



Developing Good Character Begins with Planning and Modeling for Character Initiatives: Evaluating Character Development Programs Using Evolutionary Evaluation and the Systems Evaluation Protocol

Jennifer Brown Urban and Miriam R. Linver

This is an unpublished conference paper for the 6th Annual Jubilee Centre for Character and Virtues conference at Oriel College, Oxford University, Thursday 4th – Saturday 6th January 2018.

These papers are works in progress and should not be cited without author's prior permission.

Jubilee Centre for Character and Virtues

University of Birmingham, Edgbaston, Birmingham, B15 2TT United Kingdom

T: +44 (0) 121 414 3602 F: +44 (0) 121 414 4865

E: jubileecentre@contacts.bham.ac.uk W: www.jubileecentre.ac.uk



Developing Good Character Begins with Planning and Modeling for Character Initiatives: Evaluating Character Development Programs Using Evolutionary Evaluation and the Systems Evaluation Protocol

Jennifer Brown Urban & Miriam R. Linver

In order to understand and effectively cultivate virtues in the public sphere, we must embrace evaluative thinking and engage in careful program planning and evaluation. The importance of engaging in high quality program evaluation is a generally accepted principle underscored by external pressure from foundations and government entities that fund character development initiatives (e.g., National Academies of Science, 2016; Social and Character Development Research Consortium, 2010). To meet funders' requirements to demonstrate program effectiveness, programs must participate in high quality evaluations; this process includes: a focus on utilization and learning, practitioner involvement in evaluation planning and implementation (when appropriately supported by evaluation professionals and other stakeholders), and alignment of the evidence gathering method with the program's lifecycle stage and intended use. High quality evaluation necessarily begins with good evaluation planning.

Before jumping headfirst into construct measurement, a good evaluation plan starts with modeling the theory of change that underlies the program, identifying specific evaluation questions, and mapping how the specific program fits with its surrounding environment. As Aristotle famously remarked, "Well begun is half done." If data collection commences without laying the groundwork, the resulting data may not be useful for answering appropriate research questions, and/or the evaluation design may not be appropriate for the program's stage of development (e.g., using a randomized control trial with a newly implemented program; Urban,

Hargraves, & Trochim, 2014). The early investment in planning can pay off with high quality evaluation data that can be used to validate and/or improve programs.

Barriers to thorough evaluation planning include lack of time, funding, and evaluation know-how. These limitations can lead to overly simplistic and/or inappropriate data that may not be sufficient to demonstrate program effectiveness. Two steps can lead to program improvement and ultimately enhance the programs' quality and reach. The first is to engage in evaluation planning before the evaluation launches; the second is to build evaluation capacity of program staff so they are able to conduct internal evaluations when appropriate and effectively interact with external evaluators when needed.

Programs designed to promote character virtues in particular need high-quality evaluation; many of these are small, local programs (e.g., sports teams, scouting), or are affiliates of national organizations such as Boy Scouts of America or Boys and Girls Clubs (Roth & Brooks-Gunn, 2003a). Despite their proliferation and reach, the majority of these programs have not been formally evaluated and we do not know if they are working (Roth & Brooks-Gunn, 2003b, 2003c, 2016; Roth, Brooks-Gunn, Murray, & Foster, 1998). Knowledge of evaluation is important for both small scale/new programs as well as for large-scale/established programs. Evaluation expertise affords practitioners the ability to conduct internal evaluations, potentially leading to program improvements and enables practitioners to report program impact to funders. Practitioners within larger-scale and more established programs, would also benefit from increased evaluation capacity, so they can more effectively partner with external evaluators.

Evaluation capacity building is that much stronger when program staff and evaluators work as partners, with both program staff and evaluators valued equally. This partnership ensures buy-in from program staff, as well as evaluations that take into account the real-world

environments where programs are situated. One approach to evaluation capacity building that is grounded in a partnership model is Evolutionary Evaluation (EE; Urban et al., 2014) and specifically, the Systems Evaluation Protocol (SEP; Trochim et al., 2012). In this paper, we describe how three character initiatives utilized EE and the SEP and discuss the lessons learned.

First we present an overview of EE and the planning stage of the SEP, a step-by-step guide that applies EE principles. Next we discuss three unique applications of the SEP in the general arena of character virtue development: (1) a single program evaluation of a character development program in Scotland; (2) an evaluation capacity building project that promotes training and collaboration between evaluators and youth character development program staff; and (3) planning for a large center with multiple research strands, outreach goals, and character programs, all focusing broadly on the promotion of character virtues.

Evolutionary Evaluation and the Systems Evaluation Protocol

The EE perspective considers the complex factors inherent in larger systems within which a program is embedded (Trochim et al., 2012; Urban, Hargraves, Hebbard, Burgermaster, & Trochim, 2011, November; Urban et al., 2014; Urban & Trochim, 2009). The SEP is a specific approach to evaluation that applies the principles of EE, provides an integrated, actionable foundation for planning and conducting evaluations, developing and improving programs, and fostering capacity of and commitment to evaluation among participating programs (Buckley, 2015; Trochim et al., 2012; Urban et al., 2011, November; Urban et al., 2014; Urban & Trochim, 2009).

