Promoting Evaluation in Youth Character Development through Enhanced Evaluation Capacity Building: Empirical Findings from the PACE Project

Lisa M. Chauveron, Jennifer Brown Urban, Satabdi Samtani, Milira Cox, Leslie Moorman, Monica Hargraves, Jane Buckley, and *Miriam R. Linver*

Abstract

In this chapter, we present empirical findings from a mixed-methods study of an evaluation capacity building (ECB) initiative called the Partnerships for Advancing Character Program Evaluation (PACE) Project that applied the Relational Systems Evaluation (RSE) approach. We focused especially on how participation in this partnership-based ECB program is associated with changes in attitudes and behaviors related to evaluation capacity and evaluative thinking. We used a repeated measures design with semi-structured interviews and surveys to examine two research questions: (1) Do program professionals participating in PACE adopt evaluative thinking behaviors? And (2) Do program professionals participating in PACE demonstrate increased evaluation capacity, including improved attitudes about evaluation? We found that PACE successfully increased program professional's behaviors, attitudes, and capacity; in particular, participants overwhelmingly noted that the emphasis on evaluative thinking changed the way they think about and approach evaluation, such that evaluation is now more well-integrated into their practice.

Chauveron, L. M., Urban, J. B., Samtani, S., et al. (2021). Promoting evaluation in youth character development through enhanced Evaluation Capacity Building: Empirical findings from the PACE Project. *New Directions for Evaluation*, 2021, 79– 95. https://doi.org/10.1002/ev.20447

Program practitioners and evaluators each play critical, complementary roles in the evaluation of programs. The former drive practices associated with strong program outcomes, including planning evaluations, partnering with external evaluators, and using findings within organizations (e.g., Garcia-Iriarte et al., 2011; Labin, 2014), while the latter offer crucial technical expertise and experience with a diverse array of programs. Recognizing the importance of integrating practitioner and evaluator expertise, the Partnerships for Advancing Character Program Evaluation (PACE) project used an innovative design to train both evaluators and program professionals together in a Relational Systems Evaluation (RSE) approach to evaluation planning. The focus of the PACE Project was on youth character development programs, an arena where there is a recognized need for more widespread high-quality evaluation (Roth, Brooks-Gunn, Murray, & Foster, 1998). In addition to the joint training, the PACE Project adopted an evaluation partnership framework in which evaluators were partnered with program professionals to put RSE tools into practice. The overarching goal of the PACE Project was to increase the practice and quality of program evaluation in the realm of youth character development programs by building evaluation capacity among all participants. This paper presents the results of this initiative, focusing on outcomes for evaluative thinking (ET) and evaluation capacity in participating program professionals.

Evaluation Capacity Building and Relational Systems Evaluation

Evaluation capacity building (ECB) addresses the processes and practices at the individual and organizational levels necessary for sustained high-quality evaluation. For individuals, the goals of ECB are to strengthen evaluation skills, attitudes, and knowledge. It is widely recognized in the literature on ECB that individual action alone is not sufficient to ensure sustained evaluation practices throughout an organization (Labin et al., 2012). A number of ECB models emphasize the importance of learning transfer from the individual to

Chauveron, L. M., Urban, J. B., Samtani, S., et al. (2021). Promoting evaluation in youth character development through enhanced Evaluation Capacity Building: Empirical findings from the PACE Project. *New Directions for Evaluation*, 2021, 79– 95. https://doi.org/10.1002/ev.20447

the organization, supported by leadership, organizational culture, communication, systems and structures to ensure the application of the acquired capacity to the work context and into sustainable evaluation practice. ECB unleashes the motivation, knowledge, and skills for individual and organizational commitment to regular evaluation and insight integration (e.g., Labin, 2014; Labin et al., 2012; Nelson et al., 2018; Suarez-Balcazar & Taylor-Ritzler, 2014; Taylor-Ritzler et al., 2013, Wandersman, 2014). The PACE training program was not only designed to strengthen skills and ET and to increase positive attitudes toward evaluation within individual participants, but also included purposeful activities such the creation of an "Evaluative Thinking Learning-to-Action Plan" designed to facilitate the integration of these practices into the home organizations of the PACE program professionals. This tool was developed based on research that has shown that the use of evaluation to drive organizational and programmatic decisions comes, in part, from mainstreaming the promotion of evaluation within organizations by senior leadership as well as staff at various levels (Suarez-Balcazar & Taylor-Ritzler, 2014). Further, in alignment with other research on ECB (Preskill & Boyle, 2008), PACE activities included interactive, multi-faceted teaching and learning strategies to engage participants. The foundational approach to evaluation that formed the basis for the experiential learning approach in PACE was RSE, including evaluative thinking, and the Systems Evaluation Protocol (Chapter 3 of this volume; Buckley et al., 2015; Trochim et al., 2016; Urban et al., 2011; Urban et al., 2014; Urban & Trochim, 2009), described in brief below.

Relational Systems Evaluation and the Systems Evaluation Protocol

RSE is a theoretically grounded framework that situates programs within an evolutionary and ecological context and works through evaluator-practitioner partnerships to integrate diverse sources of expertise and build evaluation capacity (Chapters 1, 2, and 4 of

this volume; Urban et al., 2014). RSE is operationalized using the Systems Evaluation Protocol (SEP), which directs evaluators and program practitioners through the planning, implementation, and use of evaluations (Chapter 3 of this volume; Trochim et al., 2016). The SEP systematically helps identify stakeholders within the program's broader systemic context, including input from individuals and groups representing diverse roles and social, economic, community-based perspectives (see Chauveron, Samtani, Groner, Linver & Urban, in press). It also connects stakeholder priorities to a programmatic theory of change visualized in a pathway model, an important SEP tool whose development builds on SEP steps that determine program boundaries and program and evaluation lifecycle phases. The PACE Project used RSE as the foundational framework for evaluation planning and built evaluation capacity using an experiential learning approach that put RSE into practice.

