# CHALLENGES FACED BY EDUCATIONAL EVALUATORS: PERSONAL EXPERIENCES ENCOUNTERED AND PLANNING ISSUES ILLUMINATED

**DONNA L. FERRARA**
*Southampton College of Long Island University, Professor Emerita*
*Smith-Layne Educational Consulting Services, U.S.A.*

## ABSTRACT

*This article explores issues that had to be confronted over a three-year period in terms of two New York State funded grants for which I was the evaluator. Major, critical challenges are presented, described, and discussed.*

*Specifically, the paper addresses issues related to lack of planning at the grant design stage that resulted in challenges and problems with conducting targeted and meaningful assessments to complete evaluations. This article promotes the concept that evaluators have expertise and competencies that are valuable to grant design and that if an evaluator is involved at the planning/design stage, many challenges confronted throughout evaluation processes might be prevented or mitigated.*

*The inquiries that were conducted as part of these grants can be regarded as case studies. In the example of one investigation, it was a single-case study of an intervention in a Long Island, New York, district; in the case of the other, it was a multiple-case study wherein a Charter School disseminated an intervention to three New York City schools. Mixed methods were used for data collection.*

*Following my work with these two funded grants, with lessons learned from my experiences with these two grants, I embarked on another journey with another agency where lessons learned from the two previous grants informed my approach to evaluating school programs. Thus, while I was not part of the original design team, we were able to discuss relevant evaluation issues early in the process. My latest experiences will also be briefly discussed at the end of the article in a Postscript.*

## INTRODUCTION AND BACKGROUND

Although there are countless topics of interest to educational planners, in general, there has been limited research regarding competencies of evaluators, especially related to evaluators' contributions to grant development.

While there is little evidence that this line of inquiry has been sufficiently tackled, Volkov (2011) addressed the line of inquiry by arguing that there is a field of expertise that defines the multiplicity of roles for evaluation experts. Sever (2021), building on Volkov's work, conducted a Delphi study to identify themes related to competencies of program evaluation experts, which included theoretical professional competencies, practical professional competencies, professional values experts should have, and professional skills experts should have. Sever's research was conducted under the assumption that program evaluation is a field of expertise.

Sever's study is foundational to the arguments and observations made in this article – that evaluation is a discrete area of expertise with distinct competencies. This issue of competencies had been previously introduced by Nevo (1983), but in general, unfortunately, this topic still appears

not to have gained much traction among practitioners. Additionally, there is a lack of literature regarding the role of the evaluator in the design of funded grants.

There has been useful information from which we can extrapolate evaluator competencies and roles in related areas, such as planning models (Beach, 2022; Ferrara, 2007), evaluation standards (Yarbrough et al., 2011), evaluation approaches (Rossi et al., 2019), program evaluation methods (Posavac & Carey, 1989), and improvement from a human resource frame (Bolman & Deal, 1997). However, this type of information seldom reaches the level of practitioners, especially those who work in schools or in public agencies that deliver educational programs.

In 2016, I began work with two new grant recipients. One evaluation involved an after-school violence prevention program on Long Island, New York, to be implemented in four elementary schools in a district; the other was for the dissemination of an existing school intervention model to a middle school and two high schools in New York City to improve academic outcomes and absentee and truancy rates, as well as to address social needs.

Similarities existed between the programs in terms of goals, although the structures of the programs and activities differed. Both grants dealt with improving outcomes for students, including succeeding in school, enhancing life opportunities, and overcoming environmental factors that lead to participation in violence. In the case of the after-school program, the grant was implemented in four elementary schools in one district for third, fourth, and fifth graders. In the case of the second program, an intervention design previously implemented at a Charter School in New York City was to be disseminated to three public schools in New York City, one middle school and two high schools.

I was contacted after funding had been approved with evaluation designs already incorporated in the grants and was therefore faced with implementing the designs in the approved applications.

