

# Understanding and Improving the Effectiveness of After-School Practice

Robert C. Granger

Published online: 18 March 2010  
© Society for Community Research and Action 2010

**Abstract** The question driving research on after-school programs is shifting from “do programs make a difference,” to “why are some programs effective while others are not?” This article draws on the work in this volume and related studies to suggest that there is an emerging consensus on the importance of staff-youth interactions as a determinant of program effectiveness. The commentary recommends that future research should continue to focus on understanding and improving program practices at the point-of-service. In doing so, two lines of inquiry seem promising. The first involves linking measures of changes in staff-youth interactions to changes in developmental outcomes. The second is to examine how policies or other interventions intended to improve program effectiveness affect the practices of line staff.

**Keywords** After-school · Commentary · Future research · Policy · Program improvement

This special issue of the *American Journal of Community Psychology* is a welcome addition to the empirical literature on after-school programs. The editors asked Barton Hirsch and me to provide commentary with an emphasis on implications for policy, practice, and future research. Bart and I divided the articles between us. I will be examining the editors’ opening article and those by Cross, Gottfredson, Wilson, Rorie, and Connell; Durlak, Weissberg, and Pachan; Larson and Walker; Sanderson and Richards; Sheldon, Arbreton, Hopkins, and Grossman; Smith, Peck, Denault, Blazeviski, and Akiva; and Yohalem and Wilson-

Ahlstrom. (The William T. Grant Foundation, where I serve as president, supported the work that led to the articles authored by Durlak, Larson, Smith, Yohalem, and their colleagues. Also, Sheldon and colleagues acknowledge the Foundation’s vice chair—and former chair—Gary Walker, as a reviewer of their article.)

The articles I reviewed reflect a shift in the research question from *if* after-school programs affect important youth outcomes to *why* some programs do so while others do not. The newer work tends to fit into one of two strategies. Some are trying to link in-program practices to measures of youth development. Others are examining how policies or other interventions, such as a change in the approach to in-service staff development, affect program practice. I begin with a discussion of the editors’ conceptual model from the volume’s opening article and then comment on the other articles I reviewed, offering thoughts about future research.

## The Editors’ Conceptual Model

In their introductory article, Durlak, Mahoney, Bohnert, and Perente present a model (see Fig. 1) for how after-school programs fit into an ecology of the forces shaping youth development. Like most developmentalists, the editors see after-school programs as one element in an ecosystem that leads to changes in indices of youth development (or, youth outcomes). For example, young people’s age, gender, and social identity influence the likelihood that they will participate in programs, as do their relationships with family, friends, and other adults. In turn, what occurs in the after-school program influences the system outside the program. One of the strengths of the editors’ model is that it shows these interconnections—after-school programs do not function in a vacuum.

R. C. Granger (✉)  
William T. Grant Foundation, Washington, DC, USA  
e-mail: bgranger@wtgrantfdn.org

Even though this model is a helpful one, any conceptual model has its limits, especially when one is trying to capture something as dynamic as the interplay in how everyday activities in different settings influence youth development. One element that should be added to any ecological model is the broad societal context of secular trends, public policies, and ongoing research findings. The secular trends include changes in immigration, the labor market, the economy, and fertility rates, which, together, shape the characteristics and locations of youth who might attend an after-school program. In other instances, public policies indirectly affect youth development by influencing the elements and connections in the editors' model. For example, health, employment, and income-support policies shape parental employment rates in ways that affect the likelihood of youth participation in after-school activities. Policy decisions can also directly affect programs and activities. Decisions about where to put a new after-school site are likely to shape the probability that youth will attend activities at that location versus at other sites in the neighborhood. Policies regarding after-school program accountability may influence how staff allocate program time. And, new research findings that suggest specific types of interventions, such as a training program for staff, can affect staff practices.