EE articulates a process of program development and evaluation as an evolution with inherent lifecycle qualities. Just as we characterize periods of human development (e.g., infancy, childhood, adolescence) we can similarly discuss program development. Every program has a

lifecycle: programs are initiated; they typically go through times of rapid change and growth; they may stabilize and become more “settled”; they may be disseminated widely; and at any point they may be retired or replaced. This view of program evolution is operationalized by four phases of program lifecycles: (1) Initiation, (2) Development, (3) Stability, and (4) Dissemination.

A program’s evolution is marked by the passage of time, but also by a substantive progression that includes refinement and stabilization of program content and approach. For any given program lifecycle stage, there is an appropriate type of evaluation work. Evaluation lifecycles are divided into four phases, parallel to the four phases of program lifecycles: (1) Process and Response, (2) Change, (3) Comparison and Control, and (4) Generalizability. Alignment between program and evaluation phases is essential to ensure both that a program obtains the kind of feedback most needed at that point in the life of the program, and that program and evaluation resources are used efficiently. When program practitioners, program managers, and evaluators employ evolutionary program evaluation, better decisions can be made about whether to keep, modify, or retire a program, and about what kinds of evaluations to conduct and fund (Urban et al., 2014). These same principles are incorporated into the SEP during each step of the process, including developing program logic and pathway models, identifying key pathways and nodes (activities, outputs and outcomes), determining the boundary conditions for program models, assessing program lifecycles, and selecting evaluation designs that are appropriate to program evolution.

The Systems Evaluation Protocol

The SEP integrates principles from multiple theoretical perspectives, including developmental systems theory (e.g., Lerner, 2006; Overton, 2006; Overton, 2010; Urban et al.,

2014), evolutionary theory (Darwin, 1859; Mayr, 2001), evolutionary epistemology (Bradie & Harms, 2006; Campbell, 1974, 1988; Cziko & Campbell, 1990; Popper, 1973, 1984), and systems theory (Bertalanffy, 1995; Laszlo, 1996; Midgley, 2003; Ragsdell, West, & Wilby, 2002). This approach emphasizes: the need for constructing a causal diagram of the program's theory of change; consulting internal and external stakeholders about their perspectives on and priorities for the program; recognizing how the program is related globally to other programs, in part by identifying research on related outcomes which can link the program to universal long-term goals; and continually assessing/revising the evaluation plans to collect evidence on and improve the program (Urban & Trochim, 2009).

From these considerations, the SEP distills a series of essential steps to create and implement an evaluation plan. Although the foundations of the approach are rooted in theory potentially viewed as less accessible by non-academics, the process of undertaking a high-quality evaluation via the SEP becomes manageable, replicable, and teachable to program practitioners whose expertise lies in program development, delivery, and management, the realities of their context, and their target audiences. The Systems Evaluation Protocol includes three stages of evaluation planning: (1) Preparation, (2) Model Development, and (3) Evaluation Plan Development (see Figure 1).

The current study highlights the Model Development stage of evaluation planning. Urban and colleagues (In press) lay out a detailed description of the full SEP, including Preparation, Model Development, and Evaluation Plan Development.

Model development. The Model Development stage of the SEP is the focus of the three case studies presented in this paper. This phase emphasizes surfacing and articulating deeper understandings of the program through: stakeholder analysis and mapping (a visual depiction of

the stakeholders and their relationship to each other); group discussion and program review resulting in a written program description; boundary analysis; identification of program and evaluation lifecycle phases; and structured program modeling in two forms, (1) the more familiar columnar logic model (with Program Assumptions, Context, Inputs, Activities, Outputs, and Outcomes), and (2) a corresponding visual pathway model.

<<INSERT FIGURE 1 HERE>>

A pathway model is a visual program model that incorporates two components of the columnar logic model (Activities and Outcomes) simultaneously specifying the theory of change underlying the program. The knowledge or insight made evident by the pathway model is often held subconsciously by people involved with the program, and the opportunity to articulate and combine the insights of a number of central people is rare. The key products of the Model Development stage are the stakeholder map, lifecycle determinations, and the logic and pathway models. These products form the foundation for strategic decision-making about evaluation scope, evaluation purpose, and ultimately the specific evaluation questions.

The Systems Evaluation Protocol can be applied to a wide array of programs. Next, we turn to three case studies where the Systems Evaluation Protocol was applied: The first case study is an evaluation of a specific character development program, Inspire Aspire. This case study illustrates both the *how* and the *why* of the SEP; the case study allows us to explain *how* the SEP can be done, illustrating specific details about each stage of the SEP, as well as *why* the SEP is useful and important. The second case study, Partnerships for Advancing Character Program Evaluation (PACE Project), is an evaluation capacity building project. PACE was designed to promote training and collaboration between evaluators and youth character development program staff; this case study illustrates the importance of Evaluation Capacity

Building (ECB), as well as the value of the SEP in ensuring practitioner buy-in to the evaluation process. The third case study details the planning and modeling of third phase of the Jubilee Centre, a large-scale interdisciplinary center-based initiative to support research, policy, and practice related to character development. This case study highlights how EE and the SEP can help focus goals and streamline resources across multiple interwoven strands of emphasis and projects within a larger institution.

