioning of the program entailed. The first author attended each program session for the third cohort, along with graduation ceremonies for the second

(2017) and third (2018) cohorts, to observe the program directly and record qualitative field notes.

ANALYSIS AND FINDINGS

The five data sources discussed above were utilized to answer the four articulated research questions. Analysis and findings from the data sources are discussed below.

Available Application Data and Program Documentation

The communities served. To gain an understanding of the communities in which program participants serve, a frequency distribution was created to examine the cities in which accepted applicants work. Table 1 provides this information for all three admitted cohorts (N = 48). As indicated, New Haven is represented most frequently, with 13 participants (27.1%). Bridgeport is represented second most frequently, with seven participants (14.6%). Tied for third most frequent, with three participants (6.3%) each, are East Haven, Hartford, Middletown, Waterbury, and Wethersfield. Two participants (4.2%) worked in Stamford. The remainder of the cities listed had one (2.1%) participant each. In general, the program is reaching working professionals from across the state, with greater participation from those in closer proximity to the University of New Haven.

Table 1: Location of Accepted Applicants (2016, 2017, & 2018)

CITY	FREQUENCY	PERCENT	CUMULATIVE PERCENT
BRIDGEPORT	7	14.6	14.6
CHESHIRE	1	2.1	16.7
EAST HAVEN	3	6.3	22.9
FARMINGTON	1	2.1	25.0
HARTFORD	3	6.3	31.3
MANCHESTER	1	2.1	33.3
MIDDLETOWN	3	6.3	39.6
MILFORD	1	2.1	41.7
NEW BRITAIN	1	2.1	43.8
NEW HAVEN	13	27.1	70.8
NORWALK	1	2.1	72.9
ROCKY HILL	1	2.1	75.0
STAMFORD	2	4.2	79.2
STRATFORD	1	2.1	81.3
WALLINGFORD	1	2.1	83.3
WATERBURY	3	6.3	89.6
WEST HARTFORD	1	2.1	91.7
WEST HAVEN	1	2.1	93.8
WETHERSFIELD	3	6.3	100.0
TOTAL	48	100.0	100.0

Gender differences. To gain an understanding of the gender of accepted applicants versus those who were denied, a chi-square test was used. The results are shown in Table 2. In total, 62 females and 24 males applied for the leadership development program. Of the applicants, 30 females and 18 males were accepted. This means that 48.4% of females who applied were accepted, and 75% of males who applied were accepted. The chi-square test was statistically significant ($p < .05$), meaning that males were more likely to be accepted than females. Through discussions with program personnel, it was determined that because more females than males applied, a higher proportion of males were selected from the applicants, in order to produce more equal representation based on gender.

Table 2: Gender of Applicants (2016, 2017, & 2018)

			Gender		Total
			Female	Male	
Accepted	No	Count	32	6	38
		% within Gender	51.6%	25.0%	44.2%
	Yes	Count	30	18	48
		% within Gender	48.4%	75.0%	55.8%
Total		Count	62	24	86
		% within Gender	100.0%	100.0%	100.0%

Note: Chi-square test; $p < .05$

Education. Chi-square analysis also was used to test whether there were statistically significant differences among accepted versus denied applicants with regard to education. Table 3 shows the group with the highest number of applicants ($N = 39$), as well as the greatest representation within the accepted group ($N = 24$), were those with master's degrees. The group with the second highest number of applicants ($N = 28$), along with the second greatest representation within the accepted group ($N = 15$), were those with bachelor's degrees. The group with the highest acceptance rate was the group with associate's degrees (80%), and the group with the lowest acceptance rate was the group with doctoral level degrees (20%). However, differences based on educational level were not found to be statistically significant. This indicates that, based on educational level, the number of selected participants from each group is roughly in proportion to the number of total applicants from each group, and working professionals with varying amount of formal education are being included in the program.

Table 3: Levels of Education (2016, 2017, & 2018)

			Education Attained					Total
			High School/GED	Associates	Bachelors	Masters	Doctorate	
Accepted	No	Count	4	1	13	15	4	37
		% within Education Attained	50.0%	20.0%	46.4%	38.5%	80.0%	43.5%
	Yes	Count	4	4	15	24	1	48
		% within Education Attained	50.0%	80.0%	53.6%	61.5%	20.0%	56.5%
Total		Count	8	5	28	39	5	85
		% within Education Attained	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Work experience. To gain an understanding of the work experience of the leadership development applicants, and to assess the experience level of those who applied and those who were accepted, a t-test was used. As shown in Table 4, the average years at their current job was 7.06 for those accepted (N = 47) and 5.67 for those not accepted (N = 37), which suggests that those accepted had more experience at their current job than those not accepted. However, this difference was not statistically significant. For overall work experience, the average for those accepted was 14.80 years (N = 46) compared to 14.32 years (N = 34) for those not accepted. Again, the difference in overall work experience was not statistically significant between the two groups, but in general program participants have attained substantial work experience prior to entering the program.

Table 4: Work Experience (2016, 2017, & 2018)

	Accepted	
	Yes	No
Years at Current Job	7.06	5.67
Years of Experience	14.80	14.32

Previous participation in other leadership programs. As part of the application process, applicants are asked whether they had participated previously in another leadership development program. As shown in Table 5, the slight majority of all applicants (N = 45) indicated that they had participated in another leadership program, as compared to those who had not (N = 40). Also revealed is that a similar majority of those accepted (N = 27) had participated in a leadership program, as compared to those who had not participated in a previous leadership program but were accepted (N = 21). A chi-square test was insignificant, indicating that participant selection was not based on participation in previous leadership

programs. However, it is evident that a relatively high percentage of participants (60%) had participated in a previous leadership program.

Table 5: Previous Leadership Program Experience (2016, 2017, & 2018)

			Previous Leadership Program		Total
			No	Yes	
Accepted	No	Count	19	18	37
		% within Leadership Program	47.5%	40.0%	43.5%
	Yes	Count	21	27	48
		% within Leadership Program	52.5%	60.0%	56.5%
Total		Count	40	45	85
		% within Leadership Program	100.0%	100.0%	100.0%

Participant Session-by-Session Assessment

Tables 6 through 8 show descriptive statistics based on the session-by-session evaluations completed by participants from the first two cohorts. Survey items one through five are based on a 5-point scale utilized by respondents, with 1 being “very little,” 3 being “moderately,” and 5 being “very much.” Question one asked, “Did the session content and subject matter meet objectives?” Question two asked, “Did the facilitators/presenters exhibit subject matter expertise?” Question three asked, “Did your knowledge of the subject matter increase?” Question four asked, “Did your skills in the subject matter increase?” Lastly, question five asked, “Are you more confident in your capacity to take action on this subject matter?” Higher scores (i.e., closer to 5) indicate more positive perceptions about the program and its impact.

Table 6 reflects the data for the first two cohorts, from 2016 and 2017. Calculated in each table are the mean (average) scores of each quantitative item on the session evaluation survey distributed to participants at the end of every session. As shown, the mean scores of each session suggest an overall high satisfaction rate, with the lowest individual item average being 3.85 (N=13) for Session 4, Item 4, for Cohort 2. There are numerous sessions that received a mean score of 5.00 (the highest possible average score). In consideration of relatively low variance within each of the five survey items, to test for differences between means, a composite score was created in which all mean scores for each survey item were added to create a total satisfaction score for each session. To test for significant differences between the two cohorts, t-tests were used. The only significant difference found between Cohort 1 and Cohort 2 was for Session 5. The data show that Cohort 1 was significantly more

satisfied with Session 5 as compared to Cohort 2. Overall, these findings indicate a high level of satisfaction among program participants, along with consistency in training delivery across the sessions and between the two cohorts.

