

**Confronting “The Big Lie”:
The Need to Reframe Expectations of After-School Programs:**

**Robert Halpern
Erikson Institute
September 2004**

**Preparation of this paper was supported by the Partnership for After-School Education
Views expressed are those of the author.**

Introduction

After-school programs have long had a diffuse and flexible mission. In each historical era, they have been defined in part by providers’ own idiosyncratic visions, in part by broad societal worries about particular groups of children. Over the past decade, one such worry -- inadequate academic achievement levels among low-and moderate-income children -- has stood out, and come to influence key dimensions of the after-school field. It has shaped the expectations of funders and policy-makers, altered (to some degree) the daily work of thousands of such programs, and, finally, strongly influenced where and how evaluators look for after-school program effects.

In this paper I analyze the expectation that after-school programs help boost academic achievement. I argue the urgent need to abandon it, step back, and undertake the basic, grounded research that might yield a more consonant set of expectations, and might shed light on the range and size of program effects for children of different dispositions, ages, and life situations, and for different types and qualities of programs. I argue that a useful program of research will require a perspective that includes consideration of the breadth of developmental tasks of children of different ages, and of those tasks after-school programs are best-suited to help address; sensitivity to the fact that different children need and want different things from after-school programs; and respect for the diversity of programs in the after-school field.¹

The arguments in this paper are rooted in a particular ideological position. In my view, after-school programs are best understood and supported as an historically distinct child development institution, rather than an extension or element of any other; and as a normative developmental support, rather than a vehicle for prevention or remediation of particular social problems. More concretely, after-school programs are well-suited to providing the types and

¹ I assume -- but do not discuss here -- the importance of a complementary program of research addressing the questions of how best to support after-school programs in their efforts to provide good developmental experiences for children, and how best to strengthen the field as a whole. In fact, some would argue that the after-school field is not ready for outcome-focused research at all. I believe that research focused on clarifying appropriate expectations will point to the developmental domains that after-school programs should be attending to, and will thereby help organize and focus the daily work of programs and their staff.

qualities of developmental experiences that other institutions (i.e. the schools and public plays spaces) can no longer provide most low- and moderate-income children. These experiences, whether in the arts, humanities, sciences, civics, physical activity, or other domains, include play and sheer fun, exploration and learning from adults skilled in different domains. They are marked by respect for children's individuality, by learning and producing through collaboration and mutual assistance, a measure of choice and control by children, activity that uses all the senses and symbolic systems, adult feedback that is focused on the learning process and tasks at hand and includes recognition for tasks well done. After-school experiences nurture such capacities and dimensions of self as creativity, aesthetic sense, growing skill in specific domains, self-expression, interpersonal skill, sense of agency and voice, identification with home and community culture, individuality and relatedness, compassion, and physical vitality. It is in domains such as these -- there are many more, to be discussed -- that we would begin the gradual, difficult process of identifying and developing measures of program effects. And we would engage in this process not to determine if after-school programs deserve public funding, but to clarify what they should be supported in trying to accomplish.

Background

For most of their 125 year history, after-school programs remained a modest institution, at the margins of social provision. They were not unimportant to the children and communities served, and in fact provided memorable, defining experiences to some of those children (Halpern, 2003). Men and women involved in after-school provision took their responsibilities seriously, and argued seriously about purpose, philosophy, and role in children's lives. But, to the extent that they took notice, elected officials and the public viewed after-school programs as a community institution and a component of local philanthropic efforts. Beginning in the late 1970s, after-school programs benefited modestly from new public subsidies for child care, through Title XX, Community Development and Social Service Block Grants, and later through Child Care and Development Block Grants. Even so, they remained largely outside of national awareness and public debate.

This changed during the 1990s. After-school programs became more widely recognized and promoted, simultaneously putting them in a position to compete for additional public resources and requiring them to fight for a share of such resources. Longstanding providers were caught off-guard by the rapid pace of events in their field. Philosophically, they were inclined to continue arguing for after-school programs in broad developmental terms. But they also knew that a meaningful share of scarce resources would not be secured by arguing that low- and moderate-income children deserved the same access to fun, enrichment and challenge as their more advantaged peers. These traditional providers were nonetheless too diverse, decentralized and perhaps inexperienced in public advocacy to come together to develop the simple, resonant, problem-oriented story-line demanded of a public issue in American life.

Meanwhile, new constituencies were discovering the after-school field, and quickly “defining” it in relation to their own concerns. Most new proponents had little knowledge of the field and its history, nor much inclination to consult those organizations and individuals with years of experience in providing after-school programs. And (though delinquency, drug use and related social problems had brief tenures as defining concerns) most new proponents were preoccupied with one particular issue -- the academic achievement gap between more and less advantaged children and among different racial and ethnic groups. During the 1990s elected officials at all levels of government latched on to this issue and made it a central social problem of American society. Consequently, the over-riding argument that emerged -- in political if not public consciousness -- for support of after-school programs was that they offered potential to boost children’s academic achievement. Lip service was given to child care needs and broader developmental aims, but success or effectiveness were defined largely in terms of academic goals.²

As noted at the outset, after-school providers had always been responsive to larger social preoccupations (see Halpern, 2003). But external pressures in the 1990s were qualitatively different. Public and private funders compelled after-school providers to make promises about academic effects that those providers knew were unrealistic. (One provider at a recent after-school meeting called the perceived need to make unrealistic promises -- which continues to the present - - “the big lie”.) New public and private initiatives were creating a broader base for after-school programming in the schools, making providers further susceptible to school-related agendas.