Study 1: Inspire Aspire

Inspire Aspire: Global Citizens in the Making is a school-based character education program that was developed in the United Kingdom by Character Scotland. Directed by their teachers, S2 pupils engage in a process of self-discovery that culminates with completing a poster template. Pupils reflect on which values and virtues are their strengths and areas in need of improvement; they research and write about an inspirational figure and how that figure inspires them; and, they reflect on and write about how they can turn that inspiration into aspiration including their specific plans for working toward those aspirations. Inspire Aspire has been implemented in over 60 countries and has reached over 100,000 youth.

Methods

The authors facilitated the SEP with program developers from Character Scotland which resulted in a pathway model (Figure 1). A lifecycle analysis revealed Inspire Aspire in Phase 3 (Stability) of its program lifecycle, but in Phase 1 (Process and Response) of its evaluation lifecycle. In order to work toward program and evaluation lifecycle alignment, an initial pilot study that explored the hypothesized key constructs and used a combination of Phases 1 and 2 evaluation lifecycle approaches was warranted. The pathway model guided the constructs to be measured in a pilot study conducted during the 2014/2015 academic year. After results from the

pilot study were analyzed, the facilitators met with the team at Character Scotland again to review the study results and revise the pathway model to reflect study findings (Figure 2). We found some youth did not demonstrate statistically significant change on some quantitative measures of short-term outcomes. However, our qualitative measures established many youth were in fact exhibiting change; either our quantitative measures were not sensitive enough to detect change, or the change represented a precursor to the target outcome. We revised the pathway model, therefore, to include shorter-term outcomes to be assessed both quantitatively and qualitatively. For example, qualitative results from the pilot study indicated that youth self-reflection was a critical outcome of the program but this outcome was not present in the original pathway model. Self-reflection was added to the pathway model and was measured both quantitatively and qualitatively in a subsequent outcome evaluation. In another example, contribution to civil society was a *long-term* outcome in the original pathway model. The pilot study revealed many youth were already contributing to society in the short-term or were making plans for contributing; however, they were doing so at a more local level and by engaging in small acts of kindness. Consequently, in the revised pathway model, increased local contribution was added as a medium-term outcome. A subsequent outcome evaluation using methods consistent with a Phase 3 (Comparison and Control) evaluation lifecycle was conducted using the revised pathway model as a guide.

Results

EE and the SEP were effectively used with a standalone later program lifecycle phase program whose evaluation lifecycle was misaligned when the project commenced. During the process of developing the pathway model, the program developers had key insights that led to substantial changes to the program itself. These included making substantial edits and further

developing the teacher manual, eliminating a national award ceremony that required substantial resources but was unrelated to the program's theory of change, and a youth stakeholder identified goal setting as a critical outcome of the program (an outcome that had previously been overlooked by the program developers; Urban, Linver, Thompson, Davidson, & Lorimer, in press).

The lifecycle analysis (and misalignment of phases) revealed the need to conduct an earlier lifecycle phase evaluation before engaging in a more expensive later lifecycle evaluation. Ultimately, this enabled a more effective use of resources as substantial programmatic changes were made after analyzing the results of the pilot study and before launching the outcome evaluation. Had a later lifecycle evaluation been conducted first, this may have led to misleading results since the program was still evolving (Urban et al., 2014).

Study 2: Partnerships for Advancing Character Program Evaluation (PACE Project)

A common challenge for youth programs is they are called upon to do better evaluation, for purposes of both program improvement and accountability to funders, but often organizations do not possess the knowledge and resources to evaluate their programs effectively. There is also a dearth of evaluation professionals with expertise in character development programs. To create stronger youth programs, we need to build the evaluation capacity of program practitioners and simultaneously develop a network of evaluation professionals trained in character development as well as cutting-edge evaluation and analytical techniques. The field of evaluation has recently recognized the importance of evaluation capacity building (ECB; Preskill & Boyle, 2008).

The Partnerships for Advancing Character Program Evaluation or PACE Project is an approach to ECB that offers professional development opportunities for staff of youth character development programs as well as evaluators. For program staff, PACE was designed to

immediately increase the capacity to evaluate, improve, defend and seek funding for programs. For evaluation professionals, PACE offered an opportunity to develop expertise in evaluation capacity building and innovative evaluation techniques. All PACE participants were trained in using EE and the SEP. PACE was organized as a true partnership between evaluators and program staff, where ownership of the evaluation process was shared, and both sets of professionals brought expertise to the project. This is perhaps the first ECB effort to train evaluators and program staff simultaneously as “teams.” Our hope is that the team approach will result in sustained relationships between evaluators and program staff, and produce high quality program and evaluation work for years to come. In addition, PACE offered an opportunity to innovate, and provided a path forward for professionals focused on character development, ECB, and evaluation who were interested in building a body of evidence and sustained evaluation capacity for character development programs, organizations, and systems.

Methods

The PACE Project invited applications from staff of youth character development programs and independent evaluators from across the United States. PACE Facilitators selected 16 youth character development programs (two staff from each program for a total of 32 program staff) and eight independent evaluators who were trained as Evaluation Capacity Builders (ECBers) as part of PACE. Program staff were divided into two cohorts and ECBers were matched with one program from each cohort. Over the course of 12-15 months, all PACE participants attended: 2 workshops (both in-person for cohort 1; 1 in-person and 1 web-based for cohort 2); 3 content-based webinars on the historical context and modern advances in character development research and a two-part series on person-centered analyses (required for ECBers

and optional for practitioners); virtual and in-person partnership meetings with a designated ECBER, and consultation with a Lead Facilitator.