Table 6: Session-by-Session Mean Responses (2016 & 2017)

Item Number	Session 1	Session 2	Session 3	Session 4	Session 5	Session 6	Session 7	Session 8	Session 9
1	4.77	4.95	4.55	4.67	4.73	4.83	4.67	4.68	5.00
2	4.85	5.00	4.90	4.75	4.77	4.83	4.71	4.89	5.00
3	4.56	4.86	4.55	4.42	4.50	4.83	4.58	4.84	4.90
4	4.37	4.57	4.40	4.08	4.36	4.39	4.33	4.79	4.90
5	4.33	4.62	4.25	4.33	4.32	4.61	4.33	4.84	4.90
Total	22.89	24.00	22.65	22.25	22.68	23.48	22.63	24.05	24.68
N	27	21	20	24	22	23	24	19	19

Table 7: Session-by-Session Mean Responses Cohort 1 (2016)

Item Number	Session 1	Session 2	Session 3	Session 4	Session 5	Session 6	Session 7	Session 8	Session 9
1	4.64	4.90	4.67	4.64	4.89	5.00	4.58	4.70	5.00
2	4.73	5.00	4.83	4.82	4.89	4.83	4.75	4.90	5.00
3	4.18	4.80	4.83	4.55	4.89	4.92	4.67	4.80	4.90
4	4.09	4.30	4.83	4.36	4.89	4.67	4.58	4.70	4.90
5	4.09	4.50	4.67	4.36	4.89	4.67	4.58	4.80	4.90
Total	21.73	23.50	23.83	22.73	24.44*	24.08	23.17	23.90	24.70
N	11	10	6	11	9	12	12	10	10

Table 8: Session-by-Session Mean Responses Cohort 2 (2017)

Item Number	Session 1	Session 2	Session 3	Session 4	Session 5	Session 6	Session 7	Session 8	Session 9
1	4.88	5.00	4.50	4.69	4.62	4.64	4.75	4.67	5.00
2	4.94	5.00	4.93	4.69	4.69	4.82	4.67	4.89	5.00
3	4.81	4.91	4.43	4.31	4.23	4.73	4.50	4.89	4.89
4	4.56	4.82	4.21	3.85	4.00	4.09	4.08	4.89	4.89
5	4.50	4.73	4.07	4.31	3.92	4.55	4.08	4.89	4.89
Total	23.69	24.45	22.14	21.85	21.46*	22.82	22.08	24.22	24.67
N	16	11	14	13	13	11	12	9	9

* T-Test for difference between Cohorts in Total Session Scores; $p < .05$

To investigate these data further, a general linear model (GLM) with repeated measures ANOVA was used (Field, 2009). Mean replacement was used to account for missing data, and pairwise comparisons were used to test for significant differences between sessions. Table 9 again shows a comparison of session means for Cohort 1 and Cohort 2 combined, with the total participant satisfaction score for each of the nine sessions. This total satisfaction score ranges from 22.25 (Session 4) to 24.68 (Session 9), again indicating a high degree of participant satisfaction (the highest average score possible would be 25.00). In terms of differences that did exist, total satisfaction for Session 9 (the final session) was found to be significantly higher than in all other sessions except Session 2. Session 2 total satisfaction was significantly higher than Session 4, and total satisfaction for Session 4 was significantly lower than for Session 2, Session 8, and Session 9. In sum, there was a fairly strong degree of stability in total satisfaction scores across the nine sessions, with satisfaction being highest in Session 9 and lowest in Session 4.

Table 9: GLM Repeated Measures for Session-by-Session Mean Responses

Item #	S1	S2	S3	S4	S5	S6	S7	S8	S9
1	4.77	4.95	4.55	4.67	4.73	4.83	4.67	4.68	5.00
2	4.85	5.00	4.90	4.75	4.77	4.83	4.71	4.89	5.00
3	4.56	4.86	4.55	4.42	4.50	4.83	4.58	4.84	4.90
4	4.37	4.57	4.40	4.08	4.36	4.39	4.33	4.79	4.90
5	4.33	4.62	4.25	4.33	4.32	4.61	4.33	4.84	4.90
Total	22.89 ⁱ	24.00 ^d	22.65 ⁱ	22.25 ^{b, h, i}	22.68 ⁱ	23.48 ⁱ	22.63 ⁱ	24.05 ^d	24.68 ^{a, c, d, e, f, g}
N	27	21	20	24	22	23	24	19	19

Note:

Superscript a indicates a significant ($p < .05$) difference from session 1

Superscript b indicates a significant ($p < .05$) difference from session 2

Superscript c indicates a significant ($p < .05$) difference from session 3

Superscript d indicates a significant ($p < .05$) difference from session 4

Superscript e indicates a significant ($p < .05$) difference from session 5

Superscript f indicates a significant ($p < .05$) difference from session 6

Superscript g indicates a significant ($p < .05$) difference from session 7

Superscript h indicates a significant ($p < .05$) difference from session 8

Superscript i indicates a significant ($p < .05$) difference from session 9

Participant Pre/Post Test

Tables 10 through 12 show mean score data for the pre- and post-program surveys for the first two cohorts. Items one through five are coded as 0 = No and 1 = Yes. All other items are coded on a 5-point scale, with 1 being “very little,” 3 being “moderately,” and 5 being “very much.” Table 10 shows pre- and post-results for the 2016 and 2017 cohorts combined, while Table 11 provides the results for the 2016 cohort, and Table 12 shows results for the 2017

cohort. In general, these data can be used to assess changes associated with program participation. However, it should be noted that some items represent organizational rather than personal work behavior (i.e., item 6: agency utilizes RBA, item 8: agency utilizes EBP, item 15: I have leadership opportunities in my agency). In this analysis, when a post-test mean score is greater than the pre-test mean score, it suggests a beneficial change resulting from the program. To test for significant changes, dependent sample t-tests were utilized. Below is a summary of results for the pre/post-test analyses for the 2016 and 2017 cohorts.

Table 10: Pre/Post Test for Both Cohorts (2016 & 2017)

Item Label	Pre-Test Mean	Post-Test Mean
1. Read Any JJ Publications	.480	.677*
2. Read JJ Reports	.467	.667*
3. Attend Youth Focused Committees	.742	.581
4. JJ Policy & Oversight Committee Meetings	.291	.677**
5. Watched JJPO Committee Meetings	.355	.710**
6. Agency Utilizes RBA	3.071	3.679*
7. Familiar with RBA	3.143	3.964***
8. Agency utilizes EBP	3.370	3.889*
9. I am familiar with EBP	2.923	4.000***
10. Experienced in research, lit review, data analysis	3.035	3.345
11. Experienced in utilizing data to inform decision	3.414	3.621
12. Leading improvement in policies for my agency	3.567	3.533
13. I have been part of change in my agency	3.833	3.967
14. I have been a leader of change in my agency	3.828	4.000
15. I have leadership opportunities in my agency	4.200	4.483
16. I have been part of change in my community	3.833	4.033
17. I have been a leader of change in my community	3.448	3.828
18. Member of coalition or organization for change	3.586	3.966
19. Leader of coalition or organization for change	3.241	3.828*
20. Familiar with other JJYJ agencies in community	3.710	4.097^
21. Partnerships with other JJYJ agencies in community	3.300	3.933**
22. Familiar with CT JJ practices and procedures	3.500	4.200**
23. Familiar with CT JJ reform efforts	3.483	4.138**
24. Involved in CT JJ reform efforts	2.903	3.645**
25. Prepared to lead change in YJ reform	4.200	4.600*
Total (Items 6-25)	73.177	82.971**

* $p < .05$, ** $p < .01$, *** $p < .001$, ^ $p < .10$

Table 11: Pre/Post Test for Cohort 1 (2016)

Item Label	Pre-Test Mean	Post-Test Mean
1. Read Any JJ Publications	.27	.5333*
2. Read JJ Reports	.3333	.60*
3. Attend Youth Focused Committees	.7333	.20*
4. JJ Policy & Oversight Committee Meetings	.20	.8667**
5. Watched JJPO Committee Meetings	.20	.7333*
6. Agency Utilizes RBA	3.0769	3.8462*
7. Familiar with RBA	3.00	3.7692**
8. Agency utilizes EBP	3.3846	3.7692
9. I am familiar with EBP	2.7692	3.8462*
10. Experienced in research, lit review, data analysis	3.20	3.20
11. Experienced in utilizing data to inform decision	3.4667	3.3333
12. Leading improvement in policies for my agency	3.6667	3.4667
13. I have been part of change in my agency	3.8571	3.8571
14. I have been a leader of change in my agency	3.7143	3.7143
15. I have leadership opportunities in my agency	4.3333	4.5333
16. I have been part of change in my community	3.6667	3.7333
17. I have been a leader of change in my community	3.4286	3.4286
18. Member of coalition or organization for change	3.7143	3.6429
19. Leader of coalition or organization for change	3.20	3.40
20. Familiar with other JJYJ agencies in community	3.800	3.8667
21. Partnerships with other JJYJ agencies in community	3.2667	3.60
22. Familiar with CT JJ practices and procedures	3.2667	4.00**
23. Familiar with CT JJ reform efforts	3.0667	4.00**
24. Involved in CT JJ reform efforts	2.40	3.40**
25. Prepared to lead change in YJ reform	4.4667	4.400
Total (Items 6-25)	70.70	78.40**