Evaluative Thinking

RSE also promotes evaluative thinking, a central component of evaluation capacity (Bennett & Jessani, 2011; Taylor-Ritzler et al., 2013). In PACE, evaluative thinking was defined as:

> critical thinking applied in the context of evaluation, motivated by an attitude of inquisitiveness and a belief in the value of evidence, that involves identifying assumptions, posing thoughtful questions, pursuing deeper understanding through reflection and perspective taking, and informing decisions in preparation for action. (Buckley, Archibald, Hargraves, & Trochim, 2015, p. 378)

ET is critical to evaluation capacity because it integrates the same skills that characterize good evaluation throughout an organization's work practices (Baker & Bruner, 2012); it is characterized by "a willingness to do reality testing, to ask the question: how do we know what we think we know? … It's an analytical way of thinking that infuses

everything that goes on" (Patton, 2005, para. 10). More specifically, the ET skills and practices introduced as part of PACE offered practical and effective ways to surface assumptions, pose the optimal evaluation questions, and weigh important considerations before making key decisions (Archibald et al., 2018; Buckley et al., 2015).

The approach to facilitating ET adopted by PACE focused on establishing the habits and conditions necessary to sustain high-quality evaluation and program development. PACE participants were offered tools, strategies, and activities that could be used (with large or small groups) as daily routines or day-long retreats to promote ET at the individual, program team, and organization levels. All the ET skills addressed in PACE overlap and integrate with the SEP steps used to structure the workshops. For example, developing program pathway models is an effective way to surface program assumptions (Buckley et al., 2015). In turn, learning to identify assumptions can improve the quality of a program's pathway model, ultimately improving the illustration of connections between activities and outcomes. PACE promoted ET as an important evaluation capacity skill in its own right, while also offering the opportunity to directly apply and practice that skill in the context of developing program descriptions and evaluation plans. Taken together, the insights presented above on ECB and ET, framed by RSE as an overarching approach and put into practice using the SEP, constituted the PACE Project, described in further detail below. After describing the PACE Project, the remainder of this chapter presents the processes and results of a mixed-methods study of PACE's outcomes.

The PACE Project

The PACE Project is the first large-scale ECB initiative (that we are aware of) to promote evaluation partnerships between program professionals (PPs) and evaluators (or evaluation capacity builders [ECBers]). An innovative national project, PACE was designed

Chauveron, L. M., Urban, J. B., Samtani, S., et al. (2021). Promoting evaluation in youth character development through enhanced Evaluation Capacity Building: Empirical findings from the PACE Project. *New Directions for Evaluation*, 2021, 79– 95. https://doi.org/10.1002/ev.20447

to build shared ownership of the evaluation process for both ECBers and PPs to exchange expertise and develop knowledge, skills, and practices that support high-quality programs. PACE aimed to address the foundational conditions for motivation, communication, and careful thinking that drive sustained evaluation work throughout the life of a program. Over 15 months of a 3-year study, PACE trainings addressed the full cycle of evaluation from planning to utilization but paid particular attention to the often under-attended evaluation planning stage. PACE focused on tools and approaches that establish a strong evaluation plan, including modeling the theory of change that drives the program, identifying specific evaluation questions, and mapping how the program fits within its environmental context.

PACE Structure and Participants

A national request for proposals (RFP) process yielded applications from 30 youth character development (CD) programs, of which 16 were selected from 12 states. Each sent two PPs except one (*N*=31; one organization only had one representative). Selected programs demonstrated "readiness," defined as having a clear set of CD program activities that focused on developing character or virtues, some defined CD outcomes, and a desire for an evaluation partnership. Participating programs received a stipend, and PPs' travel costs were covered. A similar national recruitment process yielded applications from 32 professional evaluators, eight of whom were selected as ECBers for PACE. This group, drawn from seven states, had expertise in evaluation and research methods, an interest in PACE concepts, a commitment to evaluation partnerships, and the desire to connect to other ECBers. All received a stipend and travel costs were covered.

The 16 programs were split into two cohorts of eight programs, and each evaluator was partnered with one program team from Cohort 1, whose in-person training began in October 2016, and one team from Cohort 2 whose in-person training began in January 2017.

Partnership matches were made by the PACE Lead Facilitators and the project leader after the first two days of the inaugural in-person workshop, based on a desire to align evaluator expertise, experience, and interests with program type and context, and program team experience and interests.

The resulting group of 16 ECBer-PP partnerships participated in four types of activities over a period of 15 months from 2016-2018 (see Table 6.1): (1) In-person workshops and the PACE Culminating Conference, (2) Webinars, (3) Evaluation partnership work, and (4) Consultation and coaching with a PACE Lead Facilitator.

In-person Workshops and the PACE Culminating Conference

Two multi-day, in-person workshops (WS1 and WS2) covered the core program content and were hosted twice, once for each cohort (see Table 6.1). All workshops were led by the project leader, two Lead Facilitators, and a Facilitation Team. WS2 for Cohort 2 was held virtually, to reduce travel costs. The workshops used an active-learning approach to ECB through structured group work, peer review, discussion, role-plays, case study, and brainstorming. All participated in a two-day in-person Culminating Conference where PPs gave talks and presented posters that explained how their program applied PACE skills and concepts. The Culminating Conference agenda also included time for networking, both within and across cohorts and with ECBers and potential funders, a keynote address on social-emotional learning and character development, and a panel discussion of fellow PACE participants' experiences during PACE.

<<INSERT TABLE 6.1 HERE>>

Webinars

Three webinars supplemented workshops: One shared CD concepts and definitions; two others taught concepts of person-centered analysis for use in CD programs. Webinars were recorded and made available for all PACE participants at any time.

Evaluation Partnership Work

Building on the presentations and hands-on activities of the workshops, PPs and ECBers partnered to complete key products following the SEP: a written program description, stakeholder map, program pathway model, evaluation purpose statement, and evaluation questions. These were combined in a Program Evaluation Profile (PEP). A subset of 13 programs were provided an additional stipend and completed an optional full evaluation plan with evaluation design, sample, measures, analysis plan, and timeline. Partnership work was mainly remote through regular communication and online meetings.