The evaluation plan for PACE was developed using an EE approach and the SEP. The pathway model illustrates the theory of change for PACE and served as a guide for our evaluation work (Figure 3). An internal evaluation was conducted with the aim of answering the following research questions which were derived from the PACE pathway model: 1) Do PACE participants demonstrate an increase in knowledge?; 2) Do ECBERs demonstrate increased ability to facilitate PACE activities and strategies?; 3) Do program professionals demonstrate increased ability to communicate about their program and its evaluation; adopt evaluative thinking habits; and improve attitudes about evaluation? A repeated measures design was employed and quantitative and qualitative data were collected including surveys, semi-structured interviews, observation, and document review.

Results

Preliminary results indicate that the PACE approach moved beyond evaluation skill-building activities for program staff to address the foundational conditions of motivation, communication, and critical thinking that drive evaluation work forward. After participating in PACE, programs and their staff gained: enhanced evaluation capacity; a well-developed theory of change; a Program and Evaluation Profile that can be used to communicate about the program and its evaluation with stakeholders; an Evaluative Thinking site plan designed to incrementally improve the organizational culture and communication around evaluation; and an established relationship with an evaluator.

ECBERs report their PACE participation contributed to having a greater understanding and appreciation of the work practitioners do as well as a greater recognition of the barriers and

challenges practitioners face including high staff turnover, limited and/or changing budgets, and real-world issues that impact the feasibility of implementing evaluation. One ECBER noted:

I think I would have thought of evaluation as much more top-down before this and that I am a researcher and I study after school programs, and I have this general idea about how programs work. But I learned a lot from these programs, that they have a lot of expertise about how their program works. Even how you know on paper you can read about a program's theory of change, but a lot of times when you start digging in and learning about the program in practice it can look very different. And that is the expertise that I got from the program folks of basically what happens in real life, as opposed to what we read about in journal articles. A lot of that is practical piece of, you know I can tell them all day long about RCT but in reality that might be a method that does not work in practice, the way it is written about in textbooks. I learned a lot from program professionals about what is realistic in a program in terms of evaluation. There is a gap between all these gold standard ideas on doing evaluation and in practice does not always translate perfectly. And I learned from program professionals, lots of adaptation needs to be made for an evaluation to happen in context ... it was very much a "reality check" ... I would actually use that term... In terms of what I read in journal articles with what works in reality (PACE ECB participant).

ECBERs also reported an appreciation of the collaborative partnership model and the distinction between their role as an external evaluator versus an evaluation capacity builder, for example, another ECBER reported:

Before PACE my role as an evaluator was largely as a third party evaluator. Where I was part of a research institution that was hired to be an external evaluator for a youth program. So, through PACE my role as an evaluator, took on a different capacity. In that I came to learn through PACE that...evaluation is not just an external thing, and it's not something that I would do now... if I took on a similar role. I think there is a place for third party evaluation, I am not saying that. After PACE if I took on a similar role as I learned through the PACE project, that would shift to a more collaborative process, where my job is not just to give information to the program but more collaboratively and bi-directionally working with the program, to learn about the program. Whereas I think as a third party evaluator it is more gathering information and just handing it over. But I think after working through PACE it was much more collaborative and their expertise was shared much more with the program, rather than sort of being just owned by the evaluator. I think that is the biggest distinction I would make (PACE ECB participant).

Program staff also reported having a better attitude toward evaluation and an increase in evaluation knowledge after participating in PACE. Practitioners also valued the concept of

program and evaluation lifecycles and noted that for many of their programs the lifecycles were misaligned. Two program practitioner participants noted:

So based on what we learned in PACE, Evolutionary Evaluation has to do with the lifecycle of both the program and the evaluation... knowing full well that it was way too early to expect those outcomes...I feel like this gave me a new perspective on what should be changing and what impact we should be seeing based on where the program is. This is really valuable information... (PACE Program practitioner participant).

So I had it in 2 ways...together with my colleague...we had the sort of “Aha!” moment of like...Oh! We have been just doing outcome or impact evaluation, we totally skipped the whole step around understanding whether a program was working, whether it was being implemented in the same way across different schools...all this type of information that would have happened if we had really evolved our evaluation and started with what was appropriate for a new program and it made us really think, Oh! since we switched that and we had to, since we had to jump to impact evaluation because of a funder. We need to go back and additionally get some of the information that we missed from what would have been an earlier evaluation, and it gave us the confidence and a framework to talk to our boss and the president of the organization (PACE Program practitioner participant).

Preliminary results from the PACE project indicate that the pairing of program practitioners with evaluators trained as evaluation capacity builders is effective. The ECBers and program staff learned and grew together raising the overall buy-in to the PACE method.

Study 3: Phase 3 of Jubilee Centre for Character and Virtues

EE and the SEP were used to plan and model the third phase of the Jubilee Centre, a large-scale interdisciplinary center-based initiative to support research, policy, and practice related to character development. The Jubilee Centre is housed within the School of Education at the University of Birmingham and aims to strengthen character across multiple contexts including the family, school, and communities. The first phase of the Centre’s work focused on gratitude and the second phase focused on service. In the third phase of the Centre’s work, the focus is building on and extending the work of the first two phases by re-analyzing already existing data, developing and testing potentially transformative interventions, and engaging in widespread dissemination of findings to policy and practice communities.