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

Table 12: Pre/Post Test for Cohort 2 (2017)

Item Label	Pre-Test Mean	Post-Test Mean
1. Read Any JJ Publications	.690	.813
2. Read JJ Reports	.600	.733
3. Attend Youth Focused Committees	.750	.938^
4. JJ Policy & Oversight Committee Meetings	.375	.500
5. Watched JJPO Committee Meetings	.500	.688
6. Agency Utilizes RBA	3.067	3.533
7. Familiar with RBA	3.267	4.133**
8. Agency utilizes EBP	3.357	4.000^
9. I am familiar with EBP	3.077	4.154**
10. Experienced in research, lit review, data analysis	2.857	3.500
11. Experienced in utilizing data to inform decision	3.357	3.929
12. Leading improvement in policies for my agency	3.467	3.600
13. I have been part of change in my agency	3.813	4.063
14. I have been a leader of change in my agency	3.933	4.267
15. I have leadership opportunities in my agency	4.067	4.433
16. I have been part of change in my community	4.000	4.333
17. I have been a leader of change in my community	3.467	4.200
18. Member of coalition or organization for change	3.467	4.267*
19. Leader of coalition or organization for change	3.286	4.286*
20. Familiar with other JJYJ agencies in community	3.625	4.313^
21. Partnerships with other JJYJ agencies in community	3.333	4.267*
22. Familiar with CT JJ practices and procedures	3.733	4.400^
23. Familiar with CT JJ reform efforts	3.929	4.286^
24. Involved in CT JJ reform efforts	3.375	3.875^
25. Prepared to lead change in YJ reform	3.933	4.800*
Total	73.714	86.357**

* $p < .05$, ** $p < .01$, *** $p < .001$, ^ $p < .10$

First, all items are in the desired/beneficial direction, except for:

- Item 3 (Attend youth focused committees) for both cohorts
- Item 12 (Leading improvement in policies for my agency) for both cohorts together and for cohort 2016 by itself
- Item 11 (Experienced in utilizing data to inform decisions) for the 2016 cohort
- Item 18 (Member of coalition or organization for change) for the 2016 cohort
- Item 25 (Prepared to lead change in youth justice reform) for the 2016 cohort

Virtually all of the pre/post mean differences listed above were small and non-significant. The only statistically significant and non-desirable pre/post mean difference was for Item 3 (attend youth focused committees), for the first (2016) cohort ($p < .05$).

Second, although in the beneficial direction, there were insignificant differences in pre/post means for:

- Item 10 (Experienced in research, literature reviews, and data analysis) for the 2016 cohort
- Item 13 (I have been part of change in my agency) for the 2016 cohort
- Item 14 (I have been a leader of change in my agency) for the 2016 cohort
- Item 17 (I have been a leader of change in my community) for the 2016 cohort

Third, and perhaps most importantly, there were a large number of statistically significant changes in the desired/beneficial direction for the following items:

- Both cohorts combined:
 - Item 1 (Read any juvenile justice publications)
 - Item 2 (Read juvenile justice reports)
 - Item 4 (Juvenile Justice Policy & Oversight Committee Meetings)
 - Item 5 (Watched Juvenile Justice Policy Oversight Committee Meetings)
 - Item 6 (Agency Utilizes Results Based Accountability)
 - Item 7 (Familiar with Results Based Accountability)
 - Item 8 (Agency utilizes Evidence Based Practices)
 - Item 9 (I am familiar with Evidence Based Practices)
 - Item 19 (Leader of coalition or organization for change)
 - Item 21 (Partnerships with other juvenile justice/youth justice agencies in community)
 - Item 22 (Familiar with CT juvenile justice practices and procedures)
 - Item 23 (Familiar with CT juvenile justice reform efforts)
 - Item 24 (Involved in CT juvenile justice reform efforts)
 - Item 25 (Prepared to lead change in youth justice reform)
 - Total combined score for Items 6-25
- First cohort (2016) only:
 - Item 1 (Read any juvenile justice publications)
 - Item 2 (Read juvenile justice reports)
 - Item 4 (Juvenile Justice Policy & Oversight Committee Meetings)
 - Item 5 (Watched Juvenile Justice Policy Oversight Committee Meetings)
 - Item 6 (Agency Utilizes Results Based Accountability)
 - Item 7 (Familiar with Results Based Accountability)
 - Item 9 (I am familiar with Evidence Based Practices)
 - Item 22 (Familiar with CT juvenile justice practices and procedures)
 - Item 23 (Familiar with CT juvenile justice reform efforts)
 - Item 24 (Involved in CT juvenile justice reform efforts)
 - Total combined score for Items 6-25
- Second cohort (2017) only
 - Item 7 (Familiar with Results Based Accountability)
 - Item 9 (I am familiar with Evidence Based Practices)

- Item 18 (Member of coalition or organization for change)
- Item 19 (Leader of coalition or organization for change)
 - Item 21 (Partnerships with other juvenile justice/youth justice agencies in community)
 - Item 25 (Prepared to lead change in youth justice reform)

Finally, there were several borderline statistically significant ($p < .10$) changes in the desired/beneficial direction:

- Both cohorts combined:
 - Item 20 (Familiar with other juvenile justice/youth justice agencies in community)
- Second cohort (2017) only:
 - Item 3 (Attend youth focused committees)
 - Item 8 (Agency utilizes Evidence Based Practices)
 - Item 20 (Familiar with other juvenile justice/youth justice agencies in community)
 - Item 22 (Familiar with CT juvenile justice practices and procedures)
 - Item 23 (Familiar with CT juvenile justice reform efforts)
 - Item 24 (Involved in CT juvenile justice reform efforts)

Overall, the combined pre-test/post-test results indicate a number of beneficial changes associated with program completion. A large majority of the pre-test/post-test differences were in the expected direction, and many were statistically significant. When combined with the other evaluation results, these findings suggest a beneficial impact from the program.

Alumni Follow-up Survey

A follow-up survey was distributed to the first (2016) cohort 18 months after completion of the program. The survey contained 30 questions designed to examine the impact the program had on participants, progress made on the participants' capstone projects, and their educational and professional achievements since completion of the program. Ten of the 15 cohort members (66.7%) responded to the survey (there were 16 initial cohort members, but one withdrew, meaning there are 15 leadership program alumni from the first cohort). The responses are based on three types of measurements: 1) a visual analogue scale that ranged from 0 to 11 centimeters, with a score near 0 indicating a low score and a score near 11 indicating a high score; 2) multiple response questions; and 3) dichotomous (yes/no) questions.

