Evaluating Complementary Learning Projects: An Overview of Existing Practices and Challenges

Milbrey McLaughlin
Stanford University

Introduction

Young people who grow up in poverty confront the same developmental tasks as do more advantaged American youth. They must acquire the social skills, personal attitudes and intellectual competencies that will carry them to successful adulthood. But too many poor and low income youth must accomplish these goals in the context of inadequate medical care, poor nutrition, family dysfunction, unsafe neighborhoods and few opportunities for learning and development in their out-of-school time.¹ Complementary learning efforts seek to provide the range of opportunities and resources youth need to succeed in school and move toward a productive adulthood. ² These initiatives recognize that the academic success of disadvantaged youngsters requires good schools and more. If poor youth are to compete equitably with their more advantaged peers, all elements of their lives must be touched. Thus complementary learning requires increased cooperation, coordination and communication among and across public and private agencies and sectors. In these terms, complementary learning and its goals of educational equity implicate an array of youth-serving institutions and asks them to create new relationships, service arrangements, and conceptions of institutional effectiveness.

This paper considers the “state of the art” of evaluation or accountability tools and models associated with complementary learning efforts. What’s available for policymakers and practitioners to use to assess the outcomes and progress of complementary learning efforts? What challenges must be addressed if a system-wide evaluation strategy is to be put in place? Provide valid, reliable and useful information?

However, before considering the state of the evaluation art related to complementary learning, a number of issues must be considered. First, what do existing

¹ Eccles & Gootman, 2002
complementary learning systems look like? As the Harvard Family Research Project notes, “complementary learning systems take many forms, from simple to complex.” What kinds of institutional arrangements will a measurement or indicator system need to consider? Second, what are elements of a “good” measure for a complex, cross-institutional initiative like complementary learning? And finally, what functions must good measures of complementary learning play for policy makers and other decision makers at state and local levels? I take up these questions before considering “what’s out there” by way of measures, indicators, evaluation tools and the challenges associated with complementary learning initiatives.

**Complementary learning efforts**

An array of neighborhood, community, state and national efforts aim to provide more comprehensive, integrated resources to youth, most especially to those young people growing up in high-poverty settings. Most of these activities operate at the local level. Efforts such as Beacons, which originated in New York City and now operates in urban communities across the country, and community schools such as Portland OR’s SUN schools and those sponsored by the Children’s Aid Society bring health and other developmental supports together on a single site.

The Harlem Children’s Zone designates a neighborhood as location for full-service, comprehensive support to children and their families, from schools and academic coaching to parent training classes to health clinics to chess clubs. Providence, RI’s AfterZones builds on networks of public and private community partners, city departments, schools and afterschool providers to offer city-wide afterschool opportunities for the community’s middle school youth. New York City’s The After School Corporation (TASC) and LA’s BEST (Better Educated Students Tomorrow) stand out as two of America’s most successful comprehensive, city-wide after school initiatives, providing thousands of children with academic enrichment as well as quality recreational opportunities and links to community services.

---

3 Eval info  
4 Tough 2008.  
5 [http://www.wallacefoundation.org/NewsRoom/NewsRoom/PressRelease/ProvidenceAfterZones.htm](http://www.wallacefoundation.org/NewsRoom/NewsRoom/PressRelease/ProvidenceAfterZones.htm)  
6 Evaluation ref  
7 UCLA evaluation refs
Some states have begun to meliorate the fragmentation of youth services using a variety of models and incentives. Iowa’s Collaboration for Youth Development has launched an effort to implement cross-cutting youth development services and supports at local and state levels. Arizona’s “Five Keys for Youth Success” addresses the vision that “all youth in Arizona are prepared to work, contribute and succeed in the 21st century.”8 At the national level, the 21st Century Comprehensive Learning Communities is the most prominent example of federal efforts to support complementary learning opportunities9

These efforts to provide youth [and sometimes their families] with complementary learning resources and opportunities differ along multiple dimensions--location, focus, ‘curriculum’, partnerships, participants, sources of support, for instance. But they share challenges of institutional complexity, often-incompatible multi-sector regulatory or reporting frameworks, and they regularly compete for resources at the system level. Though these and other existing complementary learning efforts suggest the promise of such an approach, most especially for low-income and high-poverty youth, and offer “proof of concept,” to my knowledge no state-wide system currently exists of complementary learning of the scope imagined by a comprehensive equity initiative; thus there is little precedent to inform the challenges of delivering and evaluating a cross-system, cross-agency, cross-constituency complementary learning system for the state’s children. But, what, within the context of a system-wide, multi-actor, multi-agency undertaking, comprises a “good” measure? In addition to usual technical concerns of validity and reliability, a comprehensive learning system presents special challenges for measurement.

**What is a ‘good’ measure?**

A “good” measure with which to assess the performance and consequence of a cross-sector, cross-agency, and multiple stakeholder policy initiative derives its worth from

---

8 Governor’s Office for Children, Youth and Families (2007).
9 P/PV on Extended school services. The Supplementary Education Provision of NCLB, in principle compatible with complementary learning goals, has not been widely used because high stakes accountability standards leave few local resources to support supplementary educational services. Additionally, long-standing federal-level fragmentation of programs serving youth, and the marginalization of out-of-school initiatives has created barriers to state or local strategic use of federal funds to support complementary learning initiatives. See Dunkle 1997, for example.

McLaughlin DRAFT 3
multiple and often competing parameters.\textsuperscript{10} It must be meaningful and face valid to multiple stakeholders—legislators, parents, mayors, youth professionals, youth themselves. It must be useful to both policymakers and practitioners and share a common interpretation across users. It must provide answers to questions stakeholders consider legitimate.

A good measure must be comparable over time and contexts—defined and understood similarly in different settings. It also must be reliable—accurately represent what is being measured and not be manipulable. For example, some schools are accused of arranging field trips for their low performing students on testing day as a way to boost their achievement indicators.

A good measure also must meet criterion of practicality if it is to function as a performance indicator for a system initiative [v. a special project or one-time investment]. It must be easy, feasible and cost-efficient to collect. It must be timely and targeted to policy and practice variables and resources.

A good indicator system takes account of factors outside the control of policymakers or operating staff, or contextual considerations. For example, are declining student outcomes a reflection of changed demographics? How are social services provided youth affected by state or regional economic downturns? A good indicator system monitors context shifts critical to interpreting measures.

A complementary learning system makes additional conceptual and practical demands on measures. In the instance of a complementary learning initiative, a good measurement system looks across outcome domains—education, health, social development. It measures positive as well as negative outcomes or problem behaviors. What have participating youth achieved and accomplished developmentally? A complementary learning system also requires indicators that are accessible and relevant across agencies and levels—not just within a single sector, such as education or health.