In the end, the approach taken for both evaluations was to utilize a case study design (Yin, 1989), to use qualitative data (Bogdan & Biklen, 1989) as well as quantitative data, and to utilize mixed methods and multiple variables (Posavac & Carey, 1989). These choices will become manifest in terms of some of the challenges faced, as explained in this article.

## THE LONG ISLAND ELEMENTARY GRANT PROGRAM

The needs this grant addressed were economic disparity, juvenile/student behavior, and academic achievement. At the time of the grant application, the district ranked #3 in extraordinary need compared with other school districts in the county. The community ranked high on juvenile offenses, persons-in-need of supervision (PINS), and youths placed on probation. Active gangs within the community included MS-13.

The student body was diverse – 71% Hispanic and 21% Black. Behavioral suspensions were more numerous at the middle school level than at the high school level. This grant, among other things, addressed risk factors elementary school students faced when they entered the middle school environment.

Grant writers developed goals to address identified student needs - academic support, violence prevention, and youth development. The program framework was to implement for 218 at-risk elementary students a 20-week after-school program running three days a week for two hours daily for a total of 120 hours of instruction and support, with a student-teacher ratio of 10 to 1.

The program contained after-school homework support, humanities and math instruction, and delivery of curricula by a Community-Based Organization (CBO) and a county-wide law enforcement agency (LEA). The CBO curriculum was a character-development program for third, fourth, and fifth grades, and the LEA curriculum was a life skills program for fourth and fifth graders to help them avoid using delinquent behavior and violence to solve problems. Both curricula addressed bullying, violent behaviors, and gang resistance.

## THE NEW YORK CITY CHARTER SCHOOL DISSEMINATION GRANT PROGRAM

The applicant for this grant was a charter high school in the West Village serving New York City's most vulnerable youth. Fifty percent (50%) of the school's lottery was set aside for homeless or transitionally housed students, students who had been involved in the child welfare system, or had previously attended middle schools where more than half of the students performed below grade level. Demographic data indicated that 51.7% of students were Black, 40.4% were Hispanic, 74.2% lived at the poverty level, with 79.9% having a designated Economic Need Index.

To address academic and social issues, including trauma, homelessness, and incarceration, the Charter School had implemented an evidence-based intervention model that had resulted in improvements in its school culture, attendance, engagement, and achievement.

Grant goals were not articulated in broad statements but in terms of the components of the intervention model previously implemented at the Charter School. From the grant narrative, I was able to extrapolate measurable goals related to disseminating the intervention model; these included impacting school culture; improving attendance, behavior, decision-making, engagement, and achievement; and training staff in positive behavioral and teacher coaching methods. Charter School personnel were to provide training to the three partner schools.

## WHAT CAN HAPPEN WHEN DISCUSSIONS WITH THE EVALUATOR DO NOT OCCUR DURING THE APPLICATION PROCESS

In my experience, grant writers are not trained planners or evaluators. They have limited knowledge of planning models, may not have skills to address "problem anticipation" (Beach, 2020, p. 27), are not aware of factors critical to program evaluation, and throw everything into the "evaluation" mix using a "more is better" approach to secure funding.

What can happen is that grant writers may not neatly align goals, objectives, activities, outcomes, and performance indicators in ways useful to the evaluator. Additionally, they may not anticipate measurement issues, specify data sources, and/or determine whether data will be accessible at all or in a timely manner. Evaluators can be invaluable at the construction phase of a grant in assisting grant writers with all these issues.

Other considerations exist as well. Is the grant being implemented internal to a district or school or external to the funded party/agency: what are the implications of this, what impact may this have on the evaluation process? In the case of the Long Island grant, the grant was implemented internally in the district with a tight oversight structure: the Central Office Administrator oversaw the Program Supervisor who implemented the program who then had oversight over the lead teachers in each of the four elementary schools who then had oversight over the five teachers in each elementary school involved in the program. From an evaluation point of view, this structure was a benefit. On the other hand, there were some issues with data acquisition - assessment tools were mentioned in the grant for which no provisions for development had been discussed or planned.