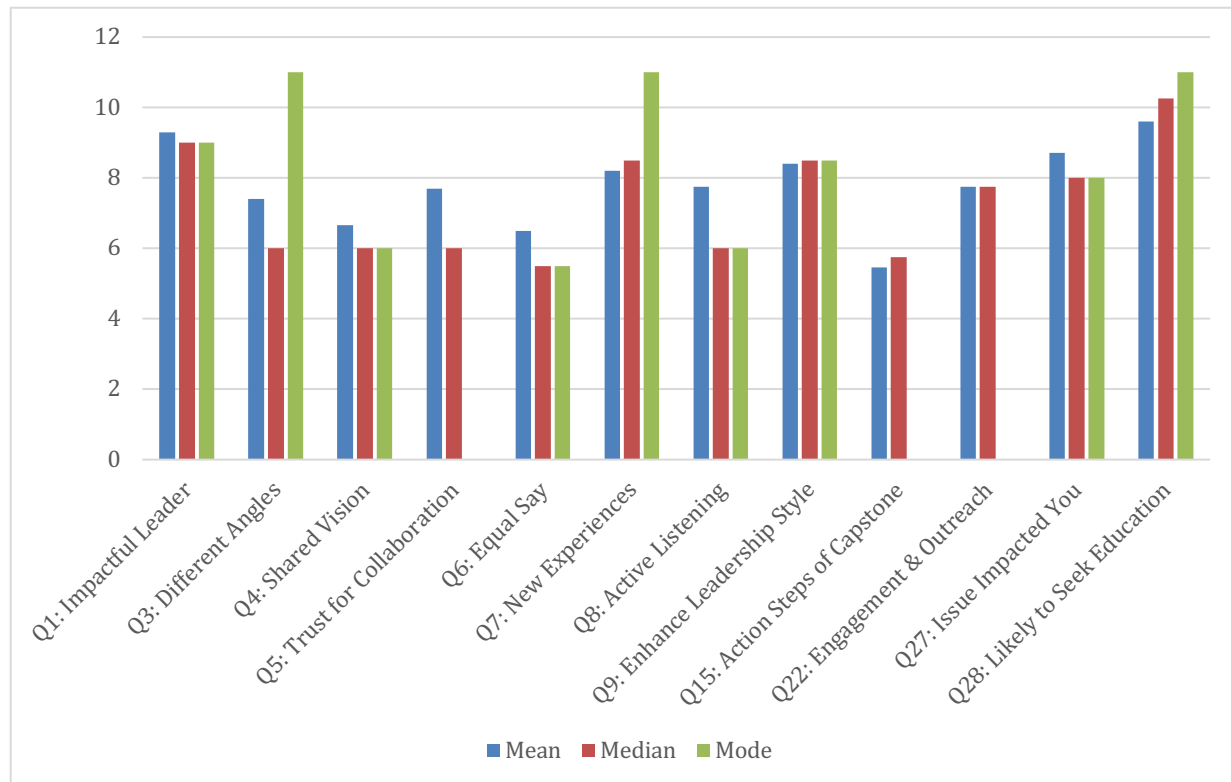
Visual analogue scale. Shown first are the survey items that were measured using a visual analogue scale. Question 1 asked if respondents define themselves now as an impactful leader, ranging from "not at all" (score of 0) to "very much" (score of 11). Question 3 stated, "I look at a perceived problem from different angles before proceeding," ranging from "strongly disagree" to "strongly agree." Other questions with responses ranging from "strongly disagree" to "strongly agree" included Question 4 ("I facilitate the development of a shared vision influenced by diverse stakeholders"), Question 5 ("I demonstrate to my peers

that I believe trust is the foundation for a successful collaboration”), Question 6 (“I create processes that ensure stakeholders have an equal say in decision making”), Question 7 (“I help people take advantage of opportunities for new experiences or learning different skills”), and Question 8 (“I listen to others actively, checking to ensure my understanding”). Question 9 asked, “In your application you noted certain leadership qualities. To what degree did your participation in the program develop or enhance your own leadership style?” ranging from “not at all” to “significantly.” Question 15 stated, “To what degree have you engaged in action steps and/or implementation of your capstone?” ranging from “not at all” to “completed.” Question 22 asked, “To what degree have you increased your outreach efforts to youth and families in your community (e.g., willingness to try a different approach, establish new commitments or cooperative agreements, and/or expanded leadership roles of others)?” ranging from “not at all” to “significantly.” Question 27 asked to what degree an issue in youth justice reform has most impacted them, their organization, and/or their community in the past year, ranging from “not at all” to “significantly.” Finally, Question 28 asked, “How likely are you to seek further education or professional development leading to career advancement?” ranging from “not at all” to “definitely.”

Table 13 and Figure 1 show the responses to the questions from the follow-up survey that were measured using a visual analogue scale. Reported in the table are the measures of central tendency (mean, median, and mode), the minimum and maximum scores for each question, and the number of respondents for each question. As indicated, all ten survey participants responded to all questions except for Item 27, which yielded 9 responses. As shown in Table 13, most of the responses were relatively high. The lowest scores were for Item 15 (5.45; the degree to which respondents have engaged in action steps or implementation of their capstone), and the highest scores were for Item 28 (9.60; the likelihood of seeking further education or professional development leading to career advancement) and Item 1 (9.30; if respondents define themselves now as an impactful leader). Figure 1 is a visualization in the form of a bar chart that shows the measures of central tendency for the responses in Table 13.

Table 13: Follow-up Survey Visual Analogue Measures

Question	Mean	Median	Mode	Min	Max	N
Q1: Impactful Leader	9.30	9.00	9.0	8.5	10.5	10
Q3: Different Angles	7.40	6.00	11.0	3.5	11.0	10
Q4: Shared Vision	6.65	6.00	6.00	3.5	11.0	10
Q5: Trust for Collaboration	7.70	6.00	6.0, 11.0	3.5	11.0	10
Q6: Equal Say	6.50	5.50	5.5	3.5	11.0	10
Q7: New Experiences	8.20	8.50	11.0	3.5	11.0	10
Q8: Active Listening	7.75	6.00	6.00	3.5	11.0	10
Q9: Enhance Leadership Style	8.40	8.50	8.50	8.0	8.50	10
Q15: Action Steps of Capstone	5.45	5.75	1.00, 8.00	1.0	10.5	10
Q22: Engagement & Outreach	7.75	7.75	7.00, 8.50	5.5	10.0	10
Q27: Issue Impacted You	8.72	8.00	8.00	8.0	10.0	9
Q28: Likely to Seek Education	9.60	10.25	11.00	6.0	11.0	10

Figure 1: Follow-up Survey Visual Analog Responses

Multiple response measures. The following figures in this section depict scores from the survey items that were measured using a multiple response measurement approach. For these items, the respondent was prompted with a question, given a list of potential responses, and told to choose the responses that were most appropriate.

Figure 2 shows the responses for Item 2 from the follow-up survey, which initially stated, “Understanding effective collaborative leadership and its traits was a major emphasis.” Participants then were asked which traits have proven most useful to them as leaders, and they were prompted to select two traits. Figure 2 indicates the “develop people” trait was most frequently chosen (28.6%, $N = 8$), and “vision & mobilize” and “self-reflection” were chosen least frequently (7.1%, $N = 2$).

Figures 3 and 4 show responses for Questions 10 and 11 from the follow-up survey. For Question 10, participants were asked, “As you interact with staff or close colleagues, which leadership skills do you practice MOST?” Question 11 followed by asking which skills they practice LEAST. Participants were asked to select their top two skills for each question. For Question 10 (MOST frequent), “shared vision” and “enable others to act” were both commonly selected (24%, $N = 6$), and “challenge the process” and “encourage the heart” were least commonly selected (16%, $N = 4$). For Question 11 (LEAST frequent), “challenge the

process” was the skill most commonly selected (40%, N = 4) and “model the way” was the only skill not selected, which indicates that it is practiced consistently among respondents.