The emphasis on academic outcomes has continued in the past half decade. The 2005 Performance Plan for the 21st Century Community Learning Centers, the largest federal after-school initiative, requires that local programs demonstrate year-over-year gains in academic achievement for children served (i.e. gains over and above those children would have made absent participation). Virtually every one of the many new public initiatives in states and cities throughout the country are justified by the need to improve academic achievement. For example, in California’s After-School Learning and Safe Neighborhoods Partnership Program, local programs will have to be re-certified every three years, based on attendance goals and children’s progress on standardized tests. Linked to this narrow outcome focus is a growing tendency to view after-school programs through the kind of social engineering lens prominent in the 1960s and 1970s. Funders want to know the optimal timing, intensity, duration, breadth and target populations for demonstrating effects. Is one year too little? How many kinds of activities per

² One irony here was that, when testing pressures forced school systems to reduce time during the school day for arts, physical education and other activities not deemed to contribute directly to higher test scores, after-school programs, whether or school- or community-based, were implicitly asked to take on the role of providing such non-test-related activities. After-school funders then turned around and told after-school providers that they would be judged by their effectiveness in contributing to higher test scores, forcing them to consider reducing time devoted to arts, physical activities, etc.

week suffice? ³ Should program slots be “held aside” for high risk children or adolescents? Public and private funders seek “promising” or “proven” models to replicate, and “outcome-driven” or “results-oriented” organizations and systems.

Expectations Translated: The Current State of Evaluation in the After-School Field

Evaluating the effects of social programs is inherently challenging, regardless of the field involved. Common problems include lack of clarity about the theoretical underpinning of a particular program approach, challenges in conceptualizing and measuring the “treatment” received, uneven implementation, necessary compromises in research design, attrition (or other changes over time) in treatment and comparison groups, and challenges in choosing the right “outcomes” to measure and measuring them adequately. Evaluators are sometimes constrained by the specifications in requests for proposals, sometimes by their own lack of knowledge of a particular type of program, and sometimes by issues of timing (programs not only need time to mature, but they evolve even after reaching maturity). It is not uncommon for evaluators to have to propose a very elaborate plan before gaining first-hand knowledge of the programs they are proposing to evaluate, and then to have little freedom to refine or alter their approach as they learn more about those programs.

Additional challenges arise from the unpredictable interaction of particular developmental settings with each person’s unique combination of capacities, vulnerabilities, predispositions, interests, and history: Individuals differ in what they bring to particular settings, how they experience them, and what they take (and get) from them. Discreet experiences are integrated into complex, evolving selves in ways that are poorly understood, difficult to partial out and difficult to measure. The influence of any one developmental setting is commingled with that of other settings. (Indeed what individuals can do -- and how well they are doing -- at any point in time in their lives is multiply determined, by biology, cumulative experience and in the present by the multiple people and settings of their daily lives.) And, especially for children and adolescents, important effects of any discreet set of experiences may become apparent only over time.⁴

³ The confusion -- and confused thinking -- surrounding participation is well-illustrated by a recent review (Chaput, 2004) that argues for the importance of “breadth” of participation. The author argues that studies she reviewed suggest that “participation in a variety of out-of-school-time activities, either within a single program or in the course of week, is associated with beneficial youth outcomes . . . youth need to participate in multiple activities within a program to maximize the benefits of participation” (pp. 3, 29). The author ignores, among other things, the exact tendency for after-school programs to involve children superficially in a wide range of short-lived activities, the value of getting into activities or projects in depth, children’s frequent preference for doing so, especially as they get older, and the importance of thematic or single-focus programs to the field, among other factors.

⁴ For all the reasons noted above, evidence from widely varying fields suggests that discreet social programs have inherently modest effects. The evidence is surprisingly consistent across such

A Distinct Set of Challenges for After-School Evaluators

Efforts to evaluate after-school programs have been subject to most of the difficulties noted above, expressed in forms that are characteristic of the field. For example, when pressed, most frontline providers in after-school programs can offer only the most general statements of what they are trying to accomplish, and why they do what they do (i.e. the theoretical rationale for the over-all design, particular use of time, daily and weekly content of activities in their programs, staff skills and qualities sought, framework for adult-child relationships in their program, etc.). Providers and programs included in a typical multi-site study often vary in mission, focus, emphasis, and structure. The nature and quality of children's experiences in after-school programs are widely variable, even in initiatives with some focus on program quality.⁵

Patterns of participation in specific after-school programs are often variable as well, for reasons that are both planned and accidental, explainable and not. The wide age range of children participating in a typical program creates a variety of distinct complications. A package of good developmental experiences for a seven or eight year old will be different than that for an eleven or twelve year old. As children grow older they tend to participate in particular after-school programs differently, for example, attending only or primarily when there is an activity that interests them. Even within programs, activities for older children are usually different in structure and emphasis than those for younger children.

While children are affected in different ways and to different degrees from any seemingly common developmental support, this variability may be especially prominent in after-school programs. Experiences during the school day affect children differently, and they want and need different things after school, psychologically, physically, cognitively and socially. In many after-school programs, children have a measure of choice in the focus and degree of engagement. Children of different ages bring different capacities to shape their own developmental experiences. (Children are better at using resources to their own ends than adults sometimes think they are.) More generally, children's out-of-school lives are complex puzzles, with many pieces;

fields as parent support and education, preschool education, compensatory education (Title I), youth employment, preventive mental health, drug and delinquency efforts . . .

⁵ Although there is widespread agreement that program quality is a problem in the after-school field, there is less agreement on the exact nature of this problem. While there are a number of obvious constraints to program quality, including inadequate funding, inadequate staffing (and high levels of staff turnover), lack of intentionality and clarity about purpose in many programs, and program isolation, the quality problem is confounded with other issues facing the field. These include lack of agreement about expectations and understandings of what after-school programs should be about (the subject of this paper); the decentralized and heterogeneous nature of the field, with many kinds and sizes of sponsors; and the difficult, resource starved community contexts in which many of the children served by after-school programs grow up. For a full discussion of the issue of program quality, see Halpern (2003, ch. 5).

any discreet programmatic experience may be a modest piece of that puzzle.