The other modification I might suggest involves the elements in the after-school program box in the editors' figure. Durlak and his co-editors title this element "after-school program features" and list organization/structure, staff characteristics, content/activities, group characteristics, program climate, and implementation as features. No single feature is conceptualized as more important than the others. In our work at the William T. Grant Foundation, we have found it useful to have a structure for categorizing various program elements and practices. We conceptualize settings such as classrooms or after-school programs as discrete systems that consist of particular *resources* (e.g., staff and youth characteristics or materials), in which the resources *are organized in particular ways* (e.g., groupings of kids and the sequence and nature of activities) that *produce certain social processes* (e.g., staff/child interactions) (Tseng and Seidman 2007; Cohen et al. 2003). This framework suggests that while resources and their arrangement are important "inputs," it is the interactions that are central for youth development. A number of the articles in this volume suggest this is true.

Many discussions of the effects of after-school programs focus only on the link between programs and outcomes. The editors' figure reminds us that there are many other forces at play and that any single research study has important limits—much is left unexamined in any effort. The solution is to have a theory or conceptual framework,

such as the editors offer, that you can test and revise against accumulated results over time.

### Articles in the Current Volume

Given the limits of any one study, one approach to advancing our knowledge is to summarize the results from a body of high-quality research. The Durlak, Weissberg, and Pachan article is such a synthetic review. It asks the question, "Do after-school programs that attempt to improve personal and social behavior affect important youth outcomes?" The article updates and expands upon a report that Durlak and Weissberg released through the Collaborative for Academic, Social, and Emotional Learning (CASEL) in 2007. I examined the 2007 report (Granger 2008) and practitioners and advocates have used it as clear evidence that after-school programs affect important youth outcomes (Afterschool Alliance 2008; National Institute on Out-of-School Time 2009). The authors have added new analyses, but the main point remains that the program evaluations that Durlak and colleagues reviewed show that, on average, the programs affected a number of important youth outcomes. However, a subset of programs created large effects, and many programs created no net effect over and above the outcomes shown by a comparison group of youth.

As in the 2007 report, the authors tell us that the best predictor of positive effects was the combination of having a sequenced, active, focused, and explicit (SAFE) program. The group of programs that had the SAFE characteristics created statistically significant impacts in all nine outcome categories while the cluster of programs that did not have all four characteristics had no positive effects.

The current article builds on the prior report by subjecting the SAFE/non-SAFE contrast to further scrutiny. Although the authors did not have enough studies to powerfully test all the questions that follow, their results suggest that there is something important about the constellation of SAFE features. Did SAFE programs appear to have more positive effects because they were more likely to be evaluated with a strong (or weak) design? Did the evaluations of the SAFE programs use more precise outcome measures? Were SAFE programs more likely to include an academic component? Serve a particular grade level of kids? At this time, in all cases, the answer is no. This review, along with a review by Lauer and colleagues (2006) of after-school programs focused on reading and math, led to the conclusion that after-school programs can make a positive difference in the lives of youth. This finding has motivated the shift toward understanding what accounts for the variation in program effectiveness and

what, if anything, can be done to improve the impact of less-effective programs.

### Inside the Program “Black Box”

Because there is little good theory about the in-program processes and mechanisms that produce program effects, it is useful for researchers to closely study and observe program activities. For several years, Reed Larson and his colleagues have been working with data from careful, longitudinal field work from a sample of 12 youth programs in Illinois. The article by Larson and Walker uses that research program to explain that one feature of an effective program may be how well staff handle everyday dilemmas that arise during program activities. Larson and colleagues have developed provisional methods of categorizing such dilemmas and describing how staff respond to these problems in ways that build and sustain youth engagement and other positive outcomes. The work is interesting, innovative, and theory-building.

One limitation of the Larson et al. work is that they selected only programs that were purported to be good quality. Programs with such good word-of-mouth may have attracted a select group of youth, unusually good staff, etc. The authors acknowledge this lack of generalizability and are cautious of making too many claims based on their findings. Despite possible concerns raised by selectivity, the qualitative vignettes from the work by Larson and his colleagues do sound typical of youth work. This is reassuring in that it suggests their findings may generalize to other programs.