These considerations suggest some of the technical and practical qualities of a good measure, and they also point to the importance of the process by which measures are selected and implemented. Indicators included in a public accountability system or an evaluation of system initiative such as complementary learning require the buy-in of the

\textsuperscript{10} This section benefits from the materials and thinking of Margaret O’Brien-Strain, Senior Research Associate at the SPHERE Institute. Burlingame, CA. I also draw on conceptualizations of a good social indicator provided by Corbett 2006; Moore 1997 and Moore & Brown 2006.
multiple stakeholders involved in the effort—educators, agency officials, policy makers, community members. Good measures result from a process of creating stakeholder consensus about what is important, how to measure it and what actions might result from it. Corbett (2006:18) puts this requirement bluntly: “If evaluators and policymakers cannot agree on measures of success, assessing effects is impossible.”

**Uses of Measures**

The “goodness” of a measure also depends on its purpose. Indicators or measures can and have been misused and misinterpreted.\(^\text{11}\) The purpose to which a measure is put gives it significance as indicator and warrant for action. Corbett (2006:14) draws on Brown and Corbett (2003) to define five purposes for social indicators of child well-being: these distinctions also work well in the instance of complementary learning.

1. *Description.* Descriptive information on key indicators provides important background and status information. For example: What are patterns of school failure around the community? State? What is the availability of health care providers? How many children and youth have access to an out-of-school program? What are participation patterns? Descriptive indicators let policymakers know about needs over all, variation among subgroups or across geographic areas.

2. *Monitoring.* Monitoring amps indicators up from simple description to provide information on trends over time. Monitoring lets policy makers know whether things getting better or worse over time and where new or different resources might be needed. The Annie E. Casey Foundation’s work on KIDS COUNT and the related local KIDS COUNT projects provide sophisticated examples of indicator projects assuming a monitoring function.

3. *Goal setting.* These indicators track a specific goal or expectation. Has the bottom 1/3 of underachieving students progressed? Has obesity decreased by 20% among teens in the community?

4. *Accountability.* Accountability indicators focus on outcomes. Is an initiative achieving expected consequences? Were participation targets met? Was a parent

\(^{11}\) See Moore & Brown 2006, for example.
outreach program put in place? Accountability indicators further up the stakes, adding consequences—incentives or sanctions—to performance.

5. **Evaluation.** Evaluations address the question of whether programs or polices are effective [or damaging] and to contribute insight to why an initiative succeeded or fell short. Evaluations consider “so what” questions.

Each indicator serves a different purpose and carries different implications for stakeholders. Brett Brown likens this typology to Matryoshkas—Russian nested dolls. Indicators move from the outside in, with each layer serving a different purpose. A complementary learning system will require indicators or measures that address all five functions if it is to inform policy and practice, and provide an accounting of its operation and consequences—a challenge of parsimony among other things.

---

**Evaluation Practices in Place: The state of the art**

**Theory of Change: Complementary Learning System**

<table>
<thead>
<tr>
<th>System-level factors, e.g.</th>
<th>Setting-level factors, e.g.</th>
<th>Individual outcomes</th>
<th>Individual-level Impact Domains</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Resources: fiscal; motivational; regulatory</td>
<td>• Partnerships</td>
<td>• Knowledge</td>
<td>• Cognitive</td>
</tr>
<tr>
<td>• Cross-agency collaboration; infrastructure</td>
<td>• Stakeholder involvement</td>
<td>• Attitudes</td>
<td>• Physical</td>
</tr>
<tr>
<td>• Political support</td>
<td>• Staff quality</td>
<td>• Beliefs</td>
<td>• Social/Emotional</td>
</tr>
<tr>
<td>• Technical assistance</td>
<td>• Outreach</td>
<td>• Behaviors</td>
<td></td>
</tr>
<tr>
<td>• Capacity</td>
<td>• Safety</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Services</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Trust</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The theory of change underlying a complementary learning initiative assumes that change in system-level factors will stimulate and support change in settings, which in turn will lead to positive change in youth outcomes. Change in these outcomes will impact youth in multiple developmental domains. The “arrows” in a theory of change focus on transactions and implementation. For example: If individual level outcomes are not apparent, is it a question of setting resources and environment? Did changes in
knowledge and other assets occur? Was change in knowledge sufficient to impact youth developmental assets? Or is action and quality at the setting level constrained by shortfalls at the system level—resources, technical support and so on.

An evaluation or accountability design moves from conditions to consequences, though a series of “if, then” assumptions, all of which require indicators and assessment if the resulting information is to provide useful guidance to policymakers and practitioners. To my knowledge, no evaluation model currently exists that looks across these three levels of indicators or attempts to tie them together to construct a narrative about how a policy investment made at one level of the system was carried out, and to what effect. Few evaluations of complementary learning efforts have moved from an explicit theory of change to identify key variables and constructs. But there is useful experience from which to learn. I turn now to those indicators in use that are relevant to the theory of change underlying complementary learning. Not surprisingly, the best developed are indicators of individual status and development. Less widely developed or used are indicators of setting and system features and change.

**Individual-level indicators.**

Most evaluations of complementary learning efforts feature individual-level outcomes for students and adopt an “impact” model of cause and effect relationships between program “inputs” and participant “outputs”. Academic performance—students’ successful school experience—motivates a complementary learning effort and is used in all evaluations, variously defined. But complementary learning sees academic outcomes as conditioned by and dependent on youth’s status on a broader array of indicators; here various conceptualizations and measures are being used to look across developmental domains.

A number of conceptualizations of a broader youth development perspective exist. For instance, Richard Lerner (2005) refers to the “5 Cs” of positive youth development: Competence, Confidence, Character, Connection and Caring and adds a sixth C to this well-known list, youth Contribution. A Connecticut handbook developed to support community youth development statewide grouped individual outcomes in terms of youth personal adjustment, social competencies, relationships with families, youth/adult relationships, youth/school connection, youth/school outcomes and
definitions, youth/peer connections and definitions, and youth/community connections.\textsuperscript{12}

Several frameworks collapse these outcomes into a smaller number of categories. The Forum for Youth Investment identified outcomes in developmental areas of Learning (basic and applied academics), Thriving (physical health), Connecting (social and emotional well-being), Working (vocational career experience) and Leading (civic and community engagement.)\textsuperscript{13} The Search Institute conceptualization of internal (or individual) developmental assets lists 20 assets grouped under four “asset types”: commitment to learning, positive values, social competencies and positive identity.\textsuperscript{14} The Edna McConnell Clark Foundation’s \textit{Youth Development Outcomes Compendium} categorizes youth outcomes in terms of Educational Achievement and Cognitive Attainment, Health and Safety, Social and Emotional Development, Self-sufficiency.\textsuperscript{15} The National Research Council’s Committee on Community-Level Programs for Youth developed a similar typology\textsuperscript{16}. The NRC report collected interdisciplinary evidence on youth development to create a four-fold typology of cognitive or intellectual development, social development, emotional or psychological development and physical health and attitudes.