In other words, deciding where, how, when and even whether to look for program effects is particularly challenging in the after-school field, given so many sources of variability -- in underlying assumptions, program emphases, program quality, patterns of participation, in the ages of children involved, in individual needs and preferences, therefore in the ways in which individual children use after-school program experiences. *One would anticipate that a significant (and illuminating) part of the story told by outcome evaluations in the after-school field would be about the challenges evaluators have faced in figuring out what to measure, how and when to do so, and in what design to use to capture so many kinds of variability and "individuality". But that has not been the case.*

Sidetracked by the Wrong Focus

Unlike neighboring fields such as public education and early childhood care and education, the after-school field has a limited history of applied research and evaluation. The 1990s brought some initial research activity, by academic researchers and contract evaluators. The first generation of research, strongly developmentally oriented, created promising ground for a longer-term research program. But it was soon side-tracked by the instrumental, and especially the academic, pressure within the after-school field. This pressure strongly influenced the focus of a number of major evaluation studies, compelling contract evaluators to devote inordinate time and energy to a search for academic outcomes. Notable among these are the studies of L.A.'s Best (Huang, 2004), the 21st Century Community Learning Centers (Dynarski et al, 2003), TASC (The After-School Corporation; Walsh et al, 2002) and the San Francisco Beacons (Walker & Arbreton, 2004).

Evaluators of major initiatives have chosen not to -- or felt they did not have the liberty to -- begin their work with a period (at least a year, preferably two) of intensive, qualitative field work aimed at inductively uncovering developmental domains in which program effects appeared to be occurring. (This lack has been especially problematic in initiatives that serve as funding streams rather than as purveyors of a particular program model.) Even when evaluators have been attuned to the broad developmental terrain in which programs were operating, as is the case with the San Francisco Beacons, they appear to have felt compelled to focus on those outcomes insisted upon by sponsors and funders (Walker & Arbreton, 2004, pp. 2, 72). In most instances, this has meant adoption of enhanced academic achievement as a central outcome and use of standardized tests of reading and math as a central measure to determine if that outcome has been achieved.

Evaluators generally have been forced to make a virtue of necessity, using tests administered by local public school systems as part of their routine testing at specific grade levels. In other words, they have worked with measures selected for -- and data collected for -- other purposes, purposes having nothing to do with the after-school programs they are evaluating.

Designs have varied, but have mostly been quasi-experimental, using matched comparison groups of various kinds. Analytic strategies have also varied, but have tended to use sophisticated statistical techniques to examine whether, in aggregate, changes in program participants' test scores over two or more points in time differed from either or both what would have been expected for those children in the absence of treatment (predicted trajectories) or actual patterns of change in comparison children.

Findings of after-school program effects on academic achievement (as measured by test scores) have been extremely modest, at best. As has been widely noted, the first round of outcome research for the 21st Century Community Learning Centers found no program-favoring academic effects for samples of elementary and middle school students (Dynarski et al, 2003). The evaluation of TASC found no effects on reading achievement test scores, and on math achievement a negative effect size of .08 after one year of participation, a positive effect size of .12 after two years of participation, and a positive effect size of .06 after three years of participation (Welsh, et al, 2002, pp. 25, 27; .2 is considered a small effect size). The San Francisco Beacons evaluation found no program effects on grades or academic achievement (Walker & Arbretton, 2004, pp. 75-76, 116-117). Likewise the L.A.'s best evaluation found no program effects on academic achievement in English language arts or math (Huang, 2004, pp. 38-44).

There is some evidence in these studies (the 21st Century CLC evaluation being an exception) of a relationship between greater participation (regularity, duration) and slightly stronger academic effects. This evidence derives from analytic strategies that are creative, but exceedingly complex, loaded with contingencies and assumptions, and thus tending to overwhelm the delicate structure of findings. While it is often very difficult to ascertain how many children are included in which analyses and which findings, there seems to be a pattern of shrinking samples, both in pre-post analyses and in much of the complex statistical modeling performed by evaluators. (Some of this is certainly due to program attrition, some to missing data, some to a need to fit children to particular profiles and rules, e.g. minimum number of days in attendance.) It appears, for example, that analytic requirements shrank sample sizes for some analyses in the TASC evaluation by a factor of ten or more. It is very difficult to ascertain how particular sub-groups of treatment children may have been different than the larger groups from which they are drawn, and how they differed as well from comparison children. (Comparability is a general problem, regardless. In a secondary review/analysis of the 21st Century Community Learning Center report, Kane, 2003, p. 9, notes of the middle school study that "participants and non-participants who were matched to be similar on an initial set of variables were subsequently revealed to be quite different when additional measures from the parent and school data bases became available.")⁶

⁶ The reports of these studies are full of interesting statements that are at best difficult to interpret, but that raise questions about designs and measures. To cite just a few examples, Huang (2004, p.

It is important to note that effects on tests of reading and math were not the only program effects examined by evaluators in these large, signal studies. Through self-report on surveys, focus groups and other qualitative methods, evaluators examined such variables as sense of self-efficacy, sensitivity to others' perspectives (Beacons), conflict resolution, cooperating with others, and decision-making (L.A.'s Best). Taken as a whole, the qualitative data from these large studies suggests a pattern of modest, mostly positive effects. (But even in these domains, most outcome measures have been pre-selected, rather than refined from a more open look at children's experiences and conversations with program staff.)

What specifically is problematic about the idea that academic achievement is an appropriate outcome for after-school programs, and the subsequent use of standardized achievement tests to assess that outcome? In the first place, academic achievement is a product or consequence of schooling, and after-school programs are neither schools nor extensions of schooling. Regardless of where they happen to be located, after-school programs are a distinct child development institution, one with its own history, logic, social role, and place in the broader human service landscape. (Conversely, academic remediation that happens to take place after school is something schools do, not something that after-school programs do.) In a world in which the concept of "academics" was fundamentally broadened to focus on many dimensions of development and include all kinds of learning experiences, the work of after-school programs might be considered "academic". But that world, if it exists at all, is receding.