Methods

An intensive two-day implementation of EE and the SEP was conducted with staff from the Jubilee Centre. The primary objective was to develop a clear, detailed, and testable theory of change for the Centre as it enters its third phase of work. Unlike a standalone program, the Centre supports multiple projects that work toward superordinate shared objectives. In this sense, the Centre is itself a system with multiple sub-parts. One of the early challenges when working with a Centre is clarifying the boundaries of the system and determining what precisely will be modeled.

Prior to beginning the SEP, staff at the Jubilee Centre had already delineated several projects that would be the focus of their phase three work. Each of these projects ultimately aimed to transform character development/education in some way and in many cases built from previous work conducted during the first two phases of the Centre. We determined early on that the goal would not be to model each of the individual projects, but rather, to develop a pathway model that encompassed the objectives of the subordinate projects. For example, several projects would likely produce results and recommendations for teachers to use in practice. Therefore, one of the activities in the pathway model is “develop continuing professional development program/intervention (in-service).” Similarly, the Centre is particularly well known for its ability to influence policy and policymakers. Therefore, another activity in the pathway model is “network-raise profile of the Centre” which leads to the short-term outcome “products of Centre and partners are highly regarded and cited.” Several of the specific projects in which the Centre is engaged serve these higher order goals for the Centre as a whole.

Results

The pathway model was developed according to the standard SEP phases. Once the pathway model was constructed, we engaged in an activity called “mining the model” where we ask several questions of the model including: (1) which boxes have lots of arrows going into them (prime destinations), lots of arrows coming out of them (gateways), and lots of arrows going both in and out of them (hubs); (2) what are the key linkages between any two boxes; and (3) who are the key stakeholders and which activities or outcomes would they most care about. Using all of this information, we identified the key “throughlines” or connections from an activity all the way out to a long-term outcome. We identified four key “throughlines” in the Jubilee Centre pathway model (Figure 4). The purple path begins with the activity “develop continuing professional development program/intervention (in-service) and ends at the long-term outcome “new generation of young people grow up with understanding of why character matters.” In this path, the focus is on teachers and parents increasingly using the language of character virtues. The orange path begins with the same activity and ends with the long-term outcome “people lead virtuous lives by cultivating moral, civic, performance and intellectual virtues integrated by phronesis.” In this path, the focus is on teachers and schools adopting character and virtue ethos in the curriculum and in the culture of the school. The blue path begins with the activity “write and disseminate research project summaries and reports” and ends with the long-term outcome “individual and societal flourishing.” Here the focus is on the utilization of the Centre’s research and products. Finally, the green path begins with the activity “network – raise profile of the Centre” and ends with the long-term outcome “people lead virtuous lives by cultivating moral, civic, performance and intellectual virtues integrated by phronesis.” In this path, the focus is on the role the Centre plays in influencing policy.

The Jubilee Centre provides an example of how EE and the SEP can help focus goals and streamline resources across a large center with multiple interwoven strands of focus and projects. Having a clearly articulated pathway model for such a large initiative can be critical when trying to elicit buy-in from funders and other key stakeholders who may want to see a clearly articulated theory of change prior to investing. The pathway model has the added benefit of serving as a blueprint for subsequent evaluation work. Although we would not recommend trying to evaluate everything in a pathway model in one evaluation cycle, a smaller evaluation scope can be carved out of the pathway model. Once the evaluation scope is determined, specific evaluation questions can be crafted and an evaluation plan can be developed.

Conclusion

The above three case studies present three different contexts in which EE and the SEP were effectively implemented with character development initiatives. One of the three cases involved a standalone character education program and one involved a large-scale Centre-based initiative. The third case study employed a modified practitioner training approach to EE and the SEP whereby both character development program practitioners and evaluators were simultaneously trained using a partnership model. These three case studies provide examples of the diverse contexts and settings in which EE and the SEP can be successfully implemented to enhance character development programs.

The field of character development is in dire need of good program evaluation and planning in order to provide the highest quality programming and serve the reporting requirements of funders. Increased evaluation capacity is also needed to sustain best practices. EE and the SEP address the need to simultaneously build evaluation capacity while maintaining a focus on high quality evaluation by providing a clear process for engaging program

professionals in developing a theory of change and a language for describing their work and the value they believe it can add. Similar to other evaluative approaches that use a collaborative, participatory, or empowerment approach (Fetterman & Wandersman, 2005; O'Sullivan, 2004), the ultimate purpose of ECB is to improve program outcomes (Labin, 2014; Suarez-Balcazar & Taylor-Ritzler, 2014; Wandersman, 2014). Our hope is that EE and the SEP will lead to better outcomes for character development.