I blend these frameworks for purposes of parsimony and conceptual clarity to consider indicators in use across three domains: Academic & Cognitive Attainment, Physical Development & Health, Social & Emotional Development. Each outcome domain includes attitudinal, behavioral, knowledge and status indicators.\textsuperscript{17}

**Academic Attainment & Attitudes.** Academic indicators routinely collected by public school systems provide status indicators of progress; they typically include data on grades, reading, math and writing skills, standardized test scores, and English proficiency. These indicators are those most commonly used to assess “impact” or success of a learning opportunity. Many initiatives use indicators of academic attainment

\textsuperscript{12} LaMotte, Stewart, Anderson, Sabatelli & Wynne (2005)\hfill 13 Ferm, Pittman with Marshall (2002)\hfill 14 Search Institute’s website elaborates the developmental assets: \url{http://www.search-institute.org/assets/}. For a community-level application of the Developmental Assets framework see Chemung County Schools 2008.\hfill 15 Hair, Moore, Hunter & Kaye 2002.\hfill 16 Eccles & Gootman 2002\hfill 17 Using indicator type specified by Hair et al. op. cit. in their youth outcome grid: domain indicator x indicator type
in addition to those found in administrative data, such as student progress on higher-order thinking skills, school and work readiness skills, knowledge of career and post-secondary education options. Attitudinal and behavioral indicators commonly related to academic outcomes include absences and tardiness, homework completion, and academic motivation.\textsuperscript{18}

**Physical development and health.** Physical or health indicators routinely include measures of teen birth, risk behaviors (drugs, alcohol), delinquency and violence rates, and obesity and well as knowledge of safe sex practices.\textsuperscript{19} The Harlem Children’s Zone evaluations take an even more comprehensive measure of youth health and physical status.\textsuperscript{20} In addition to the above commonly-used indicators, HCZ assesses such factors as participants’ asthma management and symptom reduction, physical fitness, emergency room visits, parental health insurance for children’s medical care as indicators of the success of its various youth and parent education programs. Some state-wide health indicator systems also are in place. For example, the California Healthy Kids Survey, an anonymous, confidential student and school staff report of attitudes, health risk behaviors, and protective factors, routinely gathers information on health-related behaviors such as physical activity and nutritional habits; alcohol, tobacco, and other drug use.\textsuperscript{21} Used by California schools since 1997, the CHKS consists of age-appropriate survey instruments for students in grades five, seven, nine, and eleven and is intended for use in planning and evaluating student support programs, primarily alcohol, tobacco, other drug, and violence prevention programs.

**Social and Emotional Development.** Indicators of social and emotional development tap important constructs of youth development such as connectedness to

---

\textsuperscript{18} See Afterschool Alliance July 2008 summary of academic impact of afterschool programs for academic indicators and outcomes assessed across a range of afterschool programs and settings as well as relationships between academic outcomes and other youth development indicators such as delinquency and self-esteem.

\textsuperscript{19} See Durlak & Weissberg 2007; practices reviewed in MetLife Foundation/Afterschool Alliance 2008. See also evaluations summarized in Afterschool Alliance August 2008 report on afterschool impacts on behavior, safety and family life.

\textsuperscript{20} Harlem Children’s Zone 2006.

\textsuperscript{21} \url{http://www.cde.ca.gov/ls/yl/re/chks.asp}; The CHKS is a requirement of funding for districts that accept federal Title IV Safe and Drug-Free Schools and Communities (SDFSC) funds or state Tobacco-Use Prevention Education (TUPE) funds.
society and society’s institutions, interpersonal skills and civic engagement. “Connectedness” indicators variously defined describe the extent to which a young person has a healthy, protective relationship with the institutions in their environment—school, community, family—and have been shown to have strong relationship to academic outcomes and positive life choices. Other key indicators of social and emotional development measure youth’s coping skills, self-regard, optimism and sense of self-efficacy. For example, youth development constructs addressed by an evaluation of the Children’s Aid Society’s 21st Century Learning Centers After-School Programs (2004-2007) included: resisting negative peer pressure, self-esteem, community engagement, career aspirations, decision-making and school engagement. The California Healthy Kids Survey includes an optional Resilience Module that measures youth’s ‘internal assets’ in a manner similar to the Search Institute’s framework. Connecticut’s youth development “tool kit” includes scales/items that measure such social/emotional constructs as youth personal adjustment (a strong sense of mastery, efficacy, purpose, worth, potential) and youth social competencies (empathy, cultural competence, resistance skills, conflict resolution skills).

Issues. At the individual level, many indicators exist to measure the attitudes, behavior, knowledge and status of youth across youth development domains of academic, health, social and emotional development. Until recently, positive outcomes for youth have been seen primarily in the absence of negative ones. Significant strides have been made in conceptualizing, developing and implementing individual outcome measures that move beyond the “anti’s” or problem-focused indicators—drug use, arrest, unwanted pregnancy, school failure—to include competency or strengths-based indicators of youth development such as connectedness or belonging, sense of self-efficacy, hope for the future, academic motivation. Consistent with criteria of good measures, individual level indicators look across domains, and consider positive as well as negative outcomes. Academic and indicators of physical well-being are used to serve all five functions, from description to evaluation.

---

22 See Whitlock 2004, for example. Blum
23 ACTKnowledge n.d. (~2008)
24 http://www.wested.org/cs/we/view/rs/562
25 Roth & Brooks-Gunn 2003
In short, indicator *conceptualization* at the individual level is fairly robust. However, this review of existing practices raises three issues to consider regarding individual-level indicators: feasibility, reliability/validity and relationships between outcomes and impact.

**Feasibility.** Many indicators of academic attainment and behaviors and of various health status and risk-taking behaviors exist in administrative data sets and can be harvested at reasonable cost. Problem-focused indicators—the “anti’s”—have the advantage of being easy to collect and they have face validity. A youth is either in school or not. Arrested or not. Achieving at grade level or not. For instance, most all of Iowa’s youth development indicators rely on existing administrative data from school, juvenile justice and public health systems.