Even if a funder or policy maker, understandably preoccupied with the academic achievement gap between groups of children in American society, wanted to focus all possible

19) found that L.A.'s Best participants participated in basically the same types and amounts of extracurricular activity when they were at and when they were not at the program. Huang (2004, p. vii) notes that, on a baseline survey, most program participants reported "being able to do well in school and able to finish all their homework on time. They studied hard for tests and always tried to do well in school." (She also notes that program participants "started very high" on self-reported social development, p. vii.) If this is an accurate portrayal of participants, then they would seem a quite unusual group of inner-city children. Yet L.A.'s Best seems to have been serving a cross-section of children, and would one have expected a good deal of variability in most domains. The TASC evaluators report that "students who were active participants in at least one of the years they were enrolled in a TASC after-school project showed significantly greater gains in mathematics test scores after two years of participation, compared to similar non-participants" (Welsh et al, p. 26). But who would be a "similar non-participant" to children who displayed such a particular, distinctive pattern of TASC participation? In attempting to control (or account) for potentially important unmeasured differences in characteristics between TASC participants and comparison children, the TASC evaluator included "prior-year test scores in the equation for predicting expected gains on test scores and [used] prior year school attendance" (Welsh et al, p. 23). Yet this strategy risked further confounding findings and adding another loop of circularity to what was already a design full of such complications.

resources on that problem, after-school programs would be low on any list of solutions, well below reduction of social and economic inequality, reduction of various forms of segregation (especially housing segregation), improving the quality of housing and health care for low- and moderate-income families, changing the demands on less skilled work so parents could be more available to their children, improving the conditions of urban schools (renewing physical plants, reducing over-crowding, reducing class size, strengthening non-classroom resources, etc.), strengthening instructional processes, and generally making schools more positive teaching and learning environments, environments characterized by what Tharp (1993) calls “jointness”. Put differently, the loss of what after-school programs can do well would not be worth the minute gain to be had by diverting them to this critical societal goal.

With respect to the specific issue of standardized achievement tests as an outcome measure for after-school programs, most of the experiences children have in most programs, beyond doing homework, have little or nothing to do with the narrow, specific, disembodied skills measured on such tests. Take, for example, a group of middle school children who spend a year working on a mural. How is what they learn from that experience -- about design, planning, drawing, color, form, working as part of a team, about particular cultures, their own heritage, social history, public art, etc. -- and the ways in which they personally grow through it, going to be captured on a standardized achievement test in reading or math? Take a group of children who participate in an African drumming class once a week throughout the year. How are the musical, rhythmic, kinesthetic, narrative, cultural, creative, interpersonal, and other dimensions of that experience captured on a standardized achievement test?

Take a shy child who makes a few friends, and begins to learn to hold his or her own in social situations; an English language learner (required to spend each day in a school environment that ignores, and may even prohibit, expression in his or her first language) who has opportunity to explore the traditions of his or country of origin and perhaps read literature in his or her native language; a child struggling with self-regulation, who learns through martial arts how to calm and “center” him or herself; a child who had no idea that he or she had something to say; a child whose parents are not able to help celebrate a good grade in school. One can substitute for these examples any of dozens of varied developmental experiences, types of projects or ongoing activities, from dance to photography, from ecology to urban study, from informal play to organized sports, and any of scores of individual child profiles, and ask how what is learned or acquired might be captured on standardized tests.

Even holding aside their lack of validity in capturing most of the possible effects of after-school experiences, achievement tests are inherently limited measures of children’s learning and growth, whether in school or out. They ignore important domains of learning (e.g. literature, the natural sciences, the social sciences, the arts). They ignore many important cognitive skills (e.g. synthesizing information, applying knowledge, using what one knows creatively) and social goals of schooling (e.g. perseverance, self-discipline, communication, social responsibility). They

ignore important attributes and skills in the domains they do measure (e.g. whether reading is becoming a habit, whether a child enjoys writing). Standardized tests often fail to evoke children's optimal performance in a domain (they do not allow children to express/demonstrate what they know and/or can do); they are susceptible to children's unfamiliarity with the format and/or content (put differently, responding to multiple choice questions is an unfamiliar way for children to express what they've learned); and evident skill or performance is partly, if not largely, context-dependent. Time pressure, anxiety and novelty affect performance, and some children are simply not good test-takers. Standardized tests are particularly poor measures of knowledge or skill for children of color and English language learners. They are a one-shot, high-stakes measure -- if a child is tired, distracted, worried, not feeling well on test day, it can strongly affect test scores. Finally, the meaningfulness of standardized tests as even narrow and partial measures of learning are increasingly doubtful as more teachers spend significant amounts of time on test preparation, and more children take test preparation classes from private companies hired by school districts.⁷

The most common approach to use of standardized tests to measure after-school program effects -- the departure of "treatment group" children from predicted trajectories (i.e. test score gains) over time, whether their own or those of comparison children -- is also problematic. Meaningfully altering children's "predicted" growth or trajectory over time on a standardized test score is difficult even for powerful educational interventions whose central goal is to do so. (Most after-school programs serve a wide cross-section of low- and moderate-income children, in terms of academic risk. Even if after-school programs were considered to be a form of academic remediation -- which of course they are not -- one would not expect test score effects for children who are already faring decently.) Reliably charting children's predicted trajectory is, at any rate, an uncertain art, complicated by changes in tests and test items from year to year, changes in how test data are analyzed and reported, differences in growth rates as children grow older, the shifting composition of target groups of children (due to residential mobility, changes in school populations, changing rules about who takes tests) and a variety of other confounding factors. (On a minor note, standardized testing tends to be concentrated at specific grades. There may thus be little or no test data on some after-school participants at any point in time.)

Finally, some have argued that even if we accept the after-school program as a distinct institution with its own purposes, and acknowledge that the activities of after-school programs are unrelated to skills measured on standardized achievement tests, the knowledge and skills acquired in after-school programs should "transfer" to performance in school and even to performance on

⁷ Delandshare (2002, p. 1478) argues that "current educational assessment is for the most part a relic of the past. It has limited purposes and methods that generate limited data; it is based on poorly articulated, ad hoc theories and assumptions of learning and knowing; examinees submit to the process without active and equal participation (e.g. critique, reflection, self-reflection); and secrecy, reward and punishment are still key concepts."

achievement tests. The nature, amount and conditions necessary for such transfer remains an open question. General evidence suggests that knowledge, skills, principles, and strategies developed in one setting and/or one particular domain do not transfer straightforwardly to others. (Lave & Wenger, 1991, p. 33, note that “even so called general knowledge only has power in specific circumstances.”) Each developmental setting has a distinct motivational structure, locus of control in problem-solving, degree of abstraction in learning tasks, and type of teaching-learning relationship. There are also individual differences in capacity to transfer knowledge and skill from one setting to another (Resnick, 1989).