Consultation and Coaching with PACE Lead Facilitator

Each evaluation partnership was supported by one of the two Lead Facilitators who guided and supported the partnership. In addition to phone and email communications as needed, quarterly meetings provided opportunities to check in on progress, answer questions, provide structured feedback, and collect data.

The Current Study

A mixed-methods investigation examined two research questions (RQs): (1) Do program professionals participating in PACE adopt evaluative thinking behaviors? And (2) Do program professionals participating in PACE demonstrate increased evaluation capacity, including improved attitudes about evaluation? We hypothesized that for each question, participants would show positive change. A repeated measures design included data collected from surveys and semi-structured interviews. An internal consortium from Montclair State University's RYTE Institute, Cornell University, and Virginia Tech led the study.

Methods

Sample

There were 16 programs and 31 PPs in the original group of PACE participants; program funding cuts, however, led to the withdrawal of one of the programs. Staff turnover and budget cuts at others reduced the participant group further. The sample for this study includes the 26 PPs who completed all PACE Project activities. We collected demographic data for 18 of the 26 PPs. Most participants identified as women (83.3%; 16.7% men, none identified in another way) and were between 25 and 74 years of age (M_{range}=35-44 years). Participants could describe their identity using as many options as were resonant; the final sample identified as Hispanic or Latinx (9.1%), Black or African American (11.8%), White (72.2%), and Multiracial (23.5%). Respondents had either a college (53.0%) or graduate (47.0%) degree, and most had worked for their organization one to three years (41.2%) or four to six years (35.3%); 17.6% had been at their organization for seven or more years. About a third each were in their positions for one to three years (35.2%) or four to six years (35.2%) and seven or more years (29%). While one PP entered PACE with no evaluation knowledge, the rest had very limited (68.8%) or somewhat strong (31.2%) knowledge.

Design

Semi-structured interviews were used to assess ET. A phone-based 45-60 minute interview was done before WS1 (Wave 1, Fall 2016), after the Culminating Conference (Wave 2, Fall 2017), and one year after PACE concluded (Wave 3, Fall 2018); 26 sets of Wave 1/Wave 2 interviews were analyzed through comparative analysis, described below. Twenty PPs also completed Wave 3 interviews, which were also included in analyses.

The survey used a retrospective pretest-posttest design. The design, administered at a single time point upon program completion, directs respondents to think back to their

perception of each item before the program and again after program participation. This approach reduces threats to internal validity produced by self-assessments (Howard, 1980) while also capturing potential "response shift bias," changes in respondents' internalized conceptualization of items from the pre- to posttest administration due to newfound understanding of the concept (Howard et al., 1979). Some suggest retrospective tools may better capture self-assessment changes than traditional pretest-posttests (Geldhof et al., 2018) and use it as supplemental to (Howard et al., 1979) or in lieu of (Allen & Nimon, 2007) a traditional pretest-posttest; we chose a supplemental approach to capture both traditional and retrospective views. PPs completed surveys at Wave 1 and at Wave 2. Survey data was analyzed for all PPs and did not include Cohort comparisons due to sample size.

Measures

Qualitative and quantitative data were collected and analyzed separately using the measures detailed below; results and insights were integrated for interpretation.

Evaluative Thinking Behaviors and Attitudes. The Evaluative Thinking Inventory (ETI) is a validated survey measure of ET behaviors and attitudes in individuals (McIntosh, Buckley, & Archibald, 2020). The original version of the ETI[1] has 14 items scored on a 6-point Likert-type scale (5=very frequently to 0=never) on three scales: *Posing Thoughtful Questions and Seeking Alternative Explanations* (7 items, α=.80) including "I pose questions about assumptions and claims made by others," *Describing and Illustrating Thinking* (2 items, α=.82) including "Diagrams and/or illustrations help me think about ideas," and *Believing and Practicing Evaluation* (5 items, α=.89) including "I am eager to engage in evaluation."

[1] The original version of the ETI included 14 items. A more recent publication includes four additional ETI items.

A semi-structured interview was also used to assess evaluative thinking among Program Professionals. In Waves 1 and 2, questions asked how participants would define or describe ET, how they would explain ET to a layperson, as well as what participants considered to be elements of high-quality evaluation. At Wave 3, interview questions were focused on what actions participants had taken since PACE concluded. Sample questions include: "How has PACE influenced your current thinking about evaluation," "Are there any particular tools, strategies or concepts you learned about during PACE that you have used since the end of PACE," "Have you found yourself advocating for evaluative thinking [and evaluation] more, less, or differently since PACE"? Interviews were audio recorded and transcribed for analysis.

Evaluation Capacity. The 68-item Evaluation Capacity Assessment Instrument (ECAI) measures aspects of individual and organization level evaluation practices (Taylor-Ritzler et al., 2013) in three domains: 1) *Individual* factors (α= .90) that address *Awareness* of the benefits of evaluation (11 items), *Motivation* to conduct evaluation (4 items), and *Competence* (skills and knowledge) to conduct evaluation (14 items); 2) *Organizational factors* (α=.90) that address *Leadership* for evaluation (5 items), supportive *Learning Climate* (9 items), and *Resources* available for evaluation (9 items); and 3) *Evaluation Capacity Outcomes* (α=.94) that address *Mainstreaming* evaluation into workflows (5 items) and *Use* of findings (11 items). Scale scoring for all subscales was on a Likert-type scale of (0=strongly disagree to 3=strongly agree), except *Use* where items included a Likert-type scale with response options 0=not at all to 3=to a very great extent.