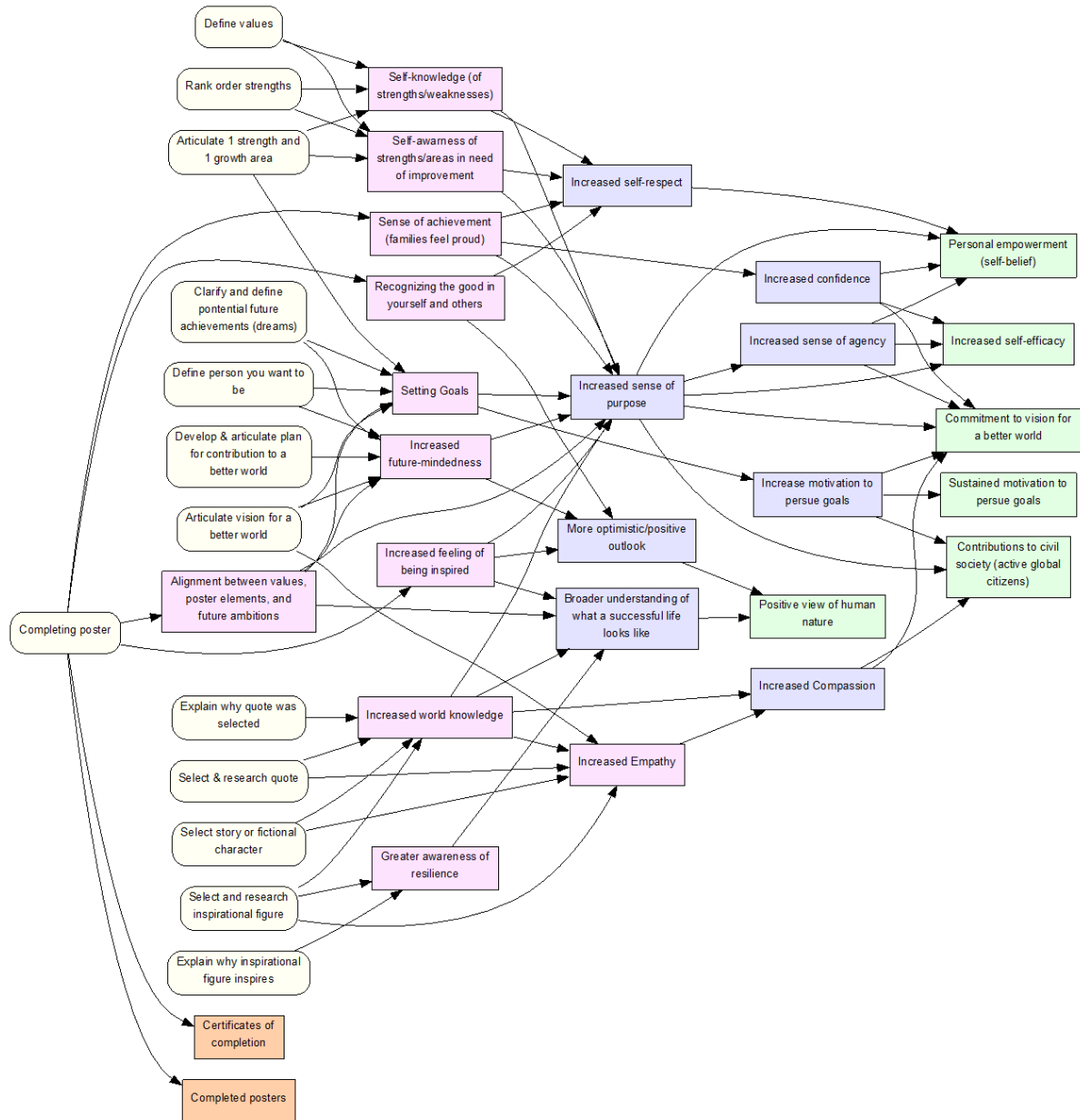


Figure 1. Initial Inspire Aspire pathway model. Yellow boxes are activities, orange boxes are outputs, pink boxes are short-term outcomes, blue boxes are medium-term outcomes, and green boxes are long-term outcomes.