An Alternative Approach to Framing Expectations, Discerning and Measuring Effects

The challenges to evaluating social programs in general and after-school programs in particular, and the invalidity of standardized tests as measures of after-school program effects, are not cause for pessimism. Lack of evidence of program effects on standardized achievement test scores says nothing about after-school programs -- their benefits and limitations, their role in children’s lives, the conceptual and practical challenges faced by the after-school field. The principal result of the focus on such tests has been *to delay the necessary work of finding appropriate ways to define expectations and measure effects, and to use evaluation activity to help program staff reflect on and as necessary refine their work with children.*

Defining appropriate expectations of after-school programs and figuring out where to look for program effects require a respect for complexity, and a process that is at once deductive (top-down, outside-in) and inductive (bottom-up, inside-out). The deductive part of the process requires, first, consideration of the full breadth of developmental tasks for children of different ages, growing up in different communities and, second, an effort to define the distinctive qualities of after-school programs as developmental settings, and therefore their distinct role in supporting children’s work on various developmental tasks. At the same time, in any actual evaluation study, the “narrowing” process has to accommodate three complicating variables: the possibly wide age range of participating children, individual differences in what children within any narrow age range want, need and get from their after-school program experience, and the diversity of programs in the after-school field, in terms of purpose, focus, and resources.

The inductive part of the process of defining expectations and, especially, deciding where to look for program effects, requires the evaluator to look carefully and in depth at the experiences children have in the after-school programs he or she evaluating -- who the children are that are being served, how time is used, how language is used, what, exactly, happens during particular activities, teaching and learning processes, types and qualities of engagement, distribution of responsibilities, relationships between adults and children and between children, how problems are handled, over-all climate, etc. The evaluator then has to marry the two perspectives: to embed what he or she sees in the larger frame of developmental tasks and after-school program roles.

On the following pages, I begin the deductive process that I believe necessary for the field, concentrating primarily on the first two elements noted above. They provide the foundation, as it were, for a new program of research.⁸

Starting With Developmental Tasks

In considering where to look for after-school program effects, it makes sense to begin by laying out the range of developmental tasks that preoccupy children of different ages. These “tasks” are generated both internally, by children’s growing physical, cognitive, social and other capacities, by their interests, concerns, and inner drives, and externally, by family, community and broader cultural demands (Silbereisen & Eyeforth, 1986; The idea of developmental tasks originated with Robert Havighurst [1953]. Other terms that have been used to capture this idea include salient developmental issues [Sroufe, 1979], developmental challenges [Peterson, xxxx], life tasks [Dittman-Kohli, 1986], and personal projects [Little, 1983]. Erikson [1950,1968], of course, also identified key tasks at each stage of development, which he variously called achievements, accomplishments or developmental crises.)⁹

The developmental tasks most relevant to after-school programs, naturally, are those of middle childhood and adolescence. The former is marked by shifts in cognitive and relational capacities, participation in new institutions whose demands children have to adapt to and master, and generally increased participation in the broader social and cultural world. Adolescence brings another significant advance in personal (and interpersonal) capacities, combined with the numerous biological and psychosocial changes of puberty, and the need to begin planning for the demands of adulthood.

Erikson (1968) characterized middle childhood as the age of industry. During this age period children begin concerted work on the tasks seen by their culture as important for effective adulthood (Weisner, 1996). In the U.S., those that stand out at first seem school-related: acquiring literacy and numeracy, developing verbal facility, developing general analytic skills, beginning to

⁸ In effect, what I am trying to do here is model the thinking that I believe needs to occur if after-school research is to be more fruitful than it has been to the present.

⁹ Until mid- to late-adolescence, with the rapid growth of self-awareness, children are not conscious of working on developmental tasks (although they can state immediate goals, such as learning to get somewhere alone). Children do become increasingly conscious of being evaluated -- by adults and peers alike -- for their competence and performance (Eccles, *Future of Children*, p. 32), and in that sense are aware of the need “to work on” what is culturally valued. More generally, developmental processes as such tend to be subterranean, or at least not transparent (Silbereisen & Eyeforth, 1986). And, as children get older, such processes have “no guaranteed direction” (Mitchell, 1998, p. xiv). As Kroger (1996, p. 174) puts it, “there is no map when it comes to matters of maturing”.

understand different knowledge domains and disciplines (and beginning to learn to think in ways characteristic of those disciplines; Sternberg, 2003, p. 5). Children also have to learn to function effectively in school as a developmental setting. They have to learn to work with “the artifacts” of school -- tests, textbooks, text book assignments (Bereiter, 2002). They have to develop a range of capacities and skills central to success in school: recognizing and regulating emotions and behavior, paying attention, persisting in abstract tasks, seeking assistance, being punctual, learning time discipline, coping with competition, coping with large groups, following adult orders without always understanding the reason, and deciphering teachers’ expectations.

There are, nonetheless, many dimensions of children’s developing selves that cannot be tied so closely -- or at least exclusively -- to the demands and opportunities found in school. Some of these involve capacities that teachers (and other school staff) lack the time or incentive to help nurture. Some involve tasks that may be more central or find different expression (or take different form) outside of school life. And some involve qualities of self that may even be antithetical to what is needed or nurtured in school, or that school experience actively undermines in particular groups of children.

Close to home, children are beginning the long-term task of “individuating” and, in some contexts, beginning to assume responsibilities for family maintenance. In the social or interpersonal domains, outside the family, children are learning to develop and maintain friendships, and generally manage social relations with peers. With both peers and adults in different settings they are learning to reciprocate, to negotiate, to influence others in appropriate ways; learning to act and function independently, to be more self-reliant (and less parent-reliant) in making decisions and solving problems. They are learning to listen to and evaluate others’ perspectives (Palincsar & Herrenkohl, 2002), to give and receive feedback, “to reflect upon [their] own interests and coordinate them with those of others” (Kroger, 1996, p. 11). Older children begin to seek a balance between connectedness and autonomy, and to learn to use solitude -- time alone -- in constructive ways, for emotional renewal, planning, re-grouping, processing, and reflecting (Bucholz, 1997).