However as practitioners and researchers acknowledge, these indicators are narrowly focused and miss youth outcomes associated with the broad range of developmental assets a complementary learning initiative seeks to foster. The inclusion of a range of youth outcomes represents a significant shift in conceptions of what matters and so what an evaluation should assess. While there is no longer much debate on whether to consider positive indicators, feasibility issues constrain implementation of an evaluation system that takes a comprehensive view of developmental goals. In particular, indicators of social and emotional development generally are not included in administrative data sets but are collected through special surveys, focus groups or interviews. Though some of these surveys of positive youth social and emotional development have been routinized at the system level [e.g. CA], for the most part they are associated with a special project or one-time evaluation. Most comprehensive evaluation of complementary learning activities have been foundation funded evaluations. These efforts provide resources by way of evaluation tools and empirical base to inform choices about indicators and accountability design but are impractical as a model to replicate. They are too expensive, too labor intensive to replicate in a state-wide data system.

Further, the generally primitive state of technology at the local level to collect and analyze measures of complementary learning poses a significant hurdle. For instance, barriers exist in reporting frameworks. Some school districts or social services agencies

---

26 Evaluations undertaken by Public/Private Ventures and Policy Studies Associates provide particularly useful appendices of survey items and scales. These evaluations also include information about scale technical properties, information missing in many reports.
have relatively sophisticated administrative data sets but they are incompatible with each other and difficult to merge. Community-based organizations—important partners in a complementary learning effort—especially lack capacity to generate data in a form that could be used by an indicator system. In our experience, most still rely on paper records and have scant ability to do much with these data. These capacity challenges are particularly significant for a data system that seeks to give voice to community-based youth serving organizations and weight to positive youth development. Because public agencies hold the majority of administrative data on youth, they play a central role in data integration. For one, with the data system owned by public agencies, community organizations may be hesitant to share data except in their role as contractors. Further, dependence of public systems on administrative data can lead to a dominance of the deficit picture—the anti’s—rather than the portrait of positive youth skills, attitudes and strengths favored by community-based organizations.

Complicating these feasibility problems is the fact that the measures that do exist often are part of extensive youth surveys or a component of a comprehensively evaluated youth development project. Surveys consisting of 40 or more items are infeasible in the context of an on-going state-wide accountability system. To be included in a system of accountability for a multi-agency, multi-stakeholder complementary learning initiative, the number of constructs addressed, and items assessed, would need to be significantly pared down and stakeholders would need to agree on them as both valid and meaningful. This criterion of a “good measure” presents no small conceptual, political or practical challenges.

Rochester, NY focused on these issues of stakeholder buy in and feasibility in developing its Rochester Evaluation of Asset Development for Youth (READY) tool. Finding that the few instruments that measure positive youth development outcomes were mainly lengthy, detailed, community-level surveys that were unlikely to be completed or used by community-based programs, two major youth-finding organizations in Monroe County, NY convened a team of researchers, funders, and program leaders to identify indicators that an array of community-based programs could agree on as valid measures youth outcomes. The group pared down a candidate list of 54 youth development

---

27 See for example the Search Institute report prepared by the Chemung County Schools, Chemug, New York, which assesses 40 developmental assets.
28 Sabaratnam & Klein 2006.
outcomes to four indicators that program leaders, funders and researchers agreed would demonstrate effectiveness and improve quality of services: caring adult relationships, basic social skills, decision making and constructive use of leisure time. Survey items addressing these constructs were adapted from existing sources or written by the team. The tool then was field tested to establish construct reliability and validity. The result of this work, READY, has been widely used by youth-serving agencies in the community.²⁹

Reliability and internal validity. Many of the indicators in use were developed several years ago, based in a population of white, middle class youth [the population most available to professors and researchers] and often are insensitive to cultural differences. Our own work in a community of African-American and Hispanic youth provides a cautionary tale on this point. Not only did the measure used [a self-efficacy scale] fall apart with these low-income youth, but black and brown youth expressed different understandings of the meaning of the same items. Likewise they understood questions about their sense of “safety” in different ways. Cultural appropriateness presents an especially critical consideration for evaluation of a comprehensive equity initiative aimed at disadvantaged youth, many of whom are youth of color.³⁰ Much work remains to create culturally sensitive measures, especially indicators useful in a state-wide evaluation system.

Similarly few measures of individual outcomes take age differences into consideration but instead pose questions in an age-neutral frame. Yet research demonstrates that children and youth understand meaning and concepts differently at different ages and underscore the importance of thinking developmentally as programs are designed.³¹ The Forum for Youth Investment arrays outcomes across the age span from early childhood to young adult³²; Connecticut’s model for youth development includes age-graded indicators [appropriate for ages 6-10; 12-14. etc] among its suggested measures.³³ These age-specific frameworks are the exception.

Positive youth indicators also are sensitive to validity threats on developmental indicators that rely on subjective assessments, especially in the domain of

²⁹ Klein (2006) provides an example of consensus development around school readiness indicators in 17 states.
³² Ferber, Pittman with Marshall 2002
³³ LaMotte et al. 2005.
McLaughlin

social/emotional development. These indicators are sensitive to respondent knowledge, comparative experience and context. For example, our interviews with high-poverty secondary school youth responding to our survey questions about expectations for the future—high school graduation? Post-secondary education? Career?—revealed that they had little conception of what it would take to prepare for a particular career [astronaut, lawyer, architect], and many had no knowledge even of where they stood in terms of credits for high school graduation. Yet their survey-based self-assessments on these indicators of hope and future planning were high—but meaningless as valid measures of these constructs. Similarly, youth’s assessment of whether they ‘could make a difference’ in their communities varied depending on whether they had had opportunity to try. Youth just beginning the Gardner Center’s youth leadership program rated themselves highly; youth who had worked at community change for a year provided more modest reality-based ratings.

Further, many individual-level indicators are static snapshots and do not take into account variable rates of change over time. The OR Schools Uniting Neighborhood Initiative (SUN) uses one of the few evaluation models that explicitly attend to this concern, spelling out what student progress would look like at different times if the community school initiative were on track.

**Outcome and Impact**. Outcome and impact are not synonymous terms but in practice they often are conflated. The expected outcome of an initiative may be achieved, but no impact seen. The relationship between youth outcomes in domains of knowledge, attitudes, beliefs and behaviors and impact on youth developmental assets remains relatively unexamined. In many instances, change in behavior—school attendance—is assumed to be proxy for impact on academic motivation. Some evaluations look for intermediate outcomes and evidence of change in related to these dimensions.34 For example, do youth have new information about healthy life styles and does that information lead to better choices? What is the relationship between increased school attendance and school outcomes? Is involvement in out-of-school activities related to increased connectedness to school? Academic success? But many evaluations do not explore these assumed connections. The actual relationship between outcomes and impact is important to understand as evidence of treatment efficacy, as well as insight

---

34 See, for example, analyses provided in LaMotte et al.
into factors that may affect individual capacity to act on new knowledge, beliefs and so on.