In the case of the New York City Grant, activities were implemented at three schools external to the applicant school. Unanticipated occurrences brought the activities to a "grinding halt" at the end of the second year of the three-year grant. This was not purely an evaluation issue – but this issue greatly impacted the ability to evaluate this program, especially for the summative three-year comprehensive evaluation. At the point of implementation, two of the three principals from the schools involved in the grant had been transferred to other schools. It took most of the first year to establish a relationship with the two new principals and to get cooperation regarding participation in grant activities. None of the three principals agreed to implement the full model as written into the grant. Resistance continued during the second year of the grant at two of the three schools. In the middle of the second year of the three-year grant, the principal of the applicant school who had written the grant was replaced. At the end of the second year, the full-time Project Coordinator, and the part-time Social Worker both left for other jobs.

By the fall of the third (final) year of the grant, the new principal at the applicant school was not able to re-establish contact via calls or emails with the three schools involved in the grant to re-engage. The principal got the consent of the New York State Education Department (NYSED) to cease activities. Nevertheless, the point-person at NYSED felt that the summative evaluation due after the end of the third year should proceed. If the evaluator had participated in the development of the grant, some challenges this grant posed might have been mitigated, resulting in discussions of "what if" scenarios, most importantly the lack of control the grantee might have had over the entire process.

## ADVANTAGES OF THE EVALUATOR'S INVOLVEMENT IN AN EVALUATION DESIGN

There are distinct advantages of evaluator involvement at the development stage of a grant application. The evaluator can assist with the following so that assessment activities go smoothly:

1) ensuring alignment of goals, objectives, and program activities with outcomes, performance indicators, and assessment activities.
2) reviewing prior to implementation of grant activities whether assessment activities are overly ambitious, redundant, and/or need modification/refinement.
3) determining at the planning level who would be collecting and organizing data for the evaluator, in what form, when, and the means by which the evaluator would receive data;
4) ensuring prior to the implementation of grant activities where sources of data to be used in the evaluation are "housed," and whether data will be accessible given the grant cycle.
5) having the opportunity to have discussions about appropriate assessment techniques, particularly in relation to determining whether tools indicated in the grant will achieve their purposes in measuring specified outcomes, and whether assessment tools are valid and reliable; and
6) determining to what extent measurement tools will have to be developed, whether the grantee has the capacity and expertise to do this, and whether the evaluator has the expertise to do this if the grantee does not.

One issue that arose with one of these two grants was that assessment activities were not always aligned with the goals, objectives, and program activities in ways that would guide data collection to specifically address performance indicators. If an evaluator is involved during grant development, the evaluator can assess if there is alignment between and among goals, objectives, activities, outcomes, performance indicators, evaluation strategies, and the tools proposed to accomplish valid and reliable assessments.

With the Long Island program, there was an attempt to consider assessment activities for the goals in the grant, with performance indicators generally more tightly tied to objectives. Still, there were no indications of who would collect data and by what means data would be accessed/and or developed and delivered to the evaluator. It appeared that the implementers of the program were not totally cognizant of the plethora of indices and data the grant specified would be collected. In order to assist the grantee with alignment issues, I generated a list of data the grant indicated or suggested would be delivered, developing a checklist to indicate what I required, from whom, in what time frames, with various other comments and indications of what would be needed to assess stated outcomes.

With the New York City grant, there was also an attempt to consider assessment activities for each of the goals in the grant. However, there was a "laundry list" approach in the specification of the assessment activities. To demonstrate what needed to be done to tighten up this approach, I created a table to exhibit what it would take to align goals, objectives, program activities, outcomes, and performance indicators with assessment activities. For each objective, the laundry list was reformulated into a checklist so that the implementers could check a box to determine for each objective exactly which data would be collected. This approach was necessary so that planning for time points for data collection could be specified, and implementers could determine whether the types of assessments they intended to use would result in the collection of data targeted to their measurable objectives.