Middle childhood is the time in which the internalized experiences of early childhood -- the way in which young children were treated inside their families, as well as in other settings -- begin to “externalize” as a distinctive capacity for empathy, kindness, concern for others (and, it should be noted, a distinctive sense of humor). Formative experience, combined with growing cognitive capacities, also shape the growing capacity for moral reasoning and development of a moral compass. On a broader canvas, children are beginning to define themselves in relation to other children and developing a better understanding of social structure and social processes, of where they, their family and community fit in the larger society. Kagan (1984, p. 140) notes that once children begin to recognize the categories to which they belong, they also begin to “assume that they should match their qualities to the proper category.”

In middle childhood, children extend the settings and ways in which they test their sense

of agency. Children are learning to master the larger community environment in which they live -- to explore, to map it out in their heads, to use community resources. They begin to acquire, to play with, and later to work on becoming more competent in using, various tools for self-expression -- language, print, image, movement, etc. (They are, in other words, beginning to master the principles, techniques and “artifacts” of different symbolic systems.). They begin to learn how to express preferences and set limits on others.

Children are also beginning to channel creative impulses into culturally valued forms, especially through the arts. Developmental work in creative and artistic domains encompasses a range of new capacities. Children are beginning to engage the arts as “maker, critic and audience” (Greene, 1991). They are beginning to understand the creative and productive processes in the arts, e.g., where a work of art comes from, the background knowledge needed, the preparatory work involved, the stages and process through which it is created. Some children are ready to choose a particular art form to work at seriously. Such children will begin to learn about the many tasks that accompany commitment, for example, how to practice, and to make use of practice for personal and creative ends. In general, sources of creativity begin to change in middle childhood, with the unselfconscious creativity of early childhood beginning to be reigned in by caution, worry about criticism. In that sense, as they grow older children need to re-learn creativity -- what to draw on, how to take appropriate artistic (i.e. expressive) risks. In terms of the aesthetics -- the substance -- of creativity, it increasingly requires some education (Greene, 1991, p. 31, writes of the need “to acknowledge that imagination and the emotions including taste and sensibility, can be and ought to be educated”).

An increasingly important task for children in U. S. society is maintaining physical vitality. In the later years of middle childhood and into early adolescence, physical activity not only declines in absolute amount, but is shaped by different factors. As with creativity, what was natural and instinctive has, in many respects, to be re-learned and re-incorporated (Halpern, 2003). Being physically active becomes a matter of social learning, and an element of identity development, as children look (and listen) to others as models of who and what they might be and how they should engage the world; to parents, relatives, and other adults in the community, to siblings and friends, and, increasingly, to popular culture. In a loose sense, physical vitality is at the core of a general sense of vitality in facing the world.

Differentiating, By Age and Context: The centrality and expression of the tasks described above naturally change as children grow older. Family demands gradually increase. Children have a growing desire to make a contribution to their family and community. As children move toward and into adolescence, self awareness deepens, self appraisal becomes more accurate, views of self become more complex, and self differentiation from others (including parents) becomes clearer. Interests gradually become defined, commitments solidify and become more consistent. Children begin to imagine the future in more concrete terms. By mid-adolescence, issues of identity become prominent: deciding what is important to do, whom to identify with, learning

what one is interested in and good at, deciding what one is for and against, determining who and what one might become. Adolescents try to figure out what their place is, where they fit, and who might accept them -- in family, peer group, school, community, the economy and society.

Developmental tasks are shaped as well by the contexts in which children grow up. Growing up poor, in a devalued group, in a devalued and neglected neighborhood, for instance, complicates childrens' and adolescents' efforts to address the tasks of both age periods. It alters the meaning of almost every activity, from mastering school-related demands to learning friendship, mastering the neighborhood, and exploring interests. It alters the normative calculus of -- including the potential costs -- of both experimentation and commitment. It increases the perceived costs of curiosity and enthusiasm. It alters the normative balance between day-to-day preoccupations and long-term goals. It gives particular meanings to academic success, standing out, and in that way shapes children's and adolescent choices' about where to focus energies. It affects the ways in which teachers and other authority figures view children. In other words, in a host of ways it limits opportunity to exercise growing capacities.

Context has a particularly powerful effect on children's developmental trajectories. The longitudinal picture for low-income children differs from that of their more advantaged peers, and is marked by a steeper decline in commitment to school, sense of self as a capable learner, sense of self-efficacy, expectations (as opposed to hopes) for the future, and related variables. By the time they have reached adolescence, many inner-city children's lives have been marked by the kinds of relational experiences, day-to-day pressures, unexpected life events and crises that pull children off track -- preoccupied or erratic parenting, inordinate responsibility to care for self and siblings, loss of family members through separation or death, family or community violence, pressure from gangs, contact with police, juvenile justice, child welfare authorities. Such experiences can lead to questioning of self and mistrust of others (Lee, 1994; Nightingale, 1993), and can sap the psychic and physical energy needed to address normative developmental tasks.

Supporting Children's Work on Developmental Tasks:

The Distinctive Role of After-School Programs

Work on the wide range of developmental tasks of middle childhood and adolescence both builds personal resources and, as suggested just above, requires them. Such work requires openness to new experience and willingness to take some risks, the maintenance of curiosity and motivation, the sense of an (at least partly) open future. At each age a foundation is being built for the work of later ages. Developmental work also requires social resources -- environmental opportunities and challenges, opportunity to exercise growing capacities, to observe, learn from and identify with experts in different domains; recognition from valued, authoritative others; a sense of psychological safety combined with some freedom of movement and some privacy; times and places for moratorium and renewal (time to process, integrate the variety of new demands that come with growing up); and opportunity to develop one's own point of view without fear of

belittlement or ridicule.