At the level of individual indicators of development, then, much important ground work has been done in conceptual and technical terms. Work remains in the areas of cultural sensitivity and inter-rater subjectivity especially, but here too empirical and theoretical resources exist to undertake this effort. Feasibility concerns associated with the need to be parsimonious in an accountability system raise practical issues; they also implicate important aspects of a “good” measure—buy-in from stakeholders. And if individual indicators are to be useful to policymakers and practitioners, more attention needs to be paid to the assumed connection between individual outcomes and impacts and the initiative’s theory of change—knowing how and why in addition to what.

**Setting-level indicators.**

Setting-level indicators measure those elements of a program, a project, or a community-based organization that are thought to influence individual level outcomes. Setting level indicators in use generally focus on five aspects of setting: participation (youth and parents); youth relationships with staff/adults; professional capacity (staff credentials, recruitment, retention); youth leadership/voice; partnerships (public and private). These measures provide policy makers with guidance about the extent to which features associated with program quality are in place as well as feedback on implementation. Although evaluations focusing on setting-level measures are the exception, those that have looked at relationships between program features and participant outcomes report consistent findings; especially in the after school arena, there appears to be convergence on key quality indicators. For example, Policy Studies Associates, Inc identified five characteristics shared by high-performing projects funded by The After-School Corporation (TASC): a broad array of enrichment opportunities; opportunities for participants’ skill-building and mastery; intentional relationship-building; a strong,
experienced leader/manager supported by a trained and supervised staff; the administrative, fiscal, and professional-development support of the sponsoring organization; strong partnerships. Other evaluations of after-school programs such as the MA After-School Research Study (MARS)\textsuperscript{38}, the Local Investment Commission (LINC) of Greater Kansas City, MO’s Before and After School Program,\textsuperscript{39} P/PV & MDRC’s report on Extended-Service Schools Initiative\textsuperscript{40} emphasize similar setting-level features—a rich ‘menu’ of high-quality opportunities for participants, staff skill and professional development opportunities, positive relationships among staff and participants.

Setting-level indicators have been aggregated to the initiative level for some complementary learning programs. For example, an evaluation of New Jersey After 3, focused primarily on service delivery—increased enrollment and attendance and expanded program content.\textsuperscript{41} Data from individual programs were tallied to provide a state-wide picture of how New Jersey After 3 was operating—a monitoring function. A statewide examination of Kentucky’s Family Resource Centers looked at the relationship between site-level implementation and students’ outcome; it found that centers with stronger implementation of the program’s components had higher success rates in a number of areas, including students’ academic proficiency scores and risk of dropping out—an evaluation function.\textsuperscript{42} Whalen’s 2007 assessment of the Chicago’s Community Schools Initiative examined “indicators of progress” provided evidence on the quality of implementation after three years—the strongest and most consistently implemented features and those that were “under-realized.” This Chicago evaluation provides an example of how setting-level indicators can serve as important evidence of needed change or mid-course correction.

The evaluation model adopted by Portland’s Schools Uniting Neighborhood Initiative (SUN) drew upon its theory of change to parse setting-level measures in terms of an expected implementation time frame. It specified student, family, community, school and system (interagency) outcomes that could be expected in short, intermediate

---

\textsuperscript{38} Intercultural Center for Research in Education & National Institute on Out-of-School Time 2005
\textsuperscript{39} Bush Center in Child Development and Social Policy 2002
\textsuperscript{40} Grossman et al. 2002
\textsuperscript{41} Walking Eagle, et al. 2008
\textsuperscript{42} Kalafat, Illback, & Sanders, 2007
and long terms. What is reasonable to expect at what point in program implementation? This approach recognizes that some outcomes may take 3-5 years to appear—assuming effective implementation at setting and system levels.

**Issues.** These examples of setting-level indicators are promising ones, and provide a good starting place for system to provide information about the progress and consequences of a complementary learning initiative. However, compared to individual-level indicators, setting-level indicators are in relative infancy in terms of technical and conceptual development.

**Qualitative considerations.** In the instance of complementary learning system, for example, a number of important considerations require measurement but interpretation requires more than enumeration. For example, the strength, stability and value of partnerships undertaken as part of the initiative comprise a critical element of the initiative. New institutional arrangements, integrated services and the quality of collaboration are seldom examined in any systematic way at the setting level. Evaluation of partnership effectiveness will require more than a simple count.

Similarly, though data about staff to youth ratios provide important information about a program, numbers themselves do not speak to critical issues of quality of youth: staff relationships.\(^4^3\) Across an array of evaluations assessing the outcomes of youth development programs, program atmosphere—the quality of relationships, the sense of personal and emotional safety—trumped other elements. Variation in qualitative aspects of the “same” program carried out in different relational contexts can result in different outcomes for youth [and for staff and parents]. Quantitative survey-based assessments can provide some information about relative differences among programs, especially at the extremes, but can provide little insight into how those relationships work and how to systematically influence them. They do not provide actionable knowledge or reliable measures of program “treatment.”

Issues of interpretation also arise with survey-based indicators. In particular, setting-level indicators pose problems of subjective assessment, and comparability over settings. What is judged “a safe place” in one setting may not be experienced as such in another. A commonly used indicator of program quality measures the rigor or challenge

---

\(^4^3\) See, for example, Roth & Brooks-Gunn 2003 review.
presented by a program. But what is a “challenging” educational program? For whom and in what educational context? In complementary learning projects, setting-level measures seldom get beyond relatively crude, easily collected “input” measures such as resource allocation or target efficiency. A complementary learning initiative is more than the sum of its parts; at the setting level it represents more than any one agency could accomplish on its own. How to capture this synchronicity at the setting level? Existing indicators and measurement tools seldom address this question.

Promising responses to this issue involve focused case studies designed to provide qualitative information about implementation processes, the significance of particular contextual features, or the consequences for particular subgroups. For example, the Ohio Urban School Initiative School Age Child Care project used a non-experimental evaluation design to collect data in sites selected to represent diversity in contexts, partnerships, program design and children served. The SUN evaluation allocated evaluation resources on a small studies that explored differential participation and outcomes associated with different student groups.

**Unexamined context considerations.** Setting-level indicators that describe program quality also are sensitive to conditions in the larger environment and so risk mis-specifying causal attribution or assumed associations. For example, does the relative lack of well-qualified staff indicate a program implementation shortfall, or constraints imposed by budget or available staff pool? Or, are disappointing youth attendance figures a reflection of youth’s assessment of the program, or the realities of gang boundaries? Transportation? Need for youth to have a job? Effective policy and program responses to setting-level indicators require not only a comprehensive picture of implementation but also information about a site’s broader social, economic political context. Unexamined context factors likewise raise problems of attribution and causality.