Qualitative Analyses

To address the first research question related to the adoption of ET behaviors, a qualitative person-centered approach (see Hershberg & Johnson, 2018) was used whereby

two coders (Team A) compared Wave 1 and Wave 2 interview transcripts. A coding dictionary was developed, and the summary statements were coded for each participant using a consensus coding approach (see Appendix A in Chapter 4 of this volume for a complete discussion of this coding process). Then, a separate team of three coders (Team B) engaged in subcoding (Saldana, 2016) by first selecting a subset of ET related codes from the coding dictionary. Direct quotes coded as related to ET from Wave 3 interviews were also added to the dataset so that each participant had a summary sheet that included ET data from across all waves with available data. Team B coders independently reviewed data for each PP using both a deductive and an inductive approach. First, statements were subcoded according to the three a priori domains of ET: (1) Posing thoughtful questions and seeking alternative explanations, (2) Describing and illustrating thinking, and (3) Believing and practicing evaluation. Coders also allowed for emergent themes to arise, and iteratively reviewed the data as new themes emerged. Each coder also wrote memos after reviewing data for each participant. The three coders then met to discuss and reach consensus on coding decisions. The coders discussed the a priori codes for each participant and discussed any coding discrepancies until consensus was reached. The coders also discussed emergent themes, reviewed the data sources, and reached consensus on coding decisions.

Results

Evaluative Thinking Behaviors (RQ1)

To answer RQ1, whether PPs participating in PACE adopt ET behaviors, analysis of the ETI results were used. Paired samples *t*-tests of 26 Wave 1 and 2 responses to the ETI confirm our hypothesis that PPs adopted more evaluative thinking habits as significant increases appeared on all three scales (See Table 6.2). The biggest effect size was exhibited on *Describing and Illustrating Thinking,* indicating that PPs increased their use of diagrams

and illustrations to communicate thinking and ideas. This was exercised by PP's through the creation of their program pathway models to describe their approach to improving character outcomes. Interestingly, the effect size on *Believing in and practicing evaluation* was only slightly smaller, which shows that PACE participants increased the desire to conduct evaluations and the belief in their importance.

<<INSERT TABLE 6.2 HERE>>

The interviews provide a deeper look at PPs understanding and application of ET. By Wave 3, most PPs (77%) described adopting ET behaviors[2]. These behaviors can be categorized into three major themes that are consistent with how ET is defined (Buckley, Archibald, Hargraves, & Trochim, 2015; McIntosh, Buckley, & Archibald, 2020). Many PPs (50%) described posing thoughtful questions within their organization since participating in PACE. One participant credits PACE with teaching them to "ask questions about 'how'" (Participant 16, Wave 2). This practice appeared to extend to other staff and stakeholders at participants' organizations. For example, one participant describes how she used ET to examine how staff are trained:

> We've also be able to use ET to really look at how we're training our teams and how what they're doing is impacting the young people we serve. And so we've done some deeper scrutiny of our onboarding processes that lead to translating to our staff what it is they're trying to accomplish. (Participant 31, Wave 3)

Questioning assumptions is applied not only as a self-reflective process, but also includes taking multiple perspectives. One participant notes that ET is:

[2] Percentages are calculated based on the total number of participants (N = 26), not the total number of participants who also completed Wave 3 interviews (N = 20).

> ... this idea that nothing ever stays stagnant and we should always be asking questions of our stakeholders, of ourselves, of those that we're serving who are also our stakeholders, but I'm thinking stakeholders like board of directors, and here maybe it's the…community at large or whoever our stakeholders are for a particular program. (Participant 35, Wave 3)

The most common ET theme that emerged after PACE participation is believing and practicing evaluation (58%). This includes discussing evaluation strategies with colleagues, demonstrating an eagerness to engage in evaluation, and working to convince others that evaluation is important. Several PPs talked about how ET was incorporated into their organizational culture. One PP described how she was able to "initiate a culture of ET by helping staff understand how they know what they know. They began to also use our evaluative tools a little bit more consistently for the first time and read those results" (Participant 31, Wave 2). She also described a follow-up project where the Strategic Impact and Evaluation team at her organization was trained in ET and the Systems Evaluation Protocol. Since then, ET has been broadly applied across her organization using "ET and pathway modeling with another program in my line of service, and their goal is to do that with all areas of service... And I know that [program partner] made it her breakthrough goal to do pathways modeling [sic] and sort of this ET training with I think three programs" (Participant 31, Wave 3). Further, PPs have demystified evaluation within their organizations by:

> helping programs embed ET, and to name ET when they may not know that that's what they're doing. And just to also shed some light, kind of in the same concept of ET, but to shed light on the nature that evaluation doesn't have to be a scary and burdensome initiative if it's done in a way that you're already having some

> touchpoints with evaluation, and just to maybe bring those more to light, or to maybe improve upon it, so that the information that is being gathered is more useful to a program. (Participant 25, Wave 3)

Additionally, one PP said that knowing ET equipped her with the

> language and project confidence to talk to senior level and executive-level leadership in my organization, and better share the impact of what good evaluation can do. It has given me the language to share how measuring impact can really benefit the program and the organization as whole. (Participant 25, Wave 3)

PPs explained that adopting ET has also shifted the way their organizations engage stakeholders, asking, "...what is the target audience? ...What's the message that we're trying to get to these stakeholders? Is that message different? ...We really take the time to do more of that now when before we didn't" (Participant, 9, Wave 3). Another participant described how employing ET improved evaluation practice and changed the way they approach client work:

> I think that the most important thing for me is that [my colleague and I] most importantly understand now the importance of what it is that we learned at PACE, which is, first and foremost, the ET. We're always thinking now about why we're doing this. And I don't think we were doing that before. Before it was just like, we want to place clients… so these are the things we need to do. Now, it's like, all right, how do we look at these clients individually, and say this person has these challenges, how are we going to help them get through these challenges? What do we need to put in place to make sure that we do the best to help them to meet these challenges? The way we think about how we work is totally different. (Participant 9, Wave 3)

Slightly less than half of PPs offered examples of describing and illustrating thinking (42%). Pathway models were frequently described as a tool that PPs use to continue

clarifying their own thoughts and communicating about the program with others. Some PPs said pathway modeling was one of the most important things they learned during PACE. One PP commented that they, "have used the pathway model quite a bit in the past year, and that's the biggest thing, and it's almost the biggest takeaway" (Participant 25, Wave 3). PPs commonly described using pathway modeling with other programs as well as encouraging colleagues to use modeling to refine their thinking about programs, "They've now begun to work on their own pathway model, not just the one that [colleague's name] and I made, but they informed the pathway model we created, but now they're doing it from scratch for themselves" (Participant 31, Wave 2).