While it is obvious that no one institution or setting can assume sole responsibility for the tasks and attendant range of supports outlined above, it is not clear how the roles and boundaries of different institutions are best sorted out. Historically, roles and responsibilities have been repeatedly negotiated and renegotiated. What seems important is recognition that different institutions are not interchangeable, that they have different strengths and limitations, and that these lead in turn to distinct roles in addressing developmental tasks.

After-school programs are a particularly flexible institution, filling gaps, complementing the primary institutions of family and school, and, notably, providing opportunities, supports and resources that these other institutions cannot or will not provide, especially to low- and moderate-income children. For instance, because they are not a mass institution, because they have no fixed set of knowledge and skills that all children have to master, and because they have no sorting and labeling responsibilities, after-school programs cope well with individual differences. Children can bring --and be -- who they are without worry that they do not match some mysterious (to them) set of institutional expectations. Challenges created for children can be individualized. After-school programs can respond to children's interests and concerns, giving children a measure of control over what they do every day, putting children in active roles as learners, and attending to their point of view. Together these attributes may support the development of a sense of agency and self-efficacy, nurture self-expression, and convey a sense that it is all-right to be oneself, whoever that self is, to have one's own views and perspective, whatever that may be.

After-school programs have the flexibility to provide developmental experiences in a range of domains that schools lack time for, and that low- and moderate-income families may lack resources to purchase in the marketplace. These include, of course, the visual and performing arts, humanities, civics, physical activities and sports. (One might argue that, given declining attention to them in school, the natural sciences have to be added to this list.) After-school programs' flexible temporal framework affords time for children to sample different kinds of activities, and to pursue selected activities and projects in depth. In some programs, children will have opportunity to work with adults skilled in a specific craft, art form or discipline. Under such circumstances, children can be exposed to the basic concepts and techniques of a discipline, craft or art form, experience practice or rehearsal, begin to learn what it takes to get good at something, begin to think like an expert, and to play with particular identities. They can exercise creativity and learn about "creating", including the need to take risks "to gain the experience that helps you make good risky choices" (Bereiter, 2002, p. 358). In general, activities tend to be relatively more "real", less symbolic and abstract than in school, affording opportunity to create products and performances that are relatively concrete and authentic.

Activities in after-school programs tend to be something to enjoy, not something to "get done" for some adult-defined purpose. That in turn may help children learn to enjoy process as much as "product". Feedback from adults is easily integrated into the learning process, an

important lesson for children for whom it is too often summative and set apart from every-day performance. Because there is a relatively low risk of failure, and because adults are in non-judgmental roles, children usually feel safe psychologically as well as physically. Summative assessment, when it occurs, is “low-stakes”: The director of a martial arts center tells Musick (1999, p. 38) that “By the time we invite [students] to test [for the next rank] we have already determined that they will pass the test. There are two ways you can fail . . . one by giving up, and two by being disrespectful of yourself, your partner or teacher.”

After-school programs are supportive of the social and interpersonal dimensions of children’s development. Their activities are full of children sharing, collaborating, helping each other, working and playing together. They therefore provide opportunity to learn about the social dimensions of creating and producing. (As Bereiter, 2002, p. 352, puts it, “Very quickly ‘my’ idea must become ‘our’ project, or it will amount to nothing.) After-school programs allow children to learn how to do “friendship” in a protected setting, and when older to make new friendships in a psychologically safe setting. The social nature of both formal and informal activity provides opportunity for children to learn to listen, negotiate, work at understanding others’ intentions and interests, learn to influence others, be responsible for others, learn to monitor the effects of their behavior on others. Differences among children may contribute to growth in a sense of empathy and flexibility. Being broader than family but often rooted in children’s home communities, after-school programs can also easily incorporate children’s home and community culture, and thus are good settings in which to explore links between “a society’s cultural heritage and [children’s] personal experience” (Damon, 1990, p. 48).

After-school programs may also offer distinct potential to help address not just the normative tasks of childhood and adolescence, but the range of vulnerabilities -- self doubt, mistrust, lack of basic skills and self-disqualification -- resulting from growing up under conditions of resource scarcity, social isolation and depredation. As noted, such vulnerabilities can impede both the exploration and the commitment important to work on varied developmental tasks. After-school program staff have the luxury of attending to children’s developmental struggles without labeling or defining children by those struggles. The defining features of activities and relationships in after-school programs may support a modest re-working of self. They offer the possibility of re-building capacity for trust, for dependence, and for openness to learning, for strengthening young peoples’ sense that they have something to say, are worthy of being listened to, the sense that their aspirations and struggles matter to the larger world around them.

Enriching School Domains. Even with developmental tasks closer to the heart of schooling, after-school programs have a distinct role, offering potential to foster capacities that schools too often ignore. With respect to literacy, for instance, these include helping children explore varied reasons to read and write, strengthening children’s sense of ownership of reading and writing, and their sense that they can use literacy for their own ends (Spielberger & Halpern,

2002). They can afford children the “freedom” to have their own, deeply personal reactions to texts (Wilhelm, 1997, p. 21). They provide opportunities for children to come to know the literacies of their own heritage -- the forms, the stories, the particular uses of language; and can play a bridging role between the literacies of home or community and that of school.

After-school programs are well-suited to fostering literacy through the visual and expressive arts, and to activities that work simultaneously across different symbol systems—words, pictures, music, movement. Since each art form has its own vocabulary and grammar, children also can be challenged to make connections between creative expression and language, learn correspondences between movement and sentences, or jazz notation and writing, and better understand narrative structure. Not least, because after-school programs are peer-oriented as much as adult-oriented settings, they provide opportunity to make reading and writing social -- for instance, reading together with others, jointly writing poems or stories, writing for a broader audience than is usually possible in school, linking children in cross-age pairs for reading aloud, and so forth

Other Factors Important to Research Design: Children’s Ages, Individual Responses, Diversity of Programs

The enumeration of children’s developmental tasks, and identification of those tasks that after-school programs are well-suited to help address, point in the general direction in which evaluators (as well as sponsors and funders) might look to define appropriate expectations, conceptualize and measure program quality and consider possible program effects. Three additional variables nonetheless complicate the task of “narrowing down”, and require conceptual attention in the design of specific studies. One is the fact that after-school programs serve children in a very broad age range, with the implication that different clusters of children in after-school programs will be working on different developmental tasks (or on different dimensions of childhood-long tasks), will have different kinds and qualities of experience (including different patterns of participation), and will have differing abilities to “demonstrate” (i.e. make visible) new skills or capacities. As they grow older, children use after-school programs differently, perhaps choosing which days to attend, using programs as just another piece of an after-school puzzle. In other words, the meaning of participating in a particular after-school program is sometimes different for younger and older children.