While the researchers categorize a majority of the staff responses to dilemmas as youth-centered and productive, they acknowledge that a sizeable minority of the interactions are not productive, though they do not offer much detail about this minority. I am looking forward to learning more as this team exploits this variation in future work. What are the consequences of positive versus less positive interactions? Are individual staff consistent in how they react to dilemmas over time? Is consistency important? How do staff responses change over time and what seems to cause those change? Does training help? Do changes in staff responses tend to generate different youth responses or are youth responses a critical driver of staff?

As work with qualitative data leads to theories about relevant program-level processes and mechanisms, it is necessary to observe how these findings fit with and inform large-scale, quantitative efforts to measure those same processes. The Smith et al. and Yohalem and Wilson-Ahlstrom articles address this issue and support the general observation that staff practices, and the resultant staff-youth interactions, are key.

Smith and colleagues have been working for several years to develop and refine an observational tool that measures how well a program’s activities support youth development. Many practitioners are using this Youth Program Quality Assessment (YPQA) instrument to assess staff practices (The Forum for Youth Investment 2009; Spielberg and Lockaby 2008). The current article uses data from the measurement development work on the YPQA. Using a correlational approach known as cluster analysis, the authors find that six pedagogical clusters capture variations in how staff work with kids. The researchers group these six clusters into three related pairs: “positive youth development style” (28% of the sampled examples of staff practice fell into this pairing); productive “staff centered” pedagogy (39%); and “low quality” pedagogical approaches (33%). Although it is not clear in this article, other work by this team indicates that the style and quality of staff-child interactions vary within programs as much as they do from one program to the next (Smith et al. 2006).

The finding that one-third of the pedagogy observed in this study is of low quality is consistent with the need to understand how to improve staff skills. I am less clear about the implications of the additional finding that these pedagogical clusters had a systematic relationship to particular grade levels and content. The YPQA, per my understanding, attempts to define good staff practices in a way that is not tied to particular content or grade level. In theory, the YPQA scale can separate good practice from bad in an arts program, a general youth development program, a science and math program, etc. In addition, Smith and colleagues are working to develop indicators within their instrument that are appropriate for different age levels so that the YPQA’s definition of good practice is not tied to a particular age of youth.

Here, Smith and colleagues found that the positive youth-centered approach was more indicative of practice with high-school students while positive staff-directed pedagogy was more indicative of work with younger children. Does this reflect anything more than staff using different means to the same end? It is easy to imagine, for example, that all youth need the opportunity to make choices and to be actively engaged, but there may be appropriate differences in how staff provide those opportunities to youth of different ages. Fortunately for those using this tool, the examples of low quality staff practice were not related to a particular age/grade.

Surprisingly, the researchers found that low quality work was more often found in arts activities (as were the positive youth-centered strategies) and less commonly found in life skills activities. I suspect this is more an artifact of this particular sample than a reflection of something durable about the average quality of arts and life skills activities.

The YPQA work is one example of a thoughtful approach to observing youth practice. Reports on the YPQA and other observation instruments have been produced by Nicole Yohalem and Alicia Wilson-Ahlstrom. Their reports include technical reviews of the instruments (Yohalem et al. 2009b) along with information on how practitioners use them. The current article by Yohalem and Wilson-Ahlstrom is a synopsis of these reports; readers are also advised to review the most recent full reports at <http://www.forumforyouthinvestment.org>. The synopsis includes a brief summary of new findings about the relationship between measures of program practices and youth outcomes, an area in which much more work is needed.