In addition to the emergence of themes consistent with the definition of ET, two other notable themes emerged. Two PPs described a particularly deep understanding of and commitment to ET, for instance, as "a way of being, a way of thinking" (Participant 23, Wave 2). They explained that using ET has altered their worldview. Another PP describes the profound impact ET has had on her thinking: "evaluative, deep thinking is – is my thing now – and that's not even just – just that's just in my life. That's with my relationships, that's with my family, that's in work, personal and professional" (Participant 9, Wave 2).

Some PPs believe that ET may foster more effective work with external evaluators. In the future, one PP thinks that she will:

> probably be a bit more particular in choosing an evaluator, looking for somebody who has a skill set in ET. We hadn't had very good experience in the past with outside evaluators, so going through this process over the year, I think, taught us a lot about our organization and the type of programs we're running, and the skill set and kind of personality ... Or maybe it is just a skill set that an evaluator needs to bring to the table to help us. (Participant 23, Wave 2)

Her hope is to build a relationship with an external evaluator where PPs' expertise is appreciated. She wants to have:

> the time to sit with the evaluator so that they're willing to work with us, listen to us, work as a team, build the evaluation process together. Understand that all of us have a skill set in the process, and so it's not necessarily that the evaluator is kind of the "smart one," that we all contribute. I think that's something that going forward, we would probably look ... If we needed another outside evaluator, we'd be looking for somebody who can work as a team and can bring that kind of evaluative thinking to the table. (Participant 23, Wave 2)

Evaluation Capacity (RQ2)

To answer RQ2, whether PPs demonstrate increased evaluation capacity, including improved attitudes about evaluation, results of the ECAI were analyzed in two ways. First, an analysis of the traditional pretest-posttest design showed five of the eight ECAI domains demonstrated significant change (see Figure 6.1); they are competence, leadership, learning climate, mainstreaming, and use. Specifically, PPs increased *Individual* capacity by expanding their *Competence* with more skills and knowledge to conduct evaluation. PPs also reported greater *Organization* level capacity through increased *Leadership* support for evaluation and a *Learning Climate* that fosters evaluative practices. Finally, PPs improved their practices around *Evaluation Outcomes* by better *Mainstreaming* evaluation into their workflows and increasing the *Use* of study findings to improve programs.

<<INSERT FIGURE 6.1 HERE>>

Interestingly, the findings were more pronounced in the analysis of the retrospective pretest-posttest design than when examined through the lens of a traditional pretest-posttest model. Using the retrospective design, a significant change was found for all eight ECAI

domains. The retrospective approach was sensitive to two additional *Individual* changes (*Awareness* of the benefits of evaluation and *Motivation* to conduct evaluation), and *Organization* level capacity about necessary *Resources* for evaluations. The biggest changes appeared in individuals' *Competence* and *Motivation*, which suggests that PACE enhanced skills, knowledge, and commitment to evaluation.

We went a step further to see if a relationship exists between thinking evaluatively and the impetus to conduct evaluations; literature suggests this relationship is fostered by enhanced evaluation capacity (e.g., Taylor-Ritzler et al., 2013). We examined the connection between ET attitudes and behaviors measured by the ETI and the Individual factor of the ECAI. Table 6.3 shows that after PACE participation, significant correlations were found between the ETI and ECAI (no such correlations were found before PACE). The ETI subscales of *Believing and Practicing Evaluation* and *Posing Thoughtful Questions and Seeking Alternatives* were significantly correlated with all three ECAI Individual factors of *Awareness, Motivation,* and *Competence*. Therefore, the attention, desire, and ability to faithfully do evaluation is connected with thinking evaluatively. The strongest correlation was between *Motivation* and *Believing and Practicing Evaluation;* thus, a considerable connection exists between the drive to conduct evaluation and individual beliefs and willingness to practice ET.

<<INSERT TABLE 6.3 HERE>>

Discussion

PACE's innovative design simultaneously trained evaluators and program professionals together through an RSE approach to evaluation. PACE paired evaluators and program practitioners in partnerships to build critical knowledge and capacity associated with high-quality evaluation; the findings from the current evaluation ultimately serve as its

"proof-of-concept." Moreover, the PACE approach produced more knowledgeable, better-skilled program professionals, stronger and more equitable connections in the evaluation process, and high-quality evaluation plans (Chauveron et al., in press). Results show that PPs gained knowledge of concepts like ET as well as processes like the SEP to plan high-quality evaluations throughout program lifespans.

Surveys and interviews also showed that PPs adopted ET habits into their approach to evaluation. Not only were there significant improvements in participant ET on all three ETI subscales, but a deeper dive into ETI subscales of *Believing and Practicing Evaluation* and *Posing Thoughtful Questions and Seeking Alternatives* showed they were significantly correlated with all three Individual factors of *Awareness, Motivation,* and *Competence* on the ECAI scale. Moreover, when we assessed the link between the ETI and ECAI, we found that a considerable connection exists between the drive to conduct evaluation and individual beliefs and willingness to practice ET. This finding underscores the fact that for individuals, the relationship between motivation, awareness, and competence necessary to conduct evaluation is linked with thinking evaluatively. Since the strongest correlation was between *Motivation* and *Believing and Practicing Evaluation,* it appears that a considerable connection exists between the drive to conduct evaluation and individual beliefs and willingness to practice ET.

ET is a cornerstone of evaluation capacity, so it makes sense that evaluative capacity and attitudes towards evaluation were also changed through PACE participation. Findings show that at the individual level, PACE PPs gained *Competence* with more skills and knowledge to conduct evaluation, *Awareness* of the benefits of evaluation, and *Motivation* to conduct evaluation. At the Organization level, participants increased *Leadership* support for evaluation with *Resources* and a supportive *Learning Climate* where staff are *Mainstreaming*

evaluation into workflows and encouraging the *Use* of findings. Thus, the findings reaffirmed previous research that changing an individual's approach to evaluation capacity through awareness, motivation, and competence is a critical first step in influencing organizational attitudes on evaluation capacity and allocation of resources.