A second issue is related to overly ambitious, redundant, and/or incomplete approaches in specifying assessment activities. This was true for both grants. One example of non-aligned redundancy that resulted in modification involved the Long Island grant and related to the proposed collection of three measures, report card data, State assessment data, and digital data to measure academic outcomes. While some redundancy can be useful in conducting research, offering opportunities for triangulation, nevertheless, in this instance, report card data offered the most valid data for the purposes of this research. The stable factor was that regular classroom teachers were assessing their students involved in the program so that analyses could be conducted in a controlled pretest/posttest design. Also, the purpose of the program was not to measure academic outcomes in terms of normative data. The best assessment of student academic improvement was to examine gains across the school year wherein report cards provided three data points, with one data point at about the point of the implementation of the program, one data point at about the time of the conclusion of the 20-week program, and one data point at the end of the school year (Marzano, 2003). The involvement of the evaluator at the grant development phase could have assisted grant writers in streamlining their outcomes assessment processes and choosing the most appropriate measures for analyzing student outcomes, including academic outcomes.

A third issue that arose with both grants was that the person responsible for collecting and organizing data for the evaluator was not specified, the means by which the evaluator was to receive the data were not indicated, the form in which the data were to be received was not documented, nor was the time frame for receipt of data by the evaluator specified. This third advantage relates to the evaluator's being able to troubleshoot and encourage grant writers to document in the grant how data will be received, from whom, when, and in what form.

A fourth advantage of evaluator participation at the grant development level is that those involved in the grant can assure the evaluator that data will be available given the grant and evaluation cycles and will conform with the specifications of the grant. I have encountered issues with both of the grants described above relative to data not being available in a timely manner. In the case of the Long Island grant, student report cards were used to assess academic outcomes; report cards are normally not available until the second or third week of July. The Central Office Administrator had requested that the evaluation be completed

around the end of June or beginning of July. Assessing academic outcomes addressed one of the substantive goals of the grant. There was no way of getting around this challenge, except to wait for the generation of the report cards. If discussions had been initiated at the grant development phase, it would have been obvious that the evaluation could not have been completed according to the time frame and in the manner intended by the district administrator.

In the case of the New York City grant, due to the inability of implementing the model as outlined in the grant, all quantitative data, except for surveys distributed at training workshops, ended up being derived from external databases. The cause of this was that the grant was not implemented as written and approved and most of the implementation activities outlined in the grant never happened, as explained above.

My only solution was to see what data were available for the three schools in statewide databases that would facilitate assessing the correlates of the grant. After doing online research of available school-level data in New York State, I realized my only hope of rescuing the evaluation strand of this grant was to work with New York City and New York State databases. The data from these databases did align rather well with many of the goals of the program, if not the activities specifically.

Regarding these data, there is a partial "data dump" sometime during the summer, but some data were not available until the second September "data dump." The September data include academic outcomes. Given the grant and evaluation cycles, receipt of the evaluation had been expected by September 1. Academic outcomes are not available until the September "data dump." The Project Coordinator did keep a journal to record real-time events in his interactions with the three New York City schools, which proved to be valuable for narrative purposes. As in the case of the Long Island grant, academic outcomes were a substantive target of the grant so there was a lack of congruence between the annual evaluation deadline and data availability.

A fifth issue related to whether performance indicators are actually measurable by the tools indicated in the grant proposal. A related issue of measurability is whether tools already developed by the grantee that the grantee intends to use in measuring goals, objectives, outcomes, and performance indicators, or tools acquired from other sources are valid and reliable. Many grant writers are not in the position to assess the validity and reliability of tools. Most evaluators should be able to review psychometric tools and protocols to assess whether they are appropriate and psychometrically sound for the project they are evaluating.