The various tools that researchers and practitioners are developing are attempts to measure the program practices that matter for youth. As Yohalem and Wilson-Ahlstrom show in this article and their full reports, the instruments share certain core dimensions. All try to assess program management practices that affect staff-youth interactions, while also measuring staff-youth interactions directly. One potentially important difference among the instruments is that some are more weighted toward program management and structure, while others are more focused on staff-youth interactions. I and others have argued that this difference is important and that instruments should emphasize the point-of-service (Granger et al. 2007; Yohalem et al. 2009a). Currently, this is an untested assumption as these instruments are not well-validated in relation to youth outcomes.

### How Might Research on the Validity of Measures of Program Practices Proceed?

The implicit claim in these instruments is that they measure phenomena that *cause* youth to develop in different ways. The idea is, if programs and staff do more of certain practices, youth development will improve. However, the instruments need more work in order to substantiate the causal claims. In a few instances, the scores from an instrument measuring various staff practices are correlated with concurrent data from youth about their engagement, development, attendance, etc. The cross-sectional correlations are encouraging, but they do not tell us if a change in staff performance leads to a change in youth outcomes. For that, we need measures of change in both areas—a temporal sequence in which the changes in staff behavior precedes the change in youth behavior, and a benchmark of what would have happened to the youth if they had not experienced the change in staff behavior. Meeting the first two needs requires good longitudinal data to improve upon the current cross-sectional analyses. Meeting the third test (what would have happened to youth without the change in

staff behavior) requires experimental tests that change staff behaviors.

### Interventions to Improve Staff Practice

Practitioners and policy makers cannot wait for perfect instruments, and their positive response to the availability of this new generation of observational measures shows that the tools are a welcome aid in their current form. For example, Yohalem and Wilson-Ahlstrom describe a state-wide quality development system in Michigan that uses the YPQA, coupling it with self-assessment and on-site coaching that responds to the data from the assessment. Staff development that consists of an assessment tool, coaching or other development activities tied to assessment data, and then reassessment, is growing in popularity in the field. Researchers are studying such efforts in order to understand the effects of interventions on program practices. The article by Sheldon and colleagues from Public/Private Ventures (P/PV) describes the results from such a study on after-school programs in five California cities. The article is particularly important because there are many current examples of states, cities, and programs using this approach, but very little data on its effectiveness.

As Sheldon and colleagues describe, the Communities Organizing Resources to Advance Learning (CORAL) Initiative in California was designed to test whether after-school programs could have positive effects on student literacy. Six years ago, the team involved with CORAL was dissatisfied with the results and revamped their staff training system to try to improve outcomes. P/PV evaluated the work. The revamped system had initial staff trainings on certain key reading activities, followed by observations by site-level supervisors and counseling for staff on the implementation of the activities.

For two years, P/PV staff administered reading tests to children in a set of the sites in each city in the fall and spring and observed how the staff conducted the literacy activities. They then looked at the relationship between lesson quality and reading gains. The results are encouraging in a number of ways. First, the quality of how well staff implemented the activities improved over time. In the first year, the majority of activities were poorly implemented. However, by the second year, nearly 90% of activities were at least adequately implemented. Although there was no control group to use as a benchmark, this progress provides support for the assessment-coaching-reassessment approach to staff development.

Further, the programs that had more effectively implemented reading activities in the first year saw better gains in measured literacy. To produce this finding, the team averaged the quality rating for implementation of a

program's activities across the year and then related that measure to the change in literacy scores from fall to spring. The finding is consistent with a view that better practice leads to improved student performance. We do not, however, know if change in practice quality over the year was correlated with the change in youth performance. While the data would still be suggestive, the case would be stronger that staff performance was causing improved youth literacy if we saw related improvements in each area.