**Feasibility.** Issues such as these present difficult feasibility challenges. Numbers are easier to collect than information about the qualities associated with them. This information requires such additional [and expensive] methods such as surveys, interviews, focus groups or extensive observation. For the most part, these expensive, extensive data collection strategies have been used in one-time evaluations, or costly,
foundation-funded initiative evaluations [e.g. Beacons, TASC, LABest] and some good tools exist. However, shouldering these costs on a recurring, system-wide basis presents difficult political and financial challenges, but without such methods, the capacity of an evaluation or accountability system to provide useful guidance to practitioners or policymakers is limited.

**Generalizability.** Complementary learning efforts seldom represent standardized models to be installed in diverse settings. Of necessity, complementary learning initiatives are, in the final analysis, local creations. They reflect the needs of the community’s youth, the opportunities and resources available to them, the capacity and goals of partners. And partners change even within a single site. For example, during the four years we have been looking at the operation and outcomes of a local afterschool program, lead partners have changed from the Park and Recreation Department to the school district to the Boys and Girls Club, each partner bringing a different staff, approach and curriculum. When these local efforts are based in commitment to the target population, clarity of purpose, expertise, and meaningful collaborations, local specification is a good thing. *Whatever it takes*, writes Paul Tough (2008) of Geoffrey Canada’s mission to change opportunities and pathways for Harlem’s youth. But defining the ‘intervention’ or ‘treatment’ in these local terms poses a difficult challenge for evaluation; idiosyncratic, situated services will not be comparable across settings.

**System-level indicators.**

System-level indicators measure those elements of the relevant policy system—state, regional or local—that affect the setting-level factors associated with individual outcomes. Change at the system level assumes stability in the range of factors that impact program or project implementation, such as financial supports and the shifts in financing that may be necessary at state, regional or local levels.\(^{45}\) System-level change implicated in complementary learning initiatives also assumes stability in the partnerships or cross-agency collaborations that affect services provided youth; these new institutional arrangements may require policymakers to provide waivers to existing categorical constraints or incentives for agency officials or well-entrenched bureaucracies to enter...
into new ways to doing business. Continued political support for and stakeholder buy in to a system-wide complementary learning initiative is the *sine qua non* of stable, adequately-resourced system change.47

Yet, even those initiatives that assume system-level change, such as the youth development efforts underway in Iowa and Connecticut, rarely include system level indicators. The few evaluation models that consider how aspects of the implementing system affect operations at the setting level use indicators that tap elements of the specific initiative under study. For example, indicators selected to assess system support for New York City’s Department of Youth and Community Development out-of-school initiative targeted elements central to that undertaking—cross-agency partnerships, parent outreach and involvement. Wraparound Oregon, a statewide initiative that aims to “transform how government agencies and nonprofits respond to children with significant substance abuse related needs” by developing a coordinated, comprehensive network of community-based services specific to its health-focused mission. A web-based management information system monitors that service integration at the system level.48

Examples such as these underscore the importance of system-level indicators to assess progress and sustainability; they also suggest that an accountability system to monitor and support a comprehensive equity initiative will need to develop many of its own indicators based on stakeholders’ assessment of critical system supports. I found no examples of system-level indicators focusing on such issues as technical supports, policy integration, agency alignment or budgetary provisions. A complementary learning initiative requires information about system-level implementation issues such as these.

**Issues.** Evaluators have paid little attention to system-level indicators. And even more important, though integrated services lie at the heart of the theory of change advanced by a complementary learning initiative, few models exist that look across domains to consider questions of system integration [or integration at setting or individual levels].

---

46 For example, Russell et al. 2008 measured the extent and value of partnerships across public and private agencies as part of the New York City Department of Youth and Community Development Out-of-school time initiative.

47 For example Jones, Byer & Zeldin (2008) argue that stakeholders must share a similar “why” for their involvement and support. See also LaMotte et al. 2005 for discussion of statewide activities to build stakeholder buy in.

48 [www.wraparoundoregon.org](http://www.wraparoundoregon.org)
Yet, the newness of system initiatives such as complementary learning make this information crucial to policymakers’ ability to provide oversight and management support. Few integrated data systems exist to track the extent to which youth are receiving the array of services intended by a complementary learning initiative. No system exists to track outcomes and progress across systems, settings and individuals on a statewide basis. Currently the strongest models of social services evaluation are organized around discrete services and outcomes and tied to specific populations.49

The challenges to developing an integrated data system to support complementary learning are significant on multiple fronts. As Corbett (2006: 15) cautions: “The more ambitious the (integrated services) model, the more unlikely it is that good, unambiguous data will be available for the entire integrated system.” Progress in developing integrated data systems to inform policy makers’ actions and decisions has at best been modest. Yet without cross-agency, cross-level, cross-sector information, it is not possible to assess the complementary learning system’s theory of change in its entirety or monitor progress toward educational equity. The development and support of an integrated data system presents the greatest challenge of all to evaluation and monitoring of a complementary learning system.

**Needed: An Integrated Data System**

Though no integrated model exists to inform a data system for the proposed complementary learning initiative, experience with the general issue of data integration does exist to inform the process. For example, a few jurisdictions have successfully integrated data across public agencies at the state level. South Carolina, for example, has made substantial investments in integrated data systems, led by health specialists, but crossing a wide range of state departments. For example, the state linked data from health, mental health, disabilities, education, social services, and vocational rehabilitation departments to develop an unduplicated count of children with special health care needs, along with their service utilization, county of residence, educational performance and household structure. This information was used to target outreach, prevention, health and educational programs for these children.50

---

49 Corbett 2006.
50 Bailey 2003
Chapin Hall, at the University of Chicago, has pioneered strategies for using administrative data as a resource for policy, management, and evaluation. Serving as a contractor to the State of Illinois, Chapin Hall developed an Integrated Database on Child and Family Programs in Illinois. These integrated data provide a multilayered picture of the entire network of relationships linking children and public services, consolidating data from more than 30 statewide and city programs, serving more than 4 million children.  

A number of communities have adopted “report cards” to give an account of a community’s or a state’s youth indicators—high school dropouts, asthma rates, teen births. Philadelphia’s efforts supported by funding from the Urban Health Initiative resulted in one of the nation’s most comprehensive “report cards;” the city was successful in engaging both public and nonprofit agencies as well as community-based organizations in a system to link client-based data. Baltimore has implemented an ambitious effort to collect and analyze data from a variety of youth-serving agencies. These local efforts are at the forefront of efforts to consider how resources and opportunities add up for youth.