The second complicating variable is the fact that, even within a narrow age range, children bring different histories, predispositions and capacities to their after-school experiences, and therefore may be likely to want, need and get different things out of those experiences. Children bring varied capacities for openness, risk-taking, engagement, and for learning itself, to developmental experiences. They bring different interests, as well as different worries and preoccupations, different levels of energy, different levels of physical well-being. Some children

drop school from their minds when the school day ends, others bring the experiences of the day with them to the after-school setting, to be sorted out and processed. All these predispositions shape what and whom children are drawn to in a particular program.

A third variable, complicating design in studies of initiatives or multiple programs, is the variability among after-school programs, even those in a common network or initiative. Programs vary in mission, activity emphasis and structure, staffing, climate, and in many other ways that affect children's experience. Some programs are eclectic, some focused, some more socially oriented, some more task oriented, some highly organized, some low key and less structured. Auspices, physical setting, community context contribute to program diversity as well.

The larger point of enumerating these additional factors is that it does not make sense that a small handful of outcome measures, even carefully considered, would do justice to what are likely to be age-specific, program-specific, activity-specific and individually-distinct effects of after-school programs. It also does not make sense to employ designs and data analytic strategies that are better at capturing average effects. If an evaluator has a large number of programs and children to account for, he or she has to gain a sense of the kinds of variability that are significant for both, and then develop conceptually meaningful categories. It may then be possible to tentatively sample programs and children within each category for the inductive work so crucial to figuring out what to measure.

The Limits of Deduction. Deducing appropriate expectations, places to look for program effects, and likely sources of variability, can take the evaluator a long way. But it has to be complemented by an inside-out process that begins with simply by spending adequate time coming to know the programs one will be evaluating, on both a daily basis and over time -- for example, what kinds of experiences particular clusters of participants have every day, over the course of a week or a few months, what developmental domains are engaged and nudged, where one might look periodically for changes in skill, behavior, performance and/or products, and, more subtly in qualities of self. The evaluator may have to spend regular time over a period of months, in a number of different programs, observing, talking to staff and children, and then organizing and categorizing, before recommending sets of domains in which to concentrate measurement for different groups of participating children (and may perhaps have to hand-craft measures as well).¹⁰

In other words, evaluators have to refrain from committing themselves to particular measures of program effects for as long as possible. If they are going to use a design that involves repeated measures, they will have to wait to select or build a comparison group until they have been through the necessary deductive and inductive processes described above. If, as the approach I have outlined suggests, they are to try to stay as close as possible to participating

¹⁰ This is not the same as elaborating a logic model or theory of change, which tends to be far too general and abstract, and rarely reflects intimate knowledge of every-day work with children.

children's daily experience (and the artifacts produced in or by that experience) in looking for likely program effects, they will face the challenge of finding comparable experiential domains and measures for comparison children.

For program groups, technologies exist for enlisting program staff and children themselves in some of the work of producing evidence of growth and development. Portfolios can be a useful tool, when used systematically, and not just as "dumps" for anything a child does. Project plans provide important sources of information. For some types of projects, there will be key vocabulary, concepts, procedures, and techniques to master at different levels. Although it is important to observe children's performance and behavior, it is also important to draw on the knowledge of those who see children every day in the after-school environment; that is, the frontline staff. If they understand adequately that their observations of children's growth is not meant as a reflection of their own performance, they can be critical sources of insight about the children they work with. Discussing children's growth with frontline staff is also a tool for reflection, an activity that receives too little support in too many programs.

Conclusion

Although the after-school field is old, the research tradition in the field is young. If we are thoughtful and patient, we can build a richly creative tradition that respects after-school programs and their work, provides information helpful to programs, and clarifies appropriate expectations of them. For that tradition to take shape will require financial support and, perhaps more important, a disentangling of after-school programs from other agendas, particularly school agendas. There is a natural tendency to want to seek continuity -- what some call alignment -- between developmental settings. Nonetheless, a measure of discontinuity (or "misalignment") between schools and after-school programs, far from being a problem, is in fact critical, both for the continued development of the after-school field itself and for the millions of children who do not seem to get the nurturance, validation, identity- and capacity-building they need in the school context.

One critical task in keeping the separate identity of the two child development institutions clear is to de-enlist after-school programs from the task of closing the academic achievement gap between different groups of children. The most concrete way of de-enlisting after-school programs from that large social task is to stop using academic achievement tests as an outcome measure. It is important to note once more that, even within the field of education, the role of achievement tests remains sorely problematic. As Bereiter (2002, p. 440) recently noted, "To draw politicians and business people away from their fixation on achievement test gains one must offer them the vision of a superior kind of outcome. The failure to do that is, I believe, the most profound failure of [educators] in our epoch." The after-school field's need to move beyond this outcome measure is even more urgent, if that is possible.

What, then, should proponents and researchers tell those politicians and funders, who

seem to want a simple story-line for the benefits of after-school programs? Would these stakeholders accept the argument that good developmental experiences, in a range of domains and focused on a range of dimensions of self, may or may not spill over to school and contribute to boosting participating children's academic achievement, but that, regardless, after-school programs have to be supported to be themselves? No one knows. If we think that the scientific or developmental or moral arguments are too modest, subtle or ambiguous, then we are free to tell the politicians and funders whatever we want. There is not much difference between a small lie and a big one.

Endnotes