Cross and colleagues also explore the relationship between quality of program practices and youth outcomes (they use the term “program implementation” for program practices). Although their data are only from five program sites, they found that measures of implementation such as student engagement, program management, and staff turnover were often related to student reports of their experience. The authors need to explore the distinction between a measure of the implementation of a program and measures of student-level phenomena that might be produced by changes in quality of implementation. For example, in this article, student and staff attendance are seen as measures of program implementation. Other researchers see attendance as a proxy for engagement—in effect, a phenomenon that is affected by and affects program implementation. Another issue to consider is whether it is more useful to see an entire program as the unit of interest, or to focus on the variation in quality within a program. While an average measure of program quality may best predict the program's average effects, we know from work by others in this volume that quality varies within programs across different activities. If the goal is to improve program quality, it may be best to assess the variation within a program and work on the activities that show the greatest need for improved practice.

### Family Influences

The article by Sanderson and Richards is unique in that it focuses on families, exploring how events at home might affect youth experiences in after-school. Using data from a survey of youth and parents in three schools (and parents in a fourth school), they assessed youth and parental satisfaction with after-school programs, the program characteristics that youth and adults preferred, and what parents and youth saw as barriers to participation. The authors found that youth were more likely than parents to endorse social and recreational activities such as sports teams and dance, while adults were more likely to look for academics. Barriers to participation included concern for safety, transportation issues, and the need to help at home.

These findings are not surprising; they remind us that what happens outside the after-school program affects what

happens inside. Future work on these issues should consider several things. First, it is productive to contrast the attitudes of youth who participate with those who do not. Sanderson and Richards did this and found some differences. Unfortunately, they could not do the same when they analyzed the responses from parents, due to a problem in the way that they worded a question. Therefore, we do not know how the attitudes of parents who had children in after-school programs differed from those who did not. Second, it would be interesting to understand how attitudes about after-school activities in the same family vary and the possible consequences of that variation. In the Sanderson and Richards article, the parents surveyed were not necessarily the parents of the youth surveyed, and all responses were analyzed individually.

The issue of how to conceptualize the importance of participation seems particularly relevant. Most current research implies that more participation, with deeper engagement, is better. That is likely true *if* one is trying to estimate the effects of a particular program, but it is not necessarily true if one is trying to understand the optimal mix of participation that fosters positive development. For example, extended participation in one program may help young people master that program's activities, but may inhibit their learning in other areas. One tangible example is how family responsibilities are discussed as a “barrier” to participation. It is true that family responsibilities may limit program participation, but such responsibilities potentially carry with them many features that enhance development. From a policy or practice perspective, it seems more important to solve safety, transportation, and cost issues than to reduce the need for youth participation in family chores and other responsibilities—unless those responsibilities raise their own concerns. Such might be the case when older youth are taking care of younger siblings.

### Final Comments on Future Research

The shift in question from *if* to *why* programs make a difference is useful for many reasons. Policy makers and practitioners want information on how to improve program effectiveness, and the shift is likely to advance more fundamental knowledge about how daily activities shape youth development.

As researchers pursue these “why” questions, they should explore what can be learned from two recent trends. Over the past 10 years, there has been considerable expansion in the number of publicly funded after-school programs. Examples include the general growth due to increased federal funding via the 21st Century Community Learning Centers (CCLC) Program, the increase in state-funded programs in California, and the development of

new citywide programs in places such as New York, Chicago, and Boston. It is reasonable to assume that participation in any particular after-school activity is shaped by other activities that exist in the community. Researchers have studied participation patterns (Simpkins et al. 2008; Wimer et al. 2008), but no one has fully assessed how the number, type, and location of programs affects participation. Such information would inform practical questions about how many programs are needed, and where new sites should be located. It also would provide a more general test of how the ecosystem that includes after-school programs functions.