The results also showed that the retrospective pretest-posttest design was more sensitive to change than the traditional pre-post survey, which we discerned through the use of a supplemental structure in alignment with the literature. The retrospective approach captured two additional *Individual* changes (*Awareness* of the benefits of evaluation and *Motivation* to conduct evaluation), and *Organization* level capacity about necessary *Resources* for evaluations. The retrospective approach surfaced the fact that the biggest change appeared in individuals' *Competence* and *Motivation*, which suggests that PACE improved participants' skills, knowledge, and commitment to evaluation.

While we were hopeful that program professionals would experience positive changes related to ET as a result of PACE participation, there were several unexpected gains illustrated by interviews. In addition to the emergence of themes consistent with the definition of ET, participants overwhelmingly noted that ET changed the way they approach evaluation. ET became a new way of thinking for many participants as they adjusted their beliefs and practices to integrate ET concepts in their approach to program evaluation. Many participants described a shift in their attitudes about evaluation, and a few described a particularly deep understanding of and commitment to ET, noting that ET has even altered their worldview.

In many interviews, PPs described becoming less fearful of a process they used to dread, instead, becoming confident in their ability to discuss evaluation planning. Participants credited PACE with equipping them with the language, skills and tools to advocate for their

programs with executive leadership, stakeholders, and external evaluators. Additionally, participants suggested that ET specifically and PACE generally impacted their approach to service delivery. Participants also recognized the significance of partnering with external evaluators who value ET. They noted that ET changed their professional scope of practice and aspects of their personal lives. Moreover, many participants described cultural shifts in their organizations due to PACE including ET's influence on staff training and professional development, enhanced stakeholder engagement, and valuing stakeholder perspectives.

Conclusion

The results and key takeaways of this study suggest that the PACE Project design offers a promising strategy for building evaluation capacity in youth character programs, and beyond. Distinctive features of the PACE design include the use of Relational Systems Evaluation as the foundational approach, the evaluation partnership structure which paired professional evaluators and program professionals in ways designed to engage and integrate researcher and practitioner expertise, and mixed training modes that employed interactive, learning-by-doing strategies applied to active real-world programs. Other studies of the PACE Project have affirmed the positive value associated with the relational aspects of PACE evaluation partnerships (Chapter 4 of this volume), and the extensive and diverse mix of value propositions associated with the PACE approach to ECB (Chapter 7 of this volume). Taken together, these results from the basic PACE design lay the groundwork for future innovations. For example, we encourage the implementation of PACE more efficiently and more often with an intentional focus on developing the networking connections between program professionals and evaluators. Doing so could build a cadre of people with similar tools and language within the character development field, all of whom are working towards the development and sustainability of effective programs. We also hope to create

opportunities for more funders to participate in PACE in a meaningful way to foster buy-in to and connections with current programs. Ultimately, we envision this cadre of individuals, organizations, and funders as key motivators in supporting the provision of strong, resonant programming that is continuously assessed through high-quality evaluation over time.

Chauveron, L. M., Urban, J. B., Samtani, S., et al. (2021). Promoting evaluation in youth character development through enhanced Evaluation Capacity Building: Empirical findings from the PACE Project. *New Directions for Evaluation*, 2021, 79– 95. https://doi.org/10.1002/ev.20447

References

Allen, J.M., & Nimon, K. (2007). Retrospective pre-test: A practical technique for professional development evaluation. *Journal of Industrial Teacher Education, 44(3),* 27-42.

Archibald, T., Sharrock, G., Buckley, J., & Young, S. (2018). Every practitioner a "knowledge worker": Promoting evaluative thinking to enhance learning and adaptive management in international development. *New Directions for Evaluation, 158*, 73-91. doi:10.1002/ev.20323

Baker, A., Bruner, B. (2012*). Integrating evaluative capacity into organizational practice*. Cambridge, MA: The Bruner Foundation. Retrieved from http://www.evaluativethinking.org/docs/Integ_Eval_Capacity_Final.pdf

Buckley, J., Archibald, T., Hargraves, M., & Trochim, W.M. (2015) Defining and teaching evaluative thinking: Insights from research on critical thinking. *American Journal of Evaluation, 36*(3), 357-388. doi:10.1177/1098214015581706

Chauveron, L., Samtani, S., Groner, M., Linver, M.R., & Urban, J.B. (in press). Including Diverse Stakeholder Voices in Youth Character Program Evaluation. *American Journal of Evaluation*.

Fetterman, D., & Wandersman, A. (Eds). (2005). *Empowerment evaluation principles in practice*. New York, NY: Guilford Press.

García-Iriarte, E., Suarez-Balcazar, Y., Taylor-Ritzler, T., & Luna, M. (2011). A catalyst-for-change approach to evaluation capacity building. *American Journal of Evaluation*, *32*(2), 168-182.

Geldhof, G.J., Warner, D.A., Finders, J.K., Thogmartin, A.A., Clark, A., & Longway, K.A.

(2018). Revisiting the utility of retrospective pre-post designs: The need for mixed-method pilot data. *Evaluation and Program Planning*, *70*, 83-89. doi:10.1016/j.evalprogplan.2018.05.002

Hershberg, R.M., & Johnson, S.K. (2019). Critical reflection about socioeconomic inequalities among White young men from poor and working-class backgrounds. *Developmental Psychology, 55*(3), 562–573. doi: 10.1037/dev0000587

Howard, G.S. (1980). Response-shift bias. *Evaluation Review, 4*(1), 93-106. doi:10.1177/0193841x8000400105

Howard, G.S., Ralph, K.M., Gulanick, N.A., Maxwell, S.E., Nance, D.W., & Gerber, S.K. (1979). Internal invalidity in pre-test-post-test self-report evaluations and a re-evaluation of retrospective pre-tests. *Applied Psychological Measurement*, *3*(1), 1-23. doi:11.1177/014662167900300101

Labin, S.N. (2014). Developing common measures in evaluation capacity building: An iterative science and practice process. *American Journal of Evaluation*, *35*(1), 107-115. doi:10.1177/1098214013499965

Labin, S.N., Duffy, J.L., Meyers, D.C., Wandersman, A., & Lesesne, C.A. (2012). A research synthesis of the evaluation capacity building literature. *American Journal of Evaluation*, *33*(3), 307-338. doi:10.1177/1098214011434608

McIntosh, J.S., Buckley, J., & Archibald, T. (2020). Refining and measuring the construct of evaluative thinking: An exploratory factor analysis of the Evaluative Thinking Inventory. *Journal of MultiDisciplinary Evaluation*, *16*(34), 104-117.