The Youth Data Archive, a project of the John W. Gardner Center for Youth and Their Communities at Stanford University, extends these report card efforts and the data integration projects at Chapin Hall in at least three key ways. For one, as a partnership with participating communities, the YDA attends specifically to important issues of stakeholder buy in on indicators and questions, strategies to incorporate contextual considerations into data interpretation and use. The YDA is not a “data warehouse” but rather an active partnership between the Gardner Center and participating communities to generate questions, identify indicators and interpret data. Second, the YDA includes data from non-profit and community-based organizations such as the Boys and Girls Clubs, in addition to data generated by public systems. And third, the Youth Data Archive matches data at the individual level. This strategy permits the YDA and partners to consider how particular configurations of resources and opportunities add up for youth from different demographic, community or neighborhood contexts. I describe the YDA strategy and operation in some detail since it offers “proof of concept” and addresses in its design many of the aspects of good measures discussed previously. The YDA also offers

http://www.about.chapinhall.org/work/work.html
lessons to inform a data system to support the proposed complementary learning initiative.

**The Youth Data Archive**

The Youth Data Archive integrates data from public and private youth-serving organizations to inform analyses at the individual, program setting and community levels (See Table 1). Currently working with communities in four Bay Area counties, the Youth Data Archive links individual-level administrative data from schools and public agencies such as child welfare, health services and probation with program data from youth development programs offered through community-based youth organizations, as well as other data sources such as survey data or qualitative studies. In doing so, the YDA aims to support policymakers, practitioners, community members and others in making positive changes in the social settings—the programs, out-of-school activities, institutional contexts—within and through which youth move. At the same time, the YDA is also an instrument to measure the settings themselves.

<table>
<thead>
<tr>
<th>Y</th>
<th>School attendance, grades and test score</th>
<th>Y</th>
<th>After-school program participation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>Child welfare caseload and placement</td>
<td>*</td>
<td>Youth employment service receipt</td>
</tr>
<tr>
<td>*</td>
<td>Juvenile probation placement and services</td>
<td>*</td>
<td>Job training participation</td>
</tr>
<tr>
<td>Y</td>
<td>TANF participation and services</td>
<td>Y</td>
<td>Sports, arts and enrichment program participation</td>
</tr>
<tr>
<td>Y</td>
<td>Housing assistance</td>
<td></td>
<td>Alcohol and other drug substance abuse services</td>
</tr>
<tr>
<td>Y</td>
<td>Public health insurance participation</td>
<td>Y</td>
<td>Teen parenting and pregnancy prevention services</td>
</tr>
<tr>
<td>*</td>
<td>Public health insurance claims</td>
<td></td>
<td>Independent living assistance</td>
</tr>
<tr>
<td></td>
<td>Arrest records</td>
<td></td>
<td>Other community based service participation</td>
</tr>
<tr>
<td>*</td>
<td>Mental/behavioral health caseload</td>
<td></td>
<td>Early childhood education</td>
</tr>
</tbody>
</table>

Y - Currently in the YDA for one or more community

* - Data use currently being negotiated and/or data sources available through other research

The YDA uses integrated data to encourage system-level responsibility and accountability for resources, opportunities and youth outcomes through coherent, mutually reinforcing policies and programs.

McLaughlin

DRAFT 23
**Event histories.** By matching individual-level data across programs and agencies and over time, the YDA creates “event histories” for children and youth that offer a much richer context for understanding participation in any given program. This analytical strategy moves away from a “main effects” or “impact” approach to program evaluation to consider the value added of an organization or activity within the context of other opportunities and resources for youth. For example, what is the value added as measured by youth development outcomes of differently implemented after school programs? Do some program designs seem particularly beneficial for youth from different cultural, ethnic or socio-economic backgrounds? In combination with youth participation in such youth-serving institutions as health clinics? Or are they most effective when other resources such as a neighborhood policing program also are present? Locating youth-serving organizations in their broader institutional environment mitigates risks of under- or over-estimating program effects because youth outcomes can be understood in the context of broader opportunities and resources.

Such event histories offer a rich strategy to measure programs and their contexts. First, in the short run, the YDA helps community partners understand the full set of services and opportunities for their youth. In particular, it allows us to determine whether multiple agencies are serving the same youth, the paths they take to reach services, and how services can be better coordinated. In addition, it can be used to find gaps in service by types of service, youth characteristics and needs, and geography. Second, over time, the YDA will allow us to measure the cumulative effects of community youth development programs and services on youth outcomes, accounting for the complex dynamics of youth participation in public and private programs. Third, where we can document key features of programs’ structures or approaches, the integrated data can document the empirical connections between youth development outcomes and other outcomes such as academic achievement, youth employment or juvenile justice.

**Expertise located in the “neutral middle”** The Gardner Center serves as a neutral third party, holding data for the benefit of the community or county as a whole rather than giving a preference to the government agencies or to the non-profit agencies, including those seen as advocates sometimes in opposition to government. In this role,
we are responsible for all the technical elements of creating the data archive, including negotiating access to the data; identifying protocols for secure, low-burden data transfer, and data matching. We also provide analytical expertise, although we are also working to build the capacity of our community partners.

Because we have no authority over any of the entities providing data, the negotiation of data access is the greatest ongoing challenge. This middle role is only feasible because we have community partners already committed to the concept of community youth development and we approach the archive creation as a long term commitment. In fact, the Gardner Center’s long-term commitment has turned out to be one of the most important features of the project. In planning for a long-term role, we take a stepwise approach to building the archive. As soon as we have data contributed, we seek to develop analyses of interest to contributors, thereby encouraging other agencies to participate. For those who are less comfortable about providing data, we develop strategies for less sensitive initial data transfer. Together these strategies allow the project to move forward even when specific agencies are not initially committed to the project or have concerns about data confidentiality. Initial “foot draggers” do not stymie the project, but can be given time to feel comfortable with the project and the process.

**Oversight & buy in by data contributors** A key strategy to ensure that agencies are willing to join the YDA is the establishment of data use agreements in which agencies do not lose ownership or control of their data. In the communities, agencies express political worries, concerns that their data in the hands of another public or private youth-serving group could be used to embarrass them or initiate legal action. This worry is not unique to our communities. As Weitzman, et al. observed, in some communities “... there is almost a paranoia that runs through it that they are afraid if they work together that someone will see inefficiencies and take money away from them.” (p. 19).

The YDA operates from the premise that agreement among stakeholders is needed on which indicators to use, and what use will be made of them. And attention must be paid to the “comfort level” of agencies and programs providing data. An initial high-stakes accountability approach likely will make project and program staff uneasy
and less willing to cooperate with data collection efforts, whether by supplying administrative data sets or administering surveys.