Figure 2: Follow-up Survey Multiple Response Item 2

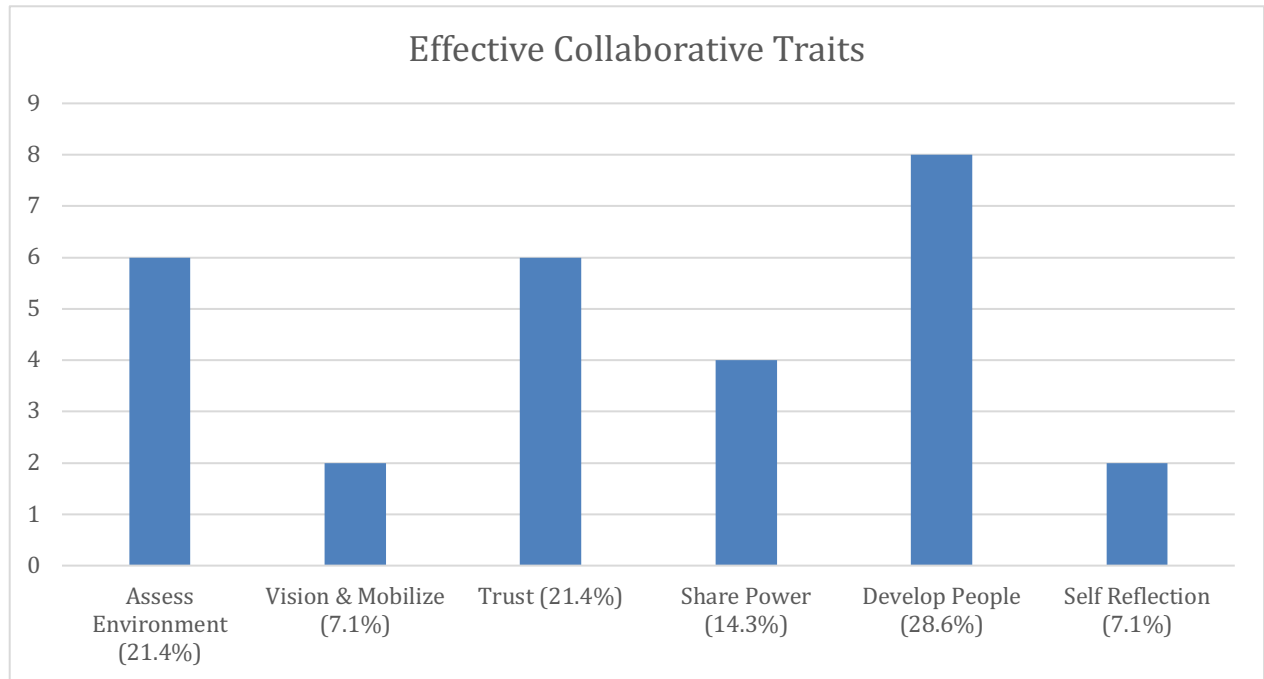


Figure 3: Follow-up Survey Multiple Response Item 10

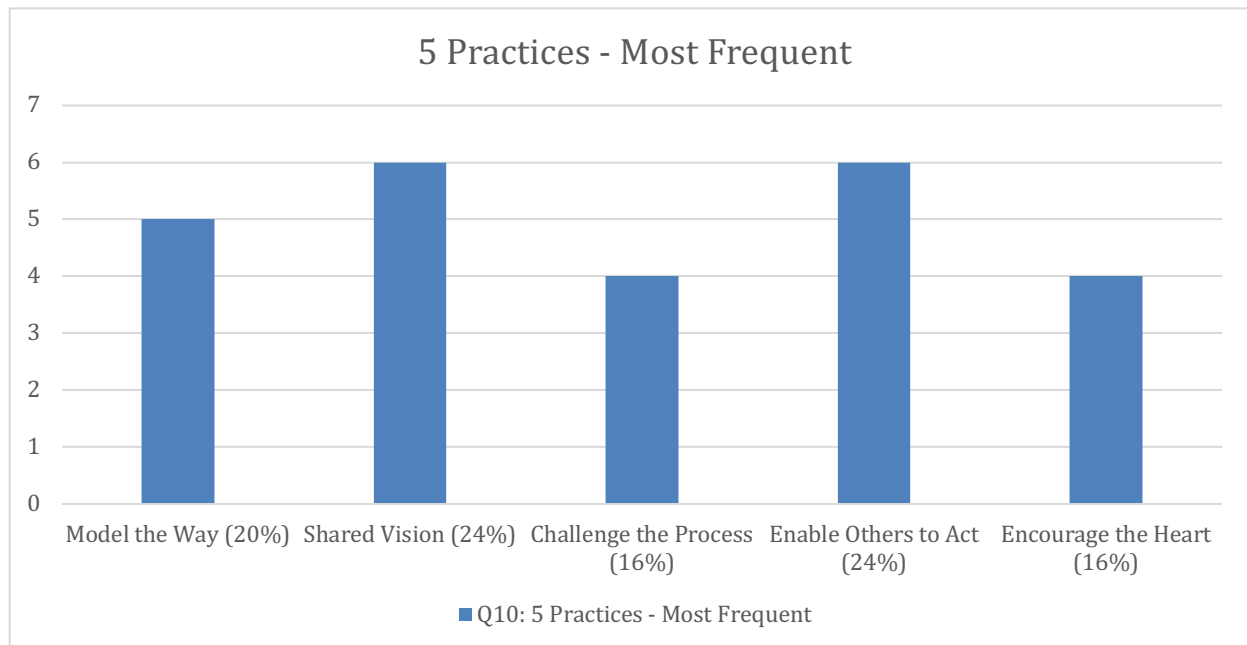
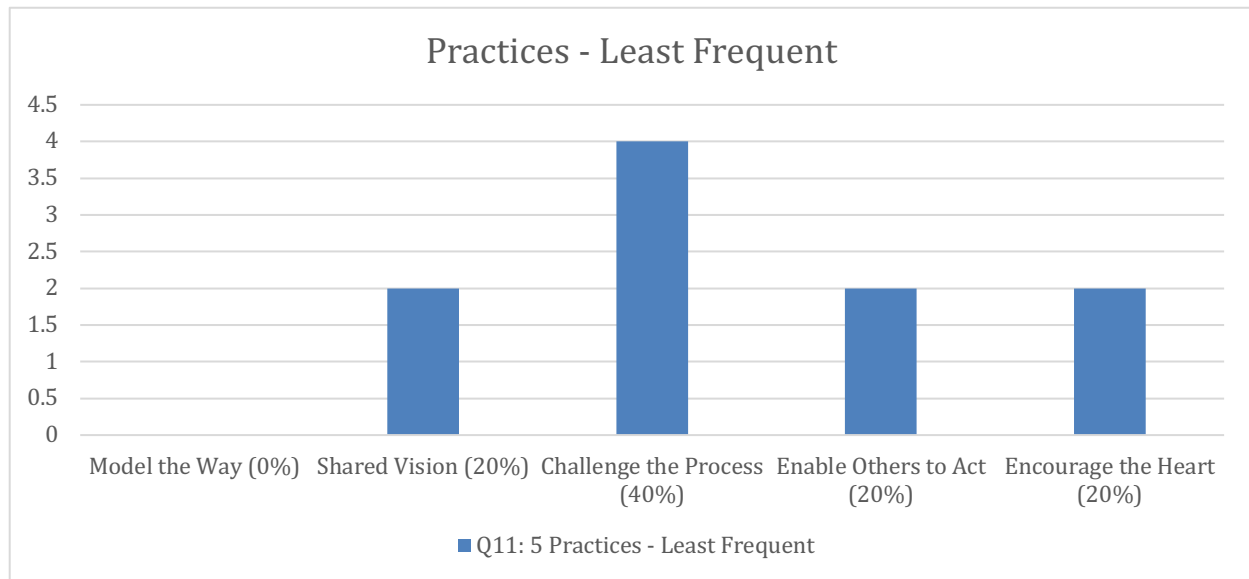
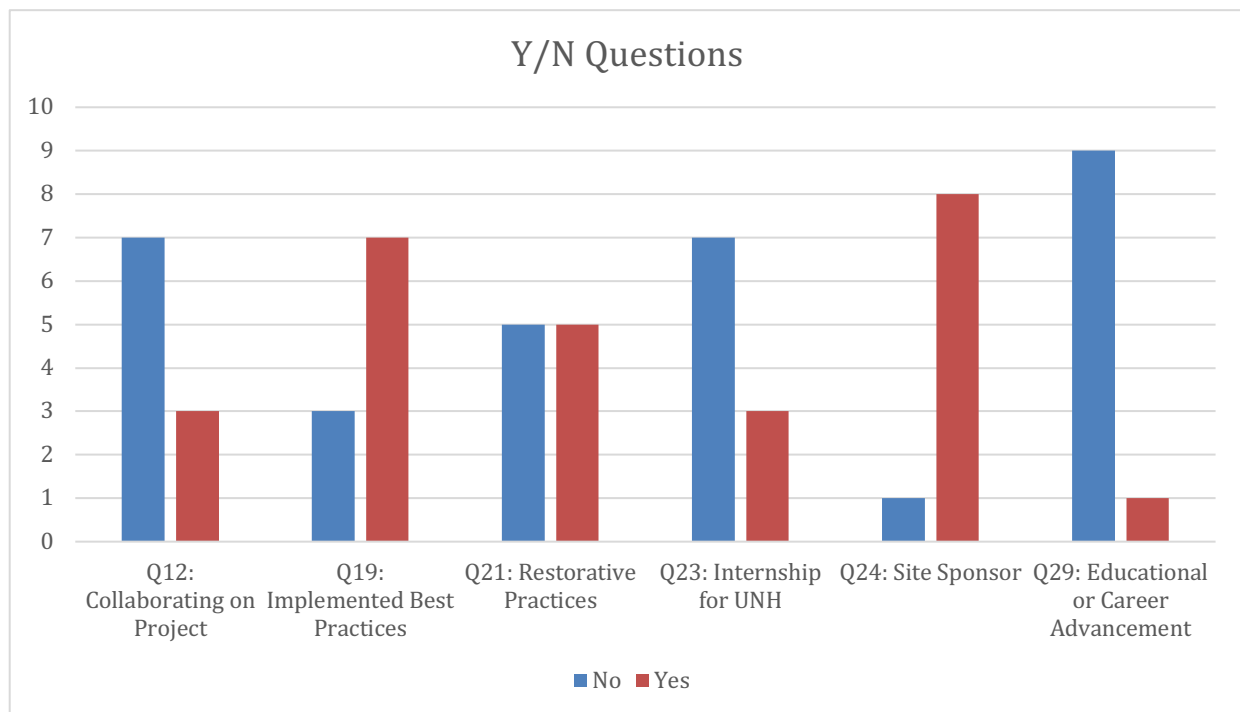