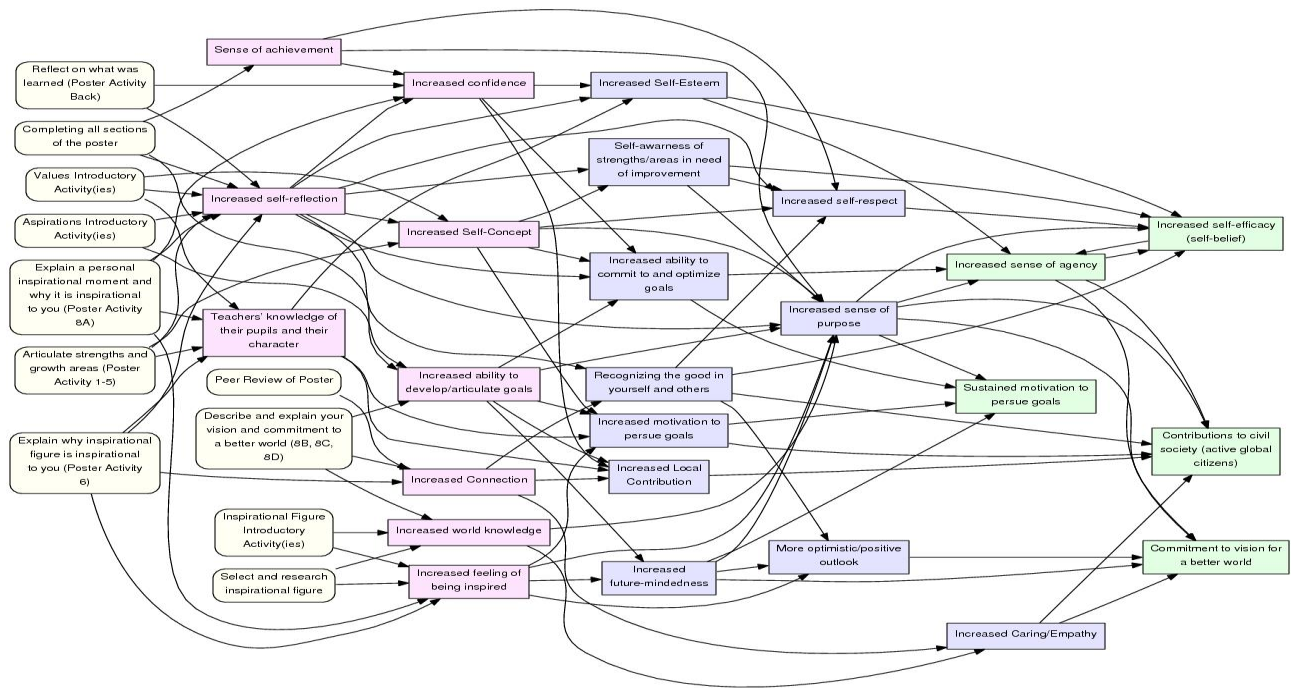


Figure 2. Revised Inspire Aspire pathway model after pilot data collection and analysis. Yellow boxes are activities, pink boxes are short-term outcomes, blue boxes are medium-term outcomes, and green boxes are long-term outcomes.

Jubilee Centre Pathway Model

Activities	Short Term Outcomes
Mid Term Outcomes	Long-Term Outcomes

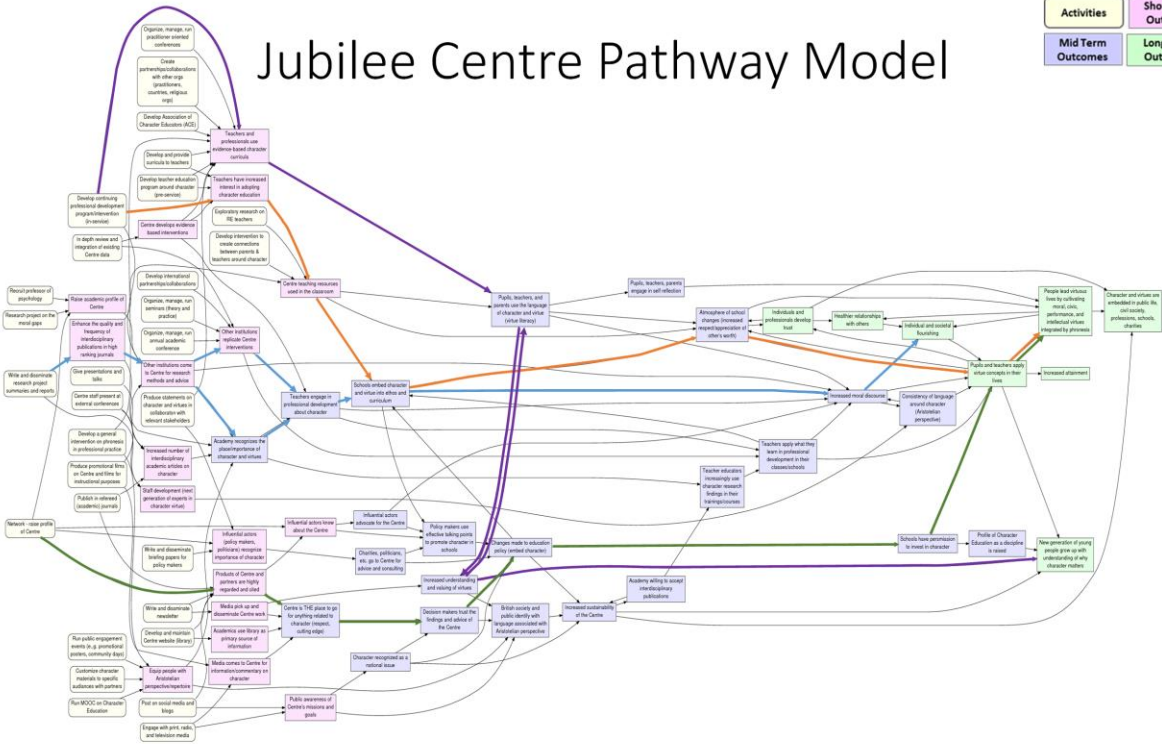


Figure 4. Pathway model for the Jubilee Centre with four key “throughlines.”