Another trend that could be explored for practical and theoretical considerations is the growth in quality improvement efforts. The general approach of creating a continuous improvement cycle that incorporates an instrument in the on-site assessment of program practices, with on-site staff development tied to the assessment, is becoming much more common (The Forum for Youth Investment 2009). This is encouraging, but we have more to learn about how to do this effectively. For example, the field needs practitioner-friendly assessment instruments that are well-validated (be they observational instruments, logs, or surveys), and an understanding of how to implement staff development in ways that are effective and affordable. These challenges require more studies of practices at the point-of-service, development of good measures of these practices, and testing and refinement of various policies and supports that could help staff improve their work with youth. Research on these efforts can help the field deploy program improvement resources effectively. Such studies also can advance our fundamental knowledge by examining if the elements that make up the after-school program respond to interventions as predicted. For example, if we intervene and change how staff involve youth in activities, do such changes improve youth performance in the ways our conceptual models predict? If so, the findings strengthen the evidence for how the editors and others are conceptualizing the after-school system. If the findings do not support our predictions, it may be time to refine our theories.

In sum, the work on after-school programs has the possibility to improve policy and practice, and, at the same time, allow us to test more theoretical notions regarding how daily environments shape youth development. The articles in this volume are useful steps toward both these goals.

## References

- Afterschool Alliance. (2008). *Evaluations backgrounder: A summary of formal evaluations of after-school programs' impact on behavior, safety, and family life*. Washington, D.C.: Afterschool Alliance.
- Cohen, D., Raudenbush, S., & Ball, D. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, 25, 1–24.
- Durlak, J. A., & Weissberg, R. P. (2007). *The impact of after-school programs that promote personal and social skills*. Chicago, IL: Collaborative for Academic, Social, and Emotional Learning.
- Granger, R. C. (2008). After-school programs and academics: Implications for policy, practice, and research. *Social Policy Report*, 22(2), 3–11.
- Granger, R. C., Durlak, J. A., Yohalem, N., & Reisner, E. (2007). *Improving after-school program quality*. New York, NY: William T. Grant Foundation.
- Lauer, P. A., Akiba, M., Wilkerson, S. B., Apthorp, H. S., Snow, D., & Martin-Glenn, M. L. (2006). Out-of-school time programs: A meta-analysis of effects for at-risk students. *Review of Educational Research*, 76, 275–313.
- National Institute on Out-of-School Time. (2009). *Making the case: A 2009 fact sheet on children and youth in out-of-school time*. Wellesley, MA: National Institute on Out-of-School Time.
- Simpkins, S. D., Eccles, J. S., & Becnel, J. N. (2008). The role of breadth in activity participation and friends in adolescents' adjustment. *Developmental Psychology*, 44, 1081–1094.
- Smith, C., Akiva, T., & Henry, B. (2006). *Quality in the out-of-school time sector: Insights from the youth PQA validation study*. Paper presented at the biennial meeting of the Society for research on adolescence, San Francisco, CA.
- Spielberger, J., & Lockaby, T. (2008). *Palm beach county's prime time initiative: Improving the quality of after-school programs*. Chicago: Chapin Hall Center for Children at the University of Chicago.
- The Forum for Youth Investment. (2009). *Raising the bar: Quality improvement systems for youth programs*. (Out-of-School Time Policy Commentary #15). Washington, D.C.: The Forum for Youth Investment.
- Tseng, V., & Seidman, E. (2007). A systems framework for understanding social settings. *American Journal of Community Psychology*, 39, 217–228.
- Wimer, C., Simpkins, S. D., Dearing, E., Caronongan, P., Bouffard, S., & Weiss, H. B. (2008). Predicting youth out-of-school time participation: Multiple risks and developmental differences. *Merrill-Palmer Quarterly*, 54, 179–207.
- Yohalem, N., Granger, R. C., & Pittman, K. J. (2009a). The quest for quality: Recent development and future directions for the out-of-school time field. In N. Yohalem, R. C. Granger, & K. J. Pittman (Eds.), *Defining and measuring quality in youth programs and classrooms*. *New Directions for Youth Development*, 121, 129–140.
- Yohalem, N., Wilson-Ahlstrom, A., Fischer, S., & Shinn, M. (2009b). *Measuring youth program quality: A guide to assessment tools* (2nd ed.). Washington, D.C.: The Forum for Youth Investment.