Figure 4: Follow-up Survey Multiple Response Item 11

Yes/No responses. Figure 5 depicts the results of the survey items that were presented in a dichotomous “yes/no” response format. For these items, respondents were asked a question that could be answered as “yes” or “no” and were prompted to select the appropriate response.

Figure 5: Follow-up Survey Yes/No Responses

Question 12 asked, “Are you presently collaborating on a special project on youth justice/reform; e.g., legislation, policy, procedure, or practice change?” As indicated, seven participants responded “no” and three survey participants responded “yes.” Question 19 asked respondents if they have implemented best practices in their agency, network, or advocacy work. As indicated, three participants responded “no” and seven responded “yes.” Question 21 asked participants if they have engaged in restorative practices (listening circle) in their agency, network, or advocacy work. As indicated, five participants responded “no” and five participants responded “yes.” Question 23 asked whether the participant’s agency/organization currently has internship or Community Federal Work Study opportunities for University of New Haven students. Seven participants responded “no” and three responded “yes.” Question 24 stated, “if no to question 23 (some participants answered this question even if they answered “yes” to question 23), would you be interested or willing to consider becoming a site sponsor for internships?” One participant responded “no” and eight responded “yes.” Finally, Question 29 asked, “have you successfully attained an educational or career advancement within the past year?” As indicated, nine participants responded “no” and one participant responded “yes.”

In-Class Direct Observation

In-class direct observation of the 2018 cohort was used to gather qualitative data to support the evaluation of the leadership development program. This was an important component because it: 1) familiarized the lead author with the program, providing deeper understanding; 2) allowed for observation of program operations that are not easily quantifiable (e.g., networking); and 3) helped explain some phenomena that were revealed in the quantitative analysis. Based on the observational data, there were certain themes identified and discussed below.

Place. Except for the TRUE unit visit, the program was hosted at one of two satellite campuses of the University of New Haven (Saw Mill Campus in West Haven, CT, and Orange Campus in Orange, CT). The Saw Mill Campus provided the program with more of a class-room-like setting, as the program primarily inhabited two classrooms. The program was granted access to the cafeteria as well. During small group activities, the groups would spread out between the two classrooms and the cafeteria.

In contrast, the Orange Campus facility, which once served as headquarters for a local business, provided the program with a conference-like setting. Cohort members often discussed the attractiveness of this facility. Here, the program had access to the executive suite, which consisted of two conference rooms and an entryway/lobby, another conference room, a classroom, an atrium area, and cafeteria.

Networking. As mentioned, direct observations allowed the lead author to observe phenomena that may not be quantified easily. An example of this was the observation of real-time and unstructured networking, such as networking occurring in the lunch line or during break. During lunch and outside of the program curriculum, it often was observed that participants would engage in both personal and professional discussions. These discussions

frequently included potential help that could be secured in working with youth who were being served by the juvenile justice system.

Another great example of the TYJI leadership development program providing an opportunity for networking was found through conversation with one of the participants during a group activity. Due to professional reasons, one of the participants (a police officer) needed to withdraw from the program. However, despite the participant's early withdrawal, networking still was occurring with others in the cohort. One of the current participants (another police officer, who did not know the other participant prior to the program) stated that they were still in contact with the participant who had withdrawn.

There also were numerous occasions in which participants either hinted at or spoke directly about a group culture that was formed among the cohort members. Participants stated they felt like everyone came to the program with their own ideas about how they should act, based on their professional titles and roles, but now they have their own group culture in which they feel they can "open up" and discuss concepts and issues freely.

Engagement. Overall, the 2018 participants were engaged throughout the leadership development program. As can be expected, there were some participants who seemed more willing to speak than others. However, every participant spoke at least once during every session. Participants often asked questions or added to the conversation, which suggests both engagement and interest in the subject matter. Some participants engage more often with program coordinators and/or facilitators, in conversations pertaining to program topics. Participants also actively engage during small group discussions on specific topics.

Facilitation. In the first two cohorts, there was one lead facilitator and one supporting facilitator (the Leadership Program Coordinator, or LPC). Prior to the start of the third (2018) cohort, the main facilitator resigned from TYJI. Therefore, an interim facilitator was utilized, with the same supporting facilitator, until a new permanent facilitator was hired. During the third session, a new permanent facilitator was brought on-board. The interim facilitator then left the leadership program, meaning the only facilitator that remained constant among the three cohorts was the supporting facilitator.

This lack of consistency seemingly had an impact on the third cohort participants. Though they were understanding of the circumstances, during a small-group interaction, participants opened-up about how they felt they were in a "transition group" and that they were "along for the ride" during the facilitator transitions. This appeared to be a minor issue, however, as the new main facilitator was in place for the majority of the third cohort's program and is expected to be the main facilitator for future cohorts.

Speakers. Outside speakers were used on multiple occasions. These speakers were experts in a specific and related topic or field (e.g., juvenile justice reform, juvenile justice research, youth development). Through observation, the lead author assessed the speakers on the information they provided and their engagement with participants. Overall, the participants

expressed interest in the subjects covered and satisfaction with the speakers themselves. All speakers were knowledgeable on the topics they covered and facilitated participation among the cohort members. Participants often asked questions and engaged with the speakers.

Circle. Circles are used deliberately as one of the key elements of the “restorative justice practices” approach, endorsed and promoted by the TYJI and by the current best practices research on juvenile justice reform. The circle process is designed to foster a sense of community, as well as to recognize individual perceptions and values that influence participants’ work and their collaboration activity within the field. Circles are an essential part of the leadership program design within this unique program focused on youth justice reform.

Participants experienced two “circle” activities per session: once at the start of the day, after a brief introduction, and at the end of the day, as a concluding activity. The lead author was not permitted to observe these circles. Therefore, to understand circle activities, the researcher relied on: 1) the comments on some session-by-session evaluations, in which participants added a comment on circle; 2) discussions with participants; and 3) observing participants after circle.

Overall, participants enjoyed the circle activities. Aside from informal networking throughout the day, circle was the place in which participants were able to connect with one-another. The participants would come back from circle either smiling and laughing, or still serious and pondering the concepts discussed, and often still talking about things that each other said during the circle activity. There also were many comments made on the session-by-session evaluations that indicated a great or thought-provoking circle activity occurred that day.

