

In ““Remind Me...Why are We Doing This?” Ten Conversations You Should Have Before Launching an Evaluation” (*GuildNotes* Fall 2011), I wrote about getting an evaluation launched and headed off in a productive direction. This article takes up where the first article left off, looking at the “mid-life” of evaluations — the time when bumps, wrinkles, and surprises often emerge. In the summer issue, I will be writing about the process of bringing an evaluation home — especially harvesting and acting on the results.

Imagine that you have successfully launched an evaluation of your program. You have a clear logic model that lays out your inputs, outputs, and outcomes. You have settled on how you will measure progress towards those outcomes. You have secured the agreements with your partners and permissions from participants. You have even raised the dollars to pay for the evaluation. The dream is that you flick the switch and the whole enterprise runs like silk and delivers you compelling findings. Only then....

- The context for the evaluation changes radically
- The findings are not what you anticipated, or
- You realize you might not be looking at the right outcomes

These mid-life “crises” are part and parcel of many evaluations — particularly those that last a year or more. When they first occur, they are disconcerting, particularly since few final reports ever acknowledge them. This is too bad, because, depending on how an organization and its evaluation partners negotiate these events, they can be crises or be occasions for deepening the evaluation process.

CRISIS 1: THE CONTEXT SHIFTS RADICALLY

Evaluations, particularly, multi-year efforts, can be laid out as elegant blueprints, with Year 1 flowing into Year 2, and seamlessly giving way to Year 3. The only flaw with that plan is that real life happens. Consider the case of ArtsRising, a Philadelphia-based initiative designed to expand arts learning opportunities for children and youth throughout the Philadelphia region. The initiative co-developed an evaluation plan in close alliance with the School District of Philadelphia to look at how increased arts learning could fuel school attendance and behavior, as well as students’ growing involvement in out-of-school arts opportunities.

But in the initiative’s third year (2011 – 2012) it became increasingly clear that the district was in very deep financial trouble — so deep that it threatened early childhood, arts, and summer programs. At this moment, the city, its families, and a School Reform Commission are reviewing plans for radical re-organization of the district. Whatever the final shape of that plan, the upshot is likely to be a much more localized and diversified system, with as many as 40 fewer public schools, smaller networks rather than a centralized district, and increasing autonomy for principals who can demonstrate strong results, winning the right to control their budgets and programs. At the same time, the city, the region, and major funders are re-thinking the role of the arts, seeing them as one of the social and cultural forces that can help to rejuvenate stressed and fractured neighborhoods. So within a matter of months, the broader context around ArtsRising’s original long-range evaluation plan has changed radically. As a result, its original plan to look closely at school-related outcomes could evolve, becoming an evaluation focused on looking at the spread and quality of arts opportunities in- and outside of school time.

What happened in Philadelphia is not an exception: In the course of any evaluation an arts-friendly mayor or superintendent can move. The state or federal government can release new standards and expectations. A financial downturn can strip arts education out of

schools and community centers. Given that reality, what are the lessons for executing an evaluation?

LESSON 1: EMBRACE EVOLUTION

In order to respond to this shifting context, ArtsRising staff and evaluators are thinking about their work in a new light: While most creative place-making efforts focus on adults, ArtsRising now has the chance to pose a prior question, how does a neighborhood “raise” the next generation of people who will make their communities vibrant? To do this both program staff and evaluators have to be:

■ **Curious:** Philadelphia is home to some of the most thoughtful work on neighborhoods and urban spaces: the *Mural Arts Program*, the *Social Impact of the Arts Project*, and the *Reinvestment Fund*, among others. As ArtsRising thinks about its new context, staff have the opportunity to reach out to new thought partners.

■ **Nimble:** ArtsRising staff are thinking about how they sustain their original commitment to increasing and improving arts education, at the same time that they re-think their own work and evaluation. At the same time, they have an important opportunity to seize. Most current conceptions of creative place-making emphasize improvements to the lives of adults: jobs, leisure activities, public spaces, as well as racial and economic integration. In its new work, ArtsRising could provoke a discussion about what it takes to support quality of life for children and youth: quality learning, safe gathering spaces, access to compelling role models, opportunities to lead and contribute. This could mean re-purposing their basic evaluation tools and data from the original evaluation (e.g., family interviews, student surveys, etc.) to address questions of how the arts foster youth development in the city’s neighborhoods. It could also mean developing new tools such as using spatial data to map the spread and growth of opportunities for youth or to identify neighborhoods needing investment. It could mean engaging youth researchers to inform and help conduct the evaluation.

■ **Resilient:** The story in Philadelphia is still evolving. This means that ArtsRising staff have to communicate with their partners and communities quite frankly about the evolving nature of the work.

CRISIS 2: THE FINDINGS ARE NOT WHAT YOU EXPECTED

Evaluations are fundamentally experiments. They ask, “If a program (or organization) does X, to what extent will Y desired outcome occur?” In that sense, the middle phase of an evaluation is all about collecting and sifting through results. More often than not, those results are mixed, even surprising or disappointing.

Consider the case of a longitudinal evaluation of the ArtsPartners program conducted by *Big Thought* in Dallas, TX. Its purpose was to examine the effects of sustained arts integration on students’ literacy. A major strand of the study looked at samples of students’ regular classroom writing and compared them to writing done in the context of arts integrated instruction using the 6 + 1 Trait writing scales developed by the Northwest Regional Education Lab (NWREL). In the first grade, student writing done in the context of arts-integrated units scored higher on dimensions such as ideas, voice, and word choice. Then, in second grade, those findings disappeared. The writing children did in the context of arts integrated lessons led by a teaching artist was no different than the writing they did when answering questions at the end of a textbook chapter. Big Thought staff, teachers, and evaluators worried, “Had the effects of having a teaching artist worn off?” “Had students developed a lockstep approach to all writing done in classrooms?” “Had this approach to evaluating the effects of arts integration run dry?” “Was the original

model of the consequences of arts-integrated writing instruction off the mark?”