Such discussions should have occurred at the level of grant development regarding the best means to measure goals, objectives, outcomes, and performance indicators appropriately. In the case of the New York City grant, there were tools that had been previously developed in-house for the applicant's improvement model. Given the issues that occurred with even getting the project started in two of the three New York City schools and the extent to which dissemination/implementation plans had to be scaled down and modified, these tools were never used, nor was I ever given copies of them to review. For the most part, the performance indicators as documented in the grant were never addressed, and I had to explore other sources of measuring outcomes that I determined were related to the overarching goals of the program.

A sixth issue relates to what extent measurement tools will have to be developed and whether the grantee has the capacity and expertise to do this. Testing and measurement is a discrete aspect of education, and many educators have not had training in nor experience with developing such tools. Educators do generally not take Related courses even at the Master's level. My experience with such

courses occurred only at the doctoral level and only because I had an interest in this area. Regarding the Long Island project, I had to develop four questionnaires (Teacher, Student, Parent, and After-school Violence Prevention Staff) and three interview Protocols (Project Coordinator, CBO, and the LEA). None of these was provided to me. I also had to develop several other questionnaires for the New York City project.

## ADDITIONAL OBSERVATIONS

There are other issues of concern related to grant development and the evaluation component that is part of assessing outcomes of interventions proposed in a grant.

Many grant writers may not be aware of the value of the contributions of an evaluator to the grant development process. In the example of the New York City grant, most of the instruments written into the grant were developed in-house for their model, and the evaluator, short of seeing these instruments, could not judge if instruments were appropriate.

When an evaluator inherits an evaluation design, the evaluator is tied to pre-existing "promises" in the approved application. Under only compelling circumstances can an approved grant be refined or amended after approval. Re-funding of grants can be affected by whether the program was implemented as planned and whether evaluation activities were carried out as written into the grant. Therefore, it is useful and informative for an evaluator to be involved from the ground up.

## ADDITIONAL ISSUES AND CHALLENGES RELATED DIRECTLY TO DATA

The above conversations dealt with what can happen when the evaluator is not part of grant development. However, there were other issues, related to the data themselves. There were multiple issues related to data collection, quality, and adequacy with both of these projects. Such issues can impact the extent to which we have faith in the outcomes of analyses, both quantitative and qualitative.

First, return rates and missing data can be problematic. With the Long Island project, return rate was important to several analyses. There was a Teacher Survey administered at the beginning of and at the end of the program each year to regular classroom teachers of students involved in the after-school program for their observations of student practices and behaviors related to program goals and objectives. A Student Survey was also administered at the beginning of and at the end of the program for all students involved in the program for self-reports regarding their classroom practices and behaviors as well as out-of-school practices and behaviors. In the second year, a concerning issue arose regarding the Teacher Survey in one school when teachers failed to return pretests. This issue was never resolved. Therefore, about 25% of analyses for pretest/posttest results for regular classroom teachers the second year were lost.

The after-school teachers and the parents/caregivers of students in the program were queried at the end of the program about practices and behaviors of the students relative to program objectives. There were no return-rate issues with after-school teachers and parents. There were three "interview" protocols, one each for the Program Supervisor, the point-person from the CBO, and the point-person from the LEA. The return of the Interview Protocol for the Program Supervisor and the CBO was never an issue but was an issue for the LEA.

There was another issue that required exploration and while unrelated strictly to the theme of this article, it was another unanticipated issue I faced and only I could address it. It dealt with the lack of congruence between the results of the Teacher Survey and the Student Survey. Teacher Surveys consistently showed gains among some of the factors assessed. All the factors assessed on

the Teacher Survey were also assessed on the Student Survey. Student results, on the other hand, showed changes in the wrong direction.