Nelson, A.G., King, J.A., Lawrenz, F., Reich, C., Bequette, M., Pattison, S., ... & Cardiel, C. L. (2018). Using a complex adaptive systems perspective to illuminate the concept of

evaluation capacity building in a network. *American Journal of Evaluation, 40*(2), 214-230.doi:10.1177/1098214018773877

O'Sullivan, R.G. (2004). *Practicing evaluation: A collaborative approach*. Thousand Oaks, CA: Sage.

Patton, M.Q. (2005). *In conversation: Michael Quinn Patton*. Interview with Lisa Waldick, from the International Development Research Center. Retrieved from http://www.idrc.ca/en/ev-30442-201-1-DO_TOPIC.html

Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation, 29*(4), 443-450. doi: 10.1177/1098214008324182

Rodríguez-Campos, L. (2005). *Collaborative evaluations: A step-by-step model for the evaluator*. Tamarac, FL: Llumina Press.

Roth, J.L., Brooks-Gunn, J., Murray, L., & Foster, W. (1998). Promoting healthy adolescents: Synthesis of youth development program evaluations. *Journal of Research on Adolescence, 8*(4), 423-459.

Saldana, J. (2016). *The coding manual for qualitative researchers.* Thousand Oaks, CA: Sage.

Stockdill, S.H., Baizerman, M., & Compton, D.W. (2002). Toward a definition of the ECB process: A conversation with the ECB literature. *New Directions for Evaluation*, *93*, 7-26. doi:10.1002/ev.39

Suarez-Balcazar, Y., & Taylor-Ritzler, T. (2014). Moving from science to practice in evaluation capacity building. *American Journal of Evaluation, 35*(1), 95-99. doi:10.117/109821401349944

Taylor-Ritzler, T., Suarez-Balcazar, Y., Garcia-Iriarte, E., Henry, D.B., & Balcazar, F.E. (2013). Understanding and measuring evaluation capacity: A model and instrument validation study. *American Journal of Evaluation*, *34*(2), 190-206.

Trochim, W.M, Urban, J.B., Hargraves, M., Hebbard, C., Buckley, J., Archibald, T., Johnson, M., & Burgermaster, M. (2016). The Guide to the Systems Evaluation Protocol (V 3.1). Ithaca, NY: Cornell Digital Print Services. https://core.human.cornell.edu/research/systems/protocol/index.cfm

Urban, J.B., Hargraves, M., Hebbard, C., Burgermaster, M., & Trochim, W.M. (2011, October). *Evaluation in the context of lifecycles: A place for everything and everything in its place*. Paper presented at the annual meetings of the American Evaluation Association, Anaheim, CA.

Urban, J.B., Hargraves, M., & Trochim, W.M. (2014). Evolutionary Evaluation: Implications for evaluators, researchers, practitioners, funders and the evidence-based program mandate. *Evaluation and Program Planning, 45*, 127-139. doi: 10.1016/j.evalprogplan.2014.03.011

Urban, J.B., & Trochim, W.M. (2009). The role of evaluation in research-practice integration: Working toward the "golden spike". *American Journal of Evaluation, 30*(4), 538-553. doi:10.1177/1098214009348327

Urban, J.B., Burgermaster, M., Archibald, T., & Byrne, A. (2015). Relationships between quantitative measures of evaluation plan and program model quality and a qualitative measure of participant perceptions of an evaluation capacity building approach. *Journal of Mixed Methods Research, 9*(2), 154-177. doi: 10.1177/1558689813516388

Chauveron, L. M., Urban, J. B., Samtani, S., et al. (2021). Promoting evaluation in youth character development through enhanced Evaluation Capacity Building: Empirical findings from the PACE Project. *New Directions for Evaluation*, 2021, 79– 95. https://doi.org/10.1002/ev.20447

Wandersman, A. (2014). Moving forward with the science and practice of evaluation capacity building (ECB): The why, how, what, and outcomes of ECB. *American Journal of Evaluation, 35*(1), 87-89. doi:10.1177/1098214013503895

Chauveron, L. M., Urban, J. B., Samtani, S., et al. (2021). Promoting evaluation in youth character development through enhanced Evaluation Capacity Building: Empirical findings from the PACE Project. *New Directions for Evaluation*, 2021, 79– 95. https://doi.org/10.1002/ev.20447

Table 6.1. PACE Program Delivery Modes and Goals

WS1 (3½ days)	WS2 (1½ days)	PACE Culminating Conference (2 days)	Webinars
Strengthen the identity of CD programs around developing virtues. Promote ET adoption through practices and activities. Increase capacities to plan high quality evaluation based on the principles of Evolutionary Evaluation and the SEP.	Increase capacities of PPs to prepare for implementation of evaluation plans and utilization of evaluation results, including selection or development of measures, analysis of results, and communication. Deepen understanding of alignment between evaluation questions, results, claims, and plans. Increase capacity to identify relevant research for individual CD programs and their evaluation. Highlight ET's role in continued evaluation implementation and utilization.	Share work completed by all programs and evaluation partnerships. Create opportunities to network with other PPs, ECBers, and potential funders. Share best practices in presentations from PACE participants.	Share CD concepts, definitions, and overview of the field. Introduce person-centered analysis concepts and methods. Share examples for use in CD programs.