During another group activity, however, a few participants expressed some negative experiences with circle. While expressing their awareness that circle is typically a good activity, they pointed to some concerns. One participant indicated difficulty with personally opening up during circle, and that this was an issue because circle works best when everyone is completely open with their feelings and thoughts. Other participants reported they felt that the structure of the circles could be improved. Specifically, some participants were concerned with how people would speak out of turn in the circle, or that they felt like the questions were “too heavy” to either start or end the day. “Too heavy” may have meant either too serious of a question or a question that required too long of a response per participant.

TRUE Unit visit. The TRUE unit exists within the Cheshire Correctional facility in Connecticut. It is a unit dedicated to the TRUE program, which connects males serving life sentences with youths (ages 18-25) serving shorter sentences in a mentor/mentee relationship. TYJI leadership development cohorts go on a field trip to visit the unit to observe the program, talk with the TRUE unit members, get a tour of the facility (from the actual TRUE unit members), and to talk to available correction personnel.

Based on the lead author's observations, the trip appears to make an impact on those that attend. It allows the participants to see and interact with the TRUE unit members in a setting different from which they are accustomed. They are "used to seeing these kids out on the street, acting tough," but in the TRUE unit, they can see young people with feelings and a willingness to change. At the unit, participants are shown success stories and hear instances of people opening up through participation. There were a few times during the visit in which TRUE unit members opened up about emotion or a desire for change. There were a number of leadership program participants who indicated they now view the youths they interact with differently, because of the new light in which they saw the TRUE unit members. Specifically, participants feel like they are more inclined to look through or past the "front" in which a youth tries acting "hard" or as if they "don't care" about the outcome of their interaction. This makes it possible to have a conversation about how youth actually do care and how their actions or demeanor could change to reflect their true feelings.

"I wish we had done this earlier." This statement, or some variation of it, was heard in session or written frequently on the session-by-session evaluations throughout the program. There were many activities, including the TRUE Unit visit and Capstone brainstorming and feedback, for which the participants would have liked to have completed earlier in the program. For example, one participant wished that they had completed the TRUE Unit visit earlier, because this could have influenced capstone project selection. On the positive side, this indicates participant appreciation, engagement, and interest in program activities and content. Moreover, it would be impossible to provide all contents that were suggested to be earlier in the program, although some consideration could be given to ordering of program content.

Each cohort is different. Finally, as mentioned previously, cohort members felt that participation in the program allowed for creating their own culture within the program. It seems likely that each cohort's culture will be different. For example, on the follow-up survey, respondents from the 2016 cohort were asked about leadership practices they use most and least frequently. During a 2018 program session, the cohort was asked the same question. While the 2016 survey respondents indicated that the practice used least frequently was "challenge the process," the 2018 cohort expressed that "challenge the process" was the practice they used most frequently.

SUMMARY AND CONCLUSIONS

Overall, the TYJI leadership development program appears to be a successful and beneficial program. Answers to the four primary research questions, program strengths and weakness, and implications of the evaluation findings are discussed below.

Answers to the Four Research Questions

Is the intended target population being reached? The evidence indicates the intended target population is being reached. The leadership program is attracting and accepting middle-management participants with moderate amounts of work experience and fairly high

educational levels. Females and males are not accepted in equal applicant proportions, with a greater number of female applicants and enrollees, but with an apparent desire to have relatively equal representation based on gender. Program applicants and participants are generally highly experienced and many from both categories have participated previously in other leadership programs.

Are the intended services being provided? As indicated in the session-by-session survey, there is an overall high and consistent satisfaction level with the program being provided. With a few exceptions, these findings apply across both sessions and cohorts. Direct observation also revealed subject matter is being presented to the participants as planned. Participants are engaged in the program and the activities provided.

Are the intended personnel and stakeholders in place and engaged? As mentioned, there was a shifting of personnel near the beginning of the third cohort. There is now a permanent main facilitator in place. Survey findings indicated participant satisfaction with the facilitators and speakers involved in the program. As indicated via the session-by-session assessments and direct observations, it appears appropriate participants are being placed in the program, and they are being trained by skilled and engaged personnel.

Are the anticipated outcomes being achieved? Overall, the anticipated outcomes are being achieved. As indicated by the pre/post survey and follow-up survey, the program is increasing awareness and helping to develop skills and knowledge of best practices in juvenile justice, reform, and organizational and community change. The program aims to be a resource to the organizations, communities, and systems serving youth, through a network of dedicated leaders. As revealed through the surveys and direct observations, a strong network of professionals is being formed through the existence and completion of this program.

Program Strengths

Based on the demographic variables available via the application data, the program is recruiting and selecting appropriate participants from the specified target population across the state of Connecticut. The session-by-session survey data suggest overall high satisfaction levels with regard to information, speakers, and activities delivered during each session. Participant satisfaction and engagement also was observed qualitatively through the direct in-class observations, along with qualitative comments provided on the session-by-session surveys. In general, great appreciation and interest in the topics were expressed by program participants.

In addition, the results of the pre/post survey analysis indicated the program is meeting its desired outcomes associated with most of the survey items. A large majority of the pre- and post-test scores showed a change that was in the desired and beneficial direction, with many of these changes being statistically significant. Furthermore, most of the responses to the follow-up survey were favorable. The highest scores were for Question 28 (the likelihood to seek further education or professional development leading to career advancement) and

Question 1 (how respondents define themselves now as an impactful leader). This indicates that program participants are highly motivated to better themselves professionally, and, perhaps most importantly, completion of the leadership program is associated with participants viewing themselves as impactful leaders.

Program Limitations

The study also uncovered a few limitations or weaknesses with the program. Males were under-represented in application and enrollment figures, leading to a higher percentage of male applicants being selected for the program. This may or may not be a program limitation, but should be considered and discussed by leadership program personnel.

The pre/post survey analysis generally indicated program success. There were a few areas lacking, however, meaning that some items produced pre- and post-test results that were not in the desired or beneficial direction. For example, Item 3 (attend youth focused committees) for both cohorts produced a lower score from pre- to post-test assessment. It may be, though, that the participants' availability for attending youth focused committees was reduced, perhaps due to new work roles or participation in the program itself (because time spent fulfilling program obligations could have reduced time available for attending youth-focused committees).

There were several other items that did not produce desired changes, but the pre- and post-test differences were small and insignificant. These items included Item 12 (leading improvement in policies for my agency) for both cohorts together and for cohort 2016 by itself, Item 11 (experience in utilizing data to inform decision) for the 2016 cohort, Item 18 (member of coalition or organization for change) for the 2016 cohort, and Item 25 (prepared to lead change in YJ reform) for the 2016 cohort. It should be noted that these insignificant changes were limited primarily to the 2016 cohort.

The follow-up survey analysis indicated that the lowest scores were for Question 15 (the degree to which respondents have engaged in action steps or implementation of their capstone). This suggests that program staff and leadership should examine how to better prepare participants for following through and implementing their capstone project. Unfortunately, though, in some cases the inability to complete a capstone may be out of the participants' hands. In some qualitative responses, participants mentioned that they were not supported by their agency or supervisor in implementing their capstones. One follow-up survey respondent reported that they are not able to complete their capstone because they are working with a group, and the group is unable to spend enough time together on the project. In contrast, however, the direct in-class observations revealed that none of the most recent participants worked in groups for their capstone projects, as previous cohorts had done. These recent program participants viewed the lack of group capstone work as a limitation, while at least one earlier cohort member felt they could not follow through with their capstone project due to a lack of time spent with the capstone group. Another important note is that only 10 participants responded to the follow-up survey. Therefore, it is possible

that more capstone projects were completed than was reported through the follow-up survey.

Research Limitations

Finally, there were several research limitations associated with this evaluation. First, the study relied primarily on participant self-report data, which leads to questions about the reliability and validity of the data. In the future, collecting survey or interview data from agency supervisors or coworkers could help address this limitation (Packard & Jones, 2015). Second, there were limited outcome data available, coming primarily from the pre/post surveys and 18-month follow-up survey. Again, future evaluations of the TYJI leadership program may want to consider collecting data from other sources (e.g., supervisor and/or coworker assessments).