When this continued to happen into the third year, I conducted additional analyses to uncover explanations for this phenomenon. First, I revisited all items on each survey for content validity. Secondly, I ran reliabilities on each of the four surveys – teacher, student, after-school staff, and parent. Thirdly, I conducted correlations on aligned pretest Teacher Surveys and pretest Student Surveys and posttest Teacher Surveys and posttest Student Surveys. Results of all the statistical analyses offered no clear explanation for the results, although the reliabilities for the Student Survey in all cases were lower than for the Teacher Surveys. Teacher reliabilities were high, .855 and .910 (pre and post), and student reliabilities were .832 and .927 (pre and post) on in-school behaviors and between .541 and .700 (pre and post) on out-of-school behaviors related to bullying, fighting, and hitting items. Results of all the statistical analyses offered no clear explanation for the directionality issue.

Finally, I did breakdowns by grade level, thinking that there might be an issue by age, with younger students, third graders, having less reliable responses than fourth and fifth graders. When I conducted a breakdown by grade level of students, there was no basis to support the hypothesis that grade level/age was a determiner regarding lower scores on the posttest than on the pretest for the Student Survey.

Then, I researched articles that dealt with the reliability of the results when younger students take surveys. The articles indicated that results from younger students are generally less reliable than those from older students or adults.

I was left with one hypothesis - that the phenomenon overall was the result of the young ages of the students – and specifically due to one possible cause: 1) that the young students enrolled in the program could have had difficulty at the beginning of the program making reliable assessments of their practices and behaviors without any context for evaluating their own behavior and practices and 2) that exposure to the program had changed their understanding so that they were more judicious and realistic in self-reporting their behavior at the point of the posttest. If pretest results were over-evaluated by students and posttest results were more "realistic," then it was possible to get results that indicated no gain or negative differences. I concluded that perhaps teachers should discuss the response scale with the students and assist students in completing the pretest surveys, taking care not to introduce bias into the responses but to assist students in interpreting items and response choices. If some assistance was required for the posttest, teachers could provide this as well.

In the fourth year of the Long Island program, due to COVID and the closure of schools in March of 2020, the program came to a halt, and the posttest surveys for teachers and students and the parent survey were never administered. After-school teachers did complete surveys. No student grades for the last marking period were available for analysis/comparison. The evaluation consisted of reporting data we had for year 2019-2020, conducting retrospective comparisons with the previous three years of the grant, and offering a holistic four-year perspective on the grant overall.

In the case of the Charter School grant, most assessments intended never occurred. The grant was never implemented as intended. However, there was a pot of gold in the "virtual sphere." Luckily, the New York City School Survey and Quality Review had rich data for all New York City schools, including: 1) a school survey completed by teachers, students, and parents; 2) assessment of school factors based on The Framework for Great Schools Model, a research-based paradigm

which focuses on six components of schooling which foster, support, enhance, and optimize the opportunity for academic achievement, including rigorous instruction, collaborative teachers, a supportive environment, effective school leadership, strong family-community ties, and trust; 3) an overall student achievement index; 4) target levels achieved for the six Framework Elements (not meeting, approaching, meeting, and exceeding); 5) academic measures in multiple Core Course areas and Regents Exam areas; and 6) numerous demographics, including student population; racial breakdowns; gender distribution; percent of ESL/ELLs, special needs, poverty, economic need index; annual attendance rates; percent with 90+ percent attendance; percent of chronically absent students; teacher attendance; and percent of four-year graduation rates. Given my dilemma, these databases were lifesavers.

As a result, assessment of outcomes focused on data available in the New York State databases which included annual academic outcomes, demographic and economic data, and the results of parent surveys by individual New York City schools. What became possible in this set of circumstances was that I could go back in these databases two years prior to the implementation of the grant so that comparisons were possible between pre-grant indices/ratings and those for the years of the grant. Since grant activities were suspended the last year of the grant, it was also possible to compare in combination the two years of implementation with the last year in which there were no activities.

These trend comparisons could result in two important findings: 1) overall, had the schools improved during the period of the grant when compared/contrasted with the two years before the grant, and 2) had the schools held any gains the third year of the grant during which time activities were suspended when compared/contrasted with the first two years of the grant when there were some activities in place.