Table 6.2 Paired Samples *T*-test Results on the ETI

ETI Subscales	Pretest M(SD)	Posttest M(SD)	Mean Difference	Effect Size (Cohen's d)
Believing in and practicing evaluation	3.78 (0.69)	4.30 (0.48)	.52***	0.80
Posing thoughtful questions and seeking alternatives	3.97 (0.50)	4.27 (0.51)	.30*	0.44
Describing and illustrating thinking	3.23 (1.04)	3.96 (0.95)	.73***	0.84

*Note: N=26, *p<.05; ***p<.001*

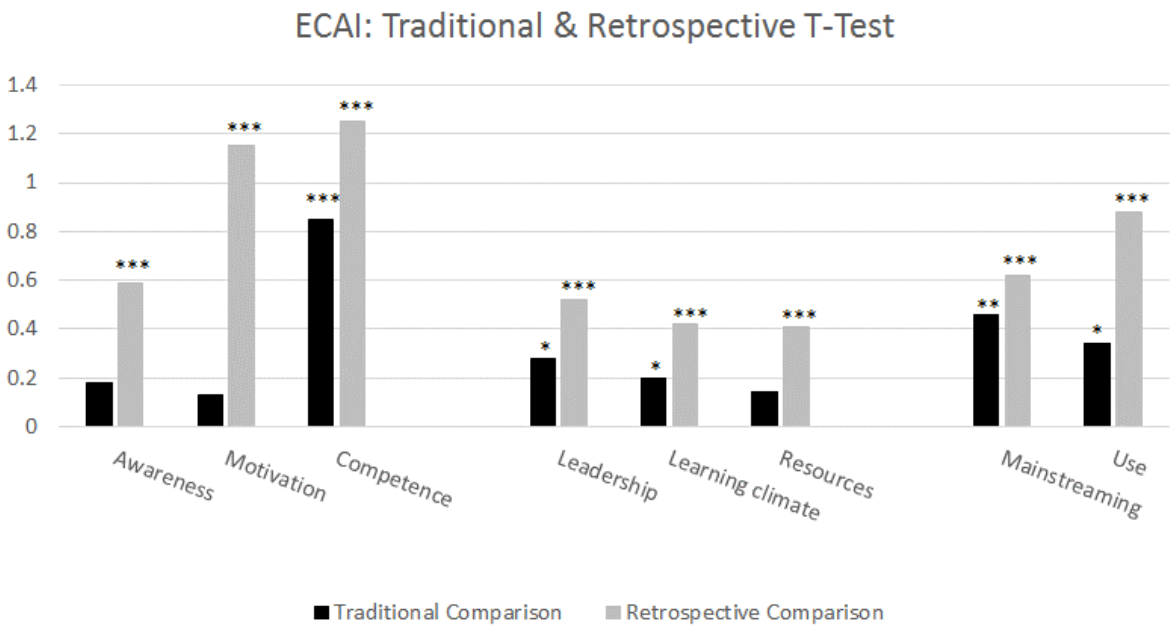
Chauveron, L. M., Urban, J. B., Samtani, S., et al. (2021). Promoting evaluation in youth character development through enhanced Evaluation Capacity Building: Empirical findings from the PACE Project. *New Directions for Evaluation*, 2021, 79– 95. https://doi.org/10.1002/ev.20447

Table 6.3. Correlations between ETI and ECAI Subscales (N=26)

		1	2	3	4	5
1.	ETI Believing	1				
2.	ETI Thoughtful	.853**	1			
3.	ETI Illustrating thinking	.089	.396*	1		
4.	ECAI Awareness	.494*	.508**	.159	1	
5.	ECAI Motivation	.704**	.601**	.054	.716**	1
6.	ECAI Competence	.613**	.555**	.256	.298	.457*

* $p < .05$; ** $p < .01$.

Chauveron, L. M., Urban, J. B., Samtani, S., et al. (2021). Promoting evaluation in youth character development through enhanced Evaluation Capacity Building: Empirical findings from the PACE Project. *New Directions for Evaluation*, 2021, 79– 95. https://doi.org/10.1002/ev.20447



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

Figure 6.1. Change in attitudes about evaluation and evaluation capacity.

LISA M. CHAUVERON is a researcher and program evaluator with a focus on urban community-based, educational, and health organizations providing programs for marginalized youth and families.

JENNIFER BROWN URBAN is professor of Family Science & Human Development, and co-Director of the Institute for Research on Youth Thriving and Evaluation, both at Montclair State University. Her research focuses on positive youth development and evaluation capacity building in youth-serving organizations.

SATABDI SAMTANI is a research scientist for the CViL project at the Center for Character and Citizenship, University of Missouri –St. Louis. She is a social science methodologist by training with research interests in monitoring and evaluation of social and community programs.

MILIRA COX a Doctoral Candidate in Family Science and Human Development and a Doctoral Research Fellow at the Montclair State University Institute for Research on Youth Thriving and Evaluation.

LESLIE K. MOORMAN is a Doctoral Candidate in Family Science and Human Development and a Doctoral Research Fellow at the Montclair State University Institute for Research on Youth Thriving and Evaluation.

Chauveron, L. M., Urban, J. B., Samtani, S., et al. (2021). Promoting evaluation in youth character development through enhanced Evaluation Capacity Building: Empirical findings from the PACE Project. *New Directions for Evaluation*, 2021, 79– 95. https://doi.org/10.1002/ev.20447

MONICA HARGRAVES is the Associate Director for Evaluation Partnerships in the Cornell Office for Research on Evaluation, at Cornell University, where her work focuses on a collaborative approach to evaluation and evaluation capacity building using the tools of Relational Systems Evaluation.

JANE BUCKLEY of JCB Consulting is an evaluation planning facilitator and evaluation capacity building consultant whose focus is on fostering a culture of evaluative thinking in partner organizations.

MIRIAM R. LINVER is professor of Family Science & Human Development, and co-Director of the Institute for Research on Youth Thriving and Evaluation, both at Montclair State University; her research focuses on diverse contexts of youth development.

Chauveron, L. M., Urban, J. B., Samtani, S., et al. (2021). Promoting evaluation in youth character development through enhanced Evaluation Capacity Building: Empirical findings from the PACE Project. *New Directions for Evaluation*, 2021, 79– 95. https://doi.org/10.1002/ev.20447