To alleviate these concerns, we have established two groups in each local community to advise the YDA work—a Policy Committee and a Data Oversight Team, with representatives from every major agency contributing data. The Policy Committee generates and approves questions for analysis, including topics proposed by outside researchers. In this way, the questions posed the YDA are “legitimate,” and reflect stakeholder concerns. The Data Oversight Team serves as a more technical review panel, considering the accuracy of data interpretation. Figure 2 below illustrates the YDA’s process of question generation, data collection, analysis and reporting. Especially in the early years of the YDA, we believe this approach has been critical to building trust among the partners who provide their data to the initiative. Equally important, we have found this kind of review critical to avoiding misinterpretations that may have enormous impact on the findings. The review process provides a forum for discussing such data issues and ideally, improving data collection over time as partners begin to value analyses as useful to them rather than a bureaucratic irritation.

Figure 2: The Youth Data Archive Strategy of Stakeholder Partnership in Research & Analysis

[Diagram showing the process of question generation, data collection, analysis, and reporting]

---

52 The Kids Integrated Data System (KIDS) run by the Cartographic Modeling Lab at the University of Pennsylvania to link data from Philadelphia city government and the city school district takes a similar governance approach.
Legal issues can complicate efforts to develop an integrated data system. Different public and private youth-serving agencies operate under different regulatory guidelines; human subjects or confidentiality requirements constrain the cross-agency data sharing assumed by a complementary learning system. These confidentiality assurances make it difficult to share data across agencies at the individual level, which arguably is the level of most significance if equity consequences of a complementary learning effort are to be tracked and assessed. Since the Gardner Center is a neutral third party providing aggregated results – albeit with much greater use of subgroup analysis and other statistical techniques – we fall clearly within the bounds of data use for research purposes, which implicate a different set of confidentiality rules and requirements.

Many of the features of the relational databases that are used for policy purposes, such as standard outcome reports, can be built into internet-based collaborative environments. Once the archive is established, the ongoing costs will largely be the staff time for specific analyses, which we expect will be supportable on a low-cost fee basis. In this way the YDA addresses the feasibility constraints associated with an integrated data system for a complementary learning initiative.

**Summing up**

Most publicly funded evaluation and accountability strategies associated with complementary learning efforts rely on administrative data sets which mean a reliance on “deficit” measures at the individual level. Academic achievement and school completion as routinely reported to state and local agencies figure most prominently in assessment of youth academic outcomes. However, several state and local initiatives have incorporated youth development indicators into evaluation and accountability schemes. At the individual level, an array of indicators exists to assess youth development on positive competencies, behaviors and attitudes. At the setting level, measures of program elements exist for after-school practices (TASC, Beacons, LABest, AfterSchool Alliance), and consensus about elements of program quality is evolving. However, at both individual and setting levels, much conceptual and technical work remains to put forward a parsimonious, reliable menu of setting-level measures appropriate to a complementary learning system and agreeable to all stakeholders.
Few indicators or strategies exist to measure implementation and outcomes at the system level. Progress in the development of setting or system-level indicators has been modest at best. Thinking about how a complementary learning system would operate—specifying its theory of change and appropriate indicators to policy and practice at each level—comprises a necessary task for reformers.

The greatest challenge for reformers, however, lies in the development of and support for an integrated data system to inform policy and practices of complementary learning. Such a strategy is essential to the effective function of a complementary learning initiative that seeks to address educational equity. Arguably the most significant obstacle to accomplishing this task may be political. Technical and practical issues are considerable to be sure; but experience suggests that solid progress toward a functioning integrated system requires the mobilizing of political will and leadership in its support.
References


Harlem Children's Zone Evaluation Highlights from July 05 - June 06. (2006). Harlem Children's Zone.
Pathways to success for youth: What counts in after-school. Massachusetts After-School  
Research Study. Authors.  

Jones, K., Byer, K., & Zeldin, S.  2008.  Youth-Adult Partnerships in Community Decision Making:  
An Evaluation of Five State 4-H Youth in Governance Programs.  

Kalafat, J., Illback, R. J., & Sanders, D. (2007). The relationship between implementation fidelity and  
educational outcomes in a school-based family support program: Development of a model for  
evaluating multidimensional full-service programs. Evaluation and Program Planning, 30,  
136–148.  

In Ben-Arieh, A. & Goerge, R.M. Eds. Indicators of children’s well-being. Dordrecht, The  

LaMotte, V., Stewart, D., Anderson, S., Sabatelli, R., & Wynn, J. (2005). A Statewide Capacity-  
Building Model for Positive Youth Development: Connecticut for Community Youth  
Development.  

Positive Youth Development. Annals of the American Academy of Political and Social Science,  
591, 125-145.  

Moore, K.A. & Brown, B.  2006.  Preparing indicators for policymakers and advocates. In Ben-Arieh,  

National Collaborative on Workforce and Disability.  2007.  Evaluating System Change Initiatives for  

Providence AfterZones Program.  


Journal of Adolescent Health. 32:170-182


Additional Sources Consulted


Austin, G., & Rolfe, M. Youth Development Strategies, Concepts, and Research: California Department of Education.


Building Partnerships for Youth. (2008). Elements of Youth Development Defined. from ag.arizona.edu/fcs/bpv/


California Healthy Kids Survey Middle School Questionnaire Module A: Core. (2007). WestEd for California Department of Education.
California Healthy Kids Survey Middle School Questionnaire Module B: Supplemental Resilience and Youth Development. (2007). WestEd for California Department of Education.

California Healthy Kids Survey - The Resilience and Youth Development Module Theoretical Framework. from www.wested.org/chks/images/RYDMFrame.gif


Theokas, C., & Lerner, R. M. Observed Ecological Assets in Families, Schools, and
Neighborhoods: Conceptualization, Measurement and Relations with Positive and
Negative Developmental Outcomes: Tufts University.
The Youth Program Quality Assessment. High/Scope Educational Research Foundation.
Trammel, M. Finding Fortune in Thirteen Out-of-School Time Programs: American Youth
Policy Forum.
Using Assessment Tools to Evaluate Afterschool Programs A Look at the Youth Program
Using the Resilience and Youth Development Module of the California Healthy Kids Survey.
Francisco: Public/Private Ventures.
on Child and Family Statistics.
Watson, B. H. Ten Lessons from the Community Change for Youth Development Initiative:
Public/Private Ventures.
Whitlock, J. (2004). Places to be and places to belong: youth connectedness in school and
community: ACT for Youth Upstate Center of Excellence.
of School Time Policy Commentary: Forum for Youth Investment.
reduction: towards strengths-based positive development policy. Journal of Community
Psychology, 36(6), 761-779.
Youth Development Results, Indicators & Strategies. (2008). Iowa Collaboration for Youth
Development.
Yu, H. C. (2007). San Francisco Youth Focus Groups for the Beacons Young Adolescent