The only disadvantage was an issue already discussed. This total database was not available until sometime in the month of September. However, given this reality, NYSED permitted me to submit my final evaluation based not on the specified deadline but based on availability of data.

Additionally, what I could not determine, nor would ever be able to determine, was the impact of NYSED mandated improvements implemented in the three schools outside the grant. Did they also explain some of the gains uncovered? Were any positive outcomes and gains the result of mandated activities, of grant activities, or of mandates and grant activities in combination?

## ISSUES RELATED TO LACK OF PROVIDING A DESIGN FOR LONG-TERM TRACKING OF STUDENTS

The two evaluations discussed in this paper dealt with issues related to impacting students positively not only for the short term but ultimately for the long term. Both interventions focused on exposing students to opportunities, activities, and reformulation of mindsets and habits that would serve them not only throughout their school experiences but also hopefully into adulthood. Specifically, both grants gave students exposure to opportunities to improve their academic outcomes, to enhance their decision making, to develop cognitive tools to foster social development and nonviolent communicative behavior, to minimize likelihood of dropping out of school, and ultimately to prevent participation in youth crime, violence, and gang activity.

Both grants provided students opportunities to interact in positive ways with caring adults who were tending to "the whole child," cognitively, socially, emotionally, and personally. Both grants also focused on improving student attendance. Additionally, grant activities addressed student

behavior and student engagement, both in school and out of school. Brain-based instructional strategies were employed, social-emotional triggers for students were addressed, and students were introduced to strategies to build resistance to negative peer pressure that they would confront as pre-adolescents (in the case of the Long Island school) and as adolescents (in the case of both the Long Island School and the New York City schools), as well as strategies for developing resilience. Trainers were involved to assist teachers in developing skills that would facilitate positive behavioral changes in students. Teachers were involved in both grants in the change effort. Principals were also involved, either directly or in supportive roles.

These grants were not designed to study a phenomenon, a teaching technique, or the quality of a culture in a moment in time. These grants were designed to change the present and future lives of these students through interventions. In that vein, while the grants were designed to address long-term effects for students, no provision had been set up for tracking students longitudinally. Inasmuch as these grants were designed to improve opportunities and life chances for these students, the schools involved should have also integrated into their planning a mechanism for tracking these students in the long term, at least to the point of high school graduation. Even if it were to be too onerous to track all of them, a representative cohort could have been identified to track a sample of the students

## LESSONS LEARNED

For several years, I have wanted to write about my experiences with these two grants as the evaluator of these programs. I have been conducting evaluations for close to 33 years. In my experience, most practitioners do not have expertise in evaluation processes and techniques. Evaluation is a specialty unto itself. Each evaluation has its own challenges, "personality," and twists and turns. Those who write grants know their needs and purposes. What they often do not have experience with and expertise in is the science of evaluation.

A valuable lesson I learned is that people do matter. I had incredible support from the Central Office Administrator and the Program Supervisor at the Long Island schools for the three years discussed. Likewise, I developed a close working relationship with the Project Coordinator at the New York City Charter School. I worked closely with all three as colleagues, developed instrumentation for them when needed, discussed best ways to address issues, and received feedback regarding my work.

Another lesson learned involved the transition from the contents and rubrics of the approved grant application to the actual work for the evaluator that entailed moving from the original assessment plan (or lack thereof) to appropriate, real-life data generation or acquisition of data to meet specifications as set out in the grant – and the frustrations for the evaluator that this transition created.

In the case of the Long Island grant, I had to develop questionnaires and access student report cards to assess outcomes. The evaluation design for the Long Island grant had challenges, but it was not as unmanageable as the New York City evaluation design.

In the case of the Charter School grant, for each objective there was a list of data-collection methods, with no connection to individual measurable objectives. I was able to address this by reconceptualizing goals and creating a means of linking measurable objectives to performance indicators, even though I had to use external data sources, different from those originally intended, to complete the final evaluation. Luckily, there was alignment between data we would have wanted to collect within the schools and data that were warehoused in New York State databases.