Third, this evaluation did not employ a control or comparison group, which would better establish cause and effect. Providing the pre/post survey and 18-month follow-up survey to a control or comparison group, along with the study group, could determine more clearly the true effect of the program on its participants. Finally, many of the analyses in this study utilized a relatively small sample size. This was especially true for the follow-up survey (N = 10). Ongoing data collection from future cohorts, including the follow-up survey, will increase the overall sample size, but techniques for ensuring a better response rate for upcoming survey distributions should be considered (e.g., more stringent policies on attendance and survey participation, collecting data at cohort networking events, and/or incentivizing the completion of surveys).

Despite the limitations of our research, we believe our mixed-methods approach produced valuable information about the evaluated leadership development program and solid evidence regarding its impact. In particular, the combination of quantitative and qualitative data provide consistent support in favor of the program and lesson concerns about the possible shortcomings of individual research methods employed (e.g., the potential bias associated with direct observation of the program sessions).

Implications

Within the modern evidence-based movement, effective juvenile justice leaders will continue to be expected to communicate a knowledge and understanding of evidence-based policies, programs, and practices; balance strong interpersonal skills with the ability to motivate staff toward higher performance; assess and collaboratively shape organizational culture; engage in strategic planning and performance measurement; and facilitate organizational capacity-building and sustainability (Howell et al., 2014; Lipsey, Howell, Kelly, Chapman, & Carver, 2010; Myers, 2013b; Seigle, Walsh, & Weber, 2014). Leadership development programs hold promise for shaping skills and abilities in these areas. Based on the findings of this study and previous research, several recommendations emerge for developing, implementing, and evaluating leadership programs in juvenile justice.

To begin, further attention should be given to specifying and measuring intended program and participant outcomes (Black & Earnest, 2009; Edwards & Turnbull, 2013; Packard & Jones, 2015). In program marketing materials and the curriculum, a clearly defined set of expectations regarding program and participant outcomes should be provided. Participants should be made aware of what is expected of them by defining successful completion of the program, skills and competencies to be gained or enhanced, and anticipated impact within the field of juvenile justice and/or their agencies (e.g., through the capstone project and/or involvement in juvenile justice leadership, reform, joining organizations). To provide guidance with this process, Black and Earnest (2009) provide a useful model for leadership development and a measurement tool for program evaluation, which includes outcomes at the individual, organizational, and community levels.

Relatedly, the current evaluation relied primarily on the pre/post-test survey and 18-month alumni follow-up survey (completed by the first [2016] cohort of alumni) to measure and assess whether intended outcomes were being achieved. Future evaluations should consider utilizing other available data sources to measure intended outcomes. Some potential sources of data include:

- Formal assessment of presentations of learning
- Formal assessment of capstone completion and follow-through
- Survey and/or interviews with participant supervisors and/or coworkers (see Packard & Jones, 2015).
- Six-month and/or yearly follow-up on capstone progress

Next, the alumni follow-up survey indicated that the lowest scores were for item 15 (the degree to which respondents have engaged in action steps or implementation of their capstone). Program developers and coordinators should consider ways to better prepare participants for developing and carrying out the action steps for their capstone. Future evaluators may also want to consider ways to understand and measure participant engagement in action steps and implementing their capstone. As indicated in some qualitative responses on the alumni follow-up survey, participants may be blocked or hindered in their ability to implement their capstone, due to reasons beyond their control. These types of restrictions may also limit the overall effectiveness of leadership development programs in producing desirable change and reform.

Finally, previous research indicates the leadership development process involves the acquisition and enhancement of a variety of skills and is shaped by factors such as personality, relationships with others, and deliberate efforts to collaborate, learn, reflect, and improve (Avolio et al., 2009; Collins & Holton, 2004; Day et al., 2014; McCauley & Hughes-James, 1994). Moreover, while a variety of life and work experiences provide valuable training for leadership, structured experiential approaches and theory-driven programs with articulated and measureable outcomes provide the best opportunity for leadership enhancement (Black & Earnest, 2009; Brungardt, 1996). When effective programs include or are combined with ongoing coaching, mentoring, networking, and opportunities to learn and grow professionally, significant behavioral change and lasting leadership effectiveness is more likely (Day, 2001). The current study examined a juvenile

justice leadership development program that included many of these aspects, such as relationship building, collaboration, reflection, networking, experiential learning, and opportunities to grow professionally. It is possible that a greater emphasis on continued coaching, mentoring, and measureable longer-term outcomes could produce even better results. We hope this study is useful to both juvenile justice professionals and researchers who are pursuing and assessing these types of goals and outcomes.

REFERENCES

- Avolio, B. J., Hannah, S., Reichard, R., Chan, A., & Walumbwa, F. (2009). A meta-analytic review of leadership impact research: Experimental and quasi-experimental studies. *The Leadership Quarterly*, 20(5), 764-784.
- Black, A. M., & Earnest, G. W. (2009). Measuring the outcomes of leadership development programs. *Journal of Leadership & Organizational Studies*, 16(2), 184-196.
- Brungardt, C. (1996). The making of leaders: A review of the research in leadership development and education. *The Journal of Leadership Studies*, 3(3), 81-95.
- Carter, H. S., & Rudd, R. D. (2000). Evaluation of the Florida leadership program for agriculture and natural resources. *Journal of Southern Agricultural Education Research*, 50(1), 193-199.
- Collins, D., & Holton, E. (2004). The effectiveness of managerial leadership development programs: A meta-analysis of studies from 1982 to 2001. *Human Resource Development Quarterly*, 15(2), 217-248.
- Day, D. V. (2001). Leadership development: A review in context. *Leadership Quarterly*, 11(4), 581-613.
- 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. *The Leadership Quarterly*, 25, 63-82.
- Edwards, G., & Turnbull, S. (2013). Special issue on new paradigms in evaluating leadership development. *Advances in Developing Human Resources*, 15(1), 3-9.
- Feld, B. C. (2017). *The evolution of the juvenile court: Race, politics, and the criminalizing of juvenile justice*. New York: New York University Press.
- Field, A. (2009). *Discovering Statistics Using IBM SPSS Statistics* (4th ed.). Washington, DC: SAGE Publications Ltd.
- Howell, J. C., Lipsey, M. W., & Wilson, J. J. (2014). *A handbook for evidence-based juvenile justice systems*. Lanham, MD: Lexington Books.
- Lipsey, M. W., Howell, J. C., Kelly, M. R., Chapman, G., & Carver, D. (2010). *Improving the effectiveness of juvenile justice programs: A new perspective on evidence-based practice*. Washington, DC: Center for Juvenile Justice Reform.
- Ma, L., Pfrommer, D., & Cooper, D. (2018). A journey through Connecticut's juvenile justice system reform. *ACJS Today*, 42(3), 23-29.
- McCauley, C. D., & Hughes-James, M. W. (1994). An evaluation of the outcomes of a leadership development program. *Center for Creative Leadership*: Greensboro, North Carolina.
- Myers, D. L. (2005). *Boys among men: Trying and sentencing juveniles as adults*. Westport, CT: Praeger Publishers.
- Myers, D. L. (2013a). Accountability and evidence-based approaches: Theory and research for juvenile justice. *Criminal Justice Studies*, 26(2), 197-212.

- Myers, D. L. (2013b). *Becoming an evidence-based organization: Demonstrating leadership and organizational growth*. Johns Creek, GA: Joyfields Institute for Professional Development.
- Packard, T., & Jones, L. (2015). An outcomes evaluation of a leadership development initiative. *Journal of Management Development*, 34(2), 153-168.
- Seigle, E., Walsh, N., & Weber, J (2014). *Core principles for reducing recidivism and improving other outcomes for youth in the juvenile justice system*. New York: Council of State Governments Justice Center.
- Tow Youth Justice Institute Newsletter (Spring 2016). CT leads nation by reducing juvenile commitment by 75%. West Haven, CT: Tow Youth Justice Institute
- Tow Youth Justice Institute Newsletter (Fall 2016). Update on transforming youth justice: A leadership development program. West Haven, CT: Tow Youth Justice Institute
- Tow Youth Justice Institute Newsletter (March 2017). Celebrating the first cohort: Transforming youth justice: Building reformers. West Haven, CT: Tow Youth Justice Institute
- W. K. Kellogg Foundation. (2001). *Leadership scan. The legacy of the ag leadership development program: Rich heritage cultivates future opportunities*. Battle Creek, MI: Author.