LESSON 2: REMEMBER AN EVALUATION IS AN INQUIRY

What the team had to remember was that an evaluation is first and foremost an inquiry into what is and isn't working. It is more like an experiment than a certification. The point is to learn — even when that's surprising or disappointing. In the case of the ArtsPartners, the evaluation team had to:

- **Re-examine:** The teams of teachers and evaluators had to ask whether there was something different about the second grade round of units and writing prompts. When they stepped back they realized that the subject matter and the prompts were not as open-ended as in earlier rounds of data collection. The unit centered on a play based on a literal-minded character (Amelia Bedelia). But rather than being a far-reaching exploration of figurative language, the teaching focused narrowly on the idioms that gave the main character such trouble (e.g., “It’s raining cats and dogs.”) As a result, students never explored the bigger questions of playing with language and wrote in less expressive ways.

- **Re-design:** As a result, the team learned some important design constraints. The teaching and writing assignments in the evaluation had to invite students to show what they learned through their work in the arts. If the content or the assignments seemed too much like regular school work, young people did not give full-reign to the what teaching artists brought to the classrooms.

Like changes in context, surprising or disappointing results crop up regularly in the mid-course of evaluations. The point is to address them head-on and learn what they are telling you about your tools, processes, or theories.

CRISIS 3: YOU REALIZE YOU AREN'T CAPTURING THE OUTCOMES

City Lore is an organization in New York City devoted to understanding cultural heritage and folk life. As part of its work, City Lore seeks to broaden the definition of arts education to include traditional arts such as classical Indian dance and the Adinkra indigo stamp designs of Africa. This conviction is reflected in its program “Nations in Neighborhoods” where students and teachers throughout the borough of Queens use the arts to explore the complex cultural composition of their communities by studying a different region of the world each year.

“Nations in Neighborhoods” was the fortunate recipient of a four-year Arts Education Model Development and Dissemination (AEMDD) grant from the U.S. Department of Education. Under the terms of that grant, City Lore developed an evaluation plan in which it collected and reported data on whether there were greater changes for children participating in “Nations in Neighborhoods,” as compared to their peers in the same school.

While never doubting the importance of literacy gains, part way into their planned evaluation, collaborating teachers raised questions about this evaluation design. Through their work with the artists in classrooms, teachers observed more than the growth of reading skills. They had seen students speaking up, performing, and thinking about the role of culture in their own lives. Partially fueled by their own sense that standardized testing was not the full measure of their own work, teachers challenged the evaluators to expand the measures. In particular, given the many English Language Learners in their Queens classrooms, City Lore teachers (along with teaching artists) wanted to capture and present a fuller picture of what students were learning through their explorations of culture.

LESSON 3: DIG DEEPER

Even though they were partway through their evaluation, City Lore staff realized they wanted to dig deeper into their evidence. To do this, the staff:

- **Reallocated resources:** Taking teachers seriously, the staff at City Lore re-distributed project funds to create stipends for an assessment seminar that would look at additional ways to index student growth.

- **Made the evaluation participatory:** The assessment seminar was open to City Lore staff, classroom teachers, and teaching artists. The seminar widened the evaluation to include an intensive look at the development of confidence, knowledge, and independent thinking as evident in student’s oral and written language. The seminar met for a semester, refined the tools, and coded student work, finding and presenting evidence for a wider set of effects.

- **Expanded their reporting:** City Lore’s evaluation reports now track the year-to-year consequences for the full range of student performance measures they have developed, as well as for the changes in students’ performance on standardized tests of reading achievement. In addition, to sending their reports to the Department of Education, City Lore staff compose a PowerPoint “movie” that teachers can share with their principals. That movie includes the multiple strands of data, plus samples of student work and excerpts from their interviews.

Through examining student work, City Lore staff have realized where they want to strengthen their curricula, what they want to ask of teaching artists, and how that work can enrich the literacy learning particularly of English Language Learners and students with special needs.

City Lore is not an exception. Seeing your first round of data and findings is like finally seeing a manuscript in print: you suddenly realize what is — *and isn't* — there. A program that takes theater into juvenile justice settings might realize that its evaluation concentrates much too much on behavior and socio-emotional outcomes, and too little on what young people are learning about acting and playwriting. A music school might look at its data about how long students persist in its programs, and realize, “Wait, but are they progressing?” These are not mistakes, but insights. This is not an open invitation to raze your original evaluation plan. Instead, the challenge is to develop a plan that preserves what is worthwhile in the original evaluation, but adds what’s possible and reasonable — within the fiscal and human resources available to the evaluation.

SURVIVING THE MIDDLE

The middle portion of any evaluation is often the rockiest — and the most critical. It is when cultural organizations, arts education providers, and their partners are most likely to come face to face with shifts in context, unpredicted data, and the need to go beyond an original blueprint. This is the phase of any evaluation in which it is critical to:

- **View an evaluation not as a rigidly fixed blueprint, but an evolving plan to collect relevant data on effectiveness.**
- **Invest time and attention to the data as it comes in.**
- **Listen to and engage thought-partners and stakeholders.**

ABOUT THE AUTHOR

Dennie Palmer Wolf is a principal at **WolfBrown**, an international consulting firm (www.wolfbrown.com). She can be reached at dennie@wolfbrown.com.