References

- Bertalanffy, L. V. (1995). *General system theory: Foundations, development, applications* (Revised ed.). New York, NY: Braziller.
- Bradie, M., & Harms, W. (2006). Evolutionary epistemology. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2016 ed.). Stanford, CA: The Metaphysics Research Lab, Center for the Study of Language and Information, Stanford University.
- Buckley, J., Archibald, T., Hargraves, M., & Trochim, W.M. (2015). Defining and teaching evaluativ thinking: Insights from research on critical thinking. *American Journal of Evaluation*, 36(3), 375-388. doi: 10.1177/1098214015581706
- Campbell, D. T. (1974). Evolutionary epistemology. In P. A. Schilpp (Ed.), *The philosophy of Karl Popper* (pp. 413-463). LaSalle, IL: Open Court.
- Campbell, D. T. (1988). Evolutionary epistemology. In E. S. Overman (Ed.), *Methodology and epistemology for social science: Selected papers of Donald T. Campbell*. Chidago, IL: University of Chicago Press.
- Cziko, G. A., & Campbell, D. T. (1990). Comprehensive evolutionary epistemology and bibliography. *Journal of Social and Biological Structures*, 13(1), 41-82. doi: 10.1016/0140-1750(90)90033-3
- Darwin, C. (1859). *On the Origin of Species by means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London, UK: John Murray.
- Fetterman, D. M., & Wandersman, A. (Eds.). (2005). *Empowerment evaluation principles in practice*. New York: NY: Guilford Press.

- 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
- Laszlo, E. (1996). *The systems view of the world: A holistic vision for our time*. Creskill, NJ: Hampton Press.
- Lerner, R. M. (2006). Developmental science, developmental systems, and contemporary theories of human development. In R. M. Lerner & W. Damon (Eds.), *Handbook of Child Psychology: Theoretical models of human development* (6th ed., Vol. 1, pp. 1-17). Hoboken, NJ: John Wiley & Sons.
- Mayr, E. (2001). *What evolution is*. New York: NY: Basic Books.
- Midgley, G. (2003). *Systems thinking*. Thousand Oaks, CA: Sage.
- National Academies of Science, Engineering & Medicine. (2016). Workshop on approaches to the development of character. Retrieved August 17, 2016, from http://sites.nationalacademies.org/DBASSE/BOTA/DBASSE_171735
- O'Sullivan, R. G. (2004). *Practicing evaluation: A collaborative approach*. Thousand Oaks, CA: Sage.
- Overton, W. F. (2006). Developmental psychology: Philosophy, concepts, methodology. In R. M. Lerner (Ed.), *Theoretica models of human development* (6th ed. ed., Vol. 1, pp. 18-88). Hoboken, NJ: Wiley.
- Overton, W. F. (2010). Life-span development: Concepts and issues. In W. F. Overton & R. M. Lerner (Eds.), *Handbook of life-span development: Cognition, biology, and methods* (Vol. 1 pp. 1-29). Hoboken, NJ: Wiley.

- Popper, K. (1973). *Evolutionary epistemology*. Paper presented at the Sections 1-VI of "the rationality of scientific revolutions" given at the Herbert Spencer lecture, University of Oxford.
- Popper, K. (1984). Evolutionary epistemology. In D. M. Miller (Ed.), *Popper selections* (pp. 78-86). Princeton, NJ: Princeton University Press.
- Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation, 29*(4), 443-450. doi: 10.1177/1098214008324182
- Ragsdell, G., West, D., & Wilby, J. (2002). *Systems theory and practice in the knowledge age*. New York, NY: Kluwer Academic/Plenum.
- Roth, J. L., & Brooks-Gunn, J. (2003a). What exactly is a youth development program? Answers from research and practice. *Applied Developmental Science, 7*(2), 94-111.
- Roth, J. L., & Brooks-Gunn, J. (2003b). What is a youth development program? Identifying defining principles. In R. M. Lerner, F. Jacobs & D. Wertlieb (Eds.), *Promoting positive child, adolescent, and family development: A handbook of program and policy innovations* (Vol. 2, pp. 197-223). Thousand Oaks, CA: Sage Publications, Inc.
- Roth, J. L., & Brooks-Gunn, J. (2003c). Youth development programs: Risk, prevention, and policy. *Journal of Adolescent Health, 32*(3), 170-182.
- Roth, J. L., & Brooks-Gunn, J. (2016). Evaluating youth development programs: Progress and promise. *Applied Developmental Science, 20*(3), 188-202. doi: 10.1080/10888691.2015.1113879
- 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.

- Social and Character Development Research Consortium. (2010). *Efficacy of Schoolwide Programs to Promote Social and Character Development and Reduce Problem Behavior in Elementary School Children*. Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.
- 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: 10.1177/1098214013499440
- Trochim, W., Urban, J. B., Hargraves, M., Hebbard, C., Buckley, J., Archibald, T., . . . Burgermaster, M. (2012). *The guide to the systems evaluation protocol*. Ithaca, NY: Cornell Digital Print Services.
- Urban, J. B., Archibald, T., Hebbard, C., Burgermaster, M., Barrios, V., & Trochim, W. M. (2015, April). *Results and implications of using the SEP and the mySEP in evaluation planning and implementation for STEM outreach programs*. Paper presented at the Eastern Evaluation Research Association, Abescon, NJ.
- 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
- Urban, J. B., Hargraves, M., Hebbard, C., Burgermaster, M., & Trochim, W. M. (2011, November). *Evaluation in the context of lifecycles: 'A place for everything, everything in its place'*. Paper presented at 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., Linver, M. R., Thompson, J., Davidson, R., & Lorimer, D. (in press). Evaluating youth character development programs using evolutionary evaluation and the systems evaluation protocol. *Applied Developmental Science*. doi:

10.1080/10888691.2017.1285236

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

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