The issues described in this article underscore the need for grant applicants to involve an evaluator in planning at the point of grant development. An experienced evaluator can assist grant developers with alignment of goals, objectives, activities, outcomes, and performances indicators, as well as with issues related to accessing or developing valid and reliable tools for data generation. In other words, an evaluator can serve as a steward for "problem anticipation."

The final lesson learned is that being proactive is better than being put in the position of being reactive. A corollary is that there may be times that evaluators must be prepared to be resourceful when everything is falling apart.

## POSTSCRIPT

Since my completion of the work of the two grant programs discussed so far, I have had the opportunity to work on three new grants, starting in the summer of 2021. Prior to my interview with this agency for consideration to be the evaluator of these grant programs, I shared the paper that I had presented at the ISEP conference in Lisbon in October of 2019 that led to this article.

From the outset, the point people at the agency took a serious look at the issues I had raised in my presentation. I believe that I was hired based on the content of that presentation. From day one, after I was hired to evaluate their grants, I was included in team meetings and in all advisory council and professional development offerings. I have been considered a part of their "team." We have collaborated on all assessment activities, and I have been part of all team planning.

These professionals had a true understanding of the challenges involved in evaluation activities once they understood how evaluations are supposed to be conducted within a grant structure, have supported me in every way possible, and were truly interested in making the most of the processes of assessment and evaluation. Their grants were exemplars, and it was easy to move from goals to objectives to activities to outcomes to performance indicators so that assessment tools could be developed.

They were sensitive to the need to measure what had been "promised" in their grants and made certain that there was fidelity to these promises and that everything that needed to be assessed was assessed in ways that would generate valid, reliable, and useful information. While I was not part of the original discussions regarding what the measurement tools would be for the programs, in the end I was responsible for designing almost all queries and questionnaires. I trust that in the future this agency will bring an evaluator into the planning and design stages of funded programs. I am assured that they are now aware of the value of keeping the evaluator close, from start to finish.

## REFERENCES


Beach, R. H. (2020). A prelude to planning. *Educational Planning, 27*(3*),* 19-30.

Bogdan, R. C., & Biklen, S. K. (1998). *Qualitative research for education: An introduction to theory and methods. 3rd edition.* Allyn and Bacon, Inc.

Bolman, L. G., & Deal, T. E. (1997). *Reframing organizations: Artistry, choice, and leadership. 2nd edition.* Jossey-Bass.

Ferrara, D. L. (2007). The school improvement and transformation system©: A multiple-target planning model to facilitate school improvement and systemically and systematically transform schools into professional learning communities. *Educational Planning*, *16*(1), 18-30.

Marzano, R. J. (2003). *What works in schools: Translating research into action.* ASCD.

Nevo, D. (1986). Conceptualization of educational evaluation: An analytical view of the literature. In E. R. House (Ed.), *New Directions in Educational Evaluation* (pp. 15-29). RoutledgeFalmer.

Posavac, E. J., & Carey, R. G. (1989). *Program evaluation: Methods and case studies. 3rd edition.* Prentice-Hall.

Rossi, P. H., Lipsey, M. W., & Henry, G. T. (2019). *Evaluation: A systematic approach.* SAGE Publication, Inc.

Sever, D. (2021). Program evaluation experts' competencies: A Delphi study. *International Journal of Contemporary Educational Research, 8*(2), 131-142. DOI: https://doi.org/10.33200/ijcer.854511

Volkov, B. B. (2011). Beyond being an evaluator: The multiplicity of roles of the internal evaluator. In B. B. Volkov & M. E. Baron (Eds.), *Internal evaluation in the 21st century. New directions for evaluation*, *132*, (pp. 25-42).

Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2011). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd. ed). SAGE Publications, Inc.

Yin, R. K. (1989). *Case study research: Design and methods.* SAGE Publications, Inc.