Program Implementation: What Do We Know?

Contributors:
Lillian Bowie, Sarah B. Garrett, Akemi Kinukawa, Krystal McKinney, Kristin A. Moore, Zakia Redd, Christina Theokas and Brooke Wilson

Principal Investigator: Kristin A. Moore, Ph.D.

Originally prepared for the John S. and James L. Knight Foundation, August 1, 2003.
Updated, revised, and expanded for The Atlantic Philanthropies, October, 2006.

Acknowledgments

The authors gratefully acknowledge the support of The Atlantic Philanthropies and the John S. and James L. Knight Foundation. Special thanks to Tamara Halle, Ann Segal, Elizabeth Hair, Kevin Cleveland, and Daniel Berry at Child Trends, and John Bare and Julie Kohler formerly with the Knight Foundation whose contributions of time and counsel greatly enriched the report.

Note: As new research becomes available, this report will be updated regularly. The current updated version will be posted on Child Trends' web site: www.childtrends.org.

Program Implementation: What Do We Know?

Table of Contents

Program Implementation: What Do We Know?

EXECUTIVE SUMMARY

When people imagine establishing a program for children, teens or youth, they typically think about the kinds of services or instruction to provide. Exactly how a program might best provide these services is often left to chance or allowed to vary across programs. Similarly, many program evaluations examine only whether a program "works" to achieve its desired outcomes, not what features within any given program contributed to or detracted from achieving these results. However, it is likely that the ways in which programs are implemented—choices made about staffing, conceptual approach, and various program practices—determine the quality of services and participant outcomes. Poor implementation may undermine program outcomes. In addition, programs may be wasting funds on practices, approaches, or staffing decisions that may not have the desired, consistently positive effect on outcomes. Alternatively, programs may be neglecting certain practices that could greatly benefit their particular participant group or program.

Regrettably, there is surprisingly little high-quality experimental research on whether, and how much, various program components help or hinder program success. The field of program implementation evaluation is underdeveloped, leaving researchers, funders, practitioners, families, and communities with little information about which implementation practices and processes are most successful. We felt that this dearth of information warranted assembling what is known about program implementation in one document, drawing on varied sources of information.

This report identifies and synthesizes what is known about specific program features for children and youth aged 6-17 that might be manipulated or regulated- components such as staff wages, group size, activities, and theoretical approach- that make up the "how" of program implementation. We have culled information from the limited extant collection of "gold standard" research—experimental studies that systematically manipulated program components and recorded their apparent effects on participant, provider, and program outcomes; this is presented in Chapter II. We have also synthesized findings from non-experimental sources, such as quasi-experimental studies, multivariate analyses, meta-analyses, and literature reviews (Chapter III). Finally, we have drawn on the wisdom of practitioners in the field who daily experience the benefits and detriments of various program components; this perspective is shared in Chapter IV. Our goal is to summarize and, where possible, synthesize information from these three types of information.

Throughout the document and in the selected findings, below, we identify specific program features as *positive*, *neutral*, or *negative*. These assessments reflect whether the component appears to benefit or detract from positive program, participant, or provider outcomes. In general, an assessment of neutral indicates that a particular component does not appear to "matter" relative to program outcomes. Specific assessment definitions

differ slightly according to the type of information from which they are drawn; we encourage the reader to review definitions in the methodology sections of Chapters II, III, and IV.

Altogether, we reviewed hundred of studies, meta-analyses, literature reviews, and provider writings, of which roughly 130 were selected for inclusion in this report. These sources address nearly 50 different program features. Below we present a small, and particularly rich, subset of findings from this collection. These fifteen components, discussed in more detail in our concluding chapter (V) and in Tables D1 and D2, comprise those findings informed by at least two, and sometimes all three, sources of information.

Overarching Conclusions

1) Staff training

Information from experimental research, non-experimental research, and provider wisdom sources indicates that staff training (i.e., courses of study, workshops, or supplemental classes relevant to the program's specific participant population or program design) is linked with positive program and participant outcomes.

2) Staff experience

Practitioner wisdom suggests that experience in combination with high-quality interpersonal skills and program-specific training is a positive and important program component in center-based or after-school programs for older children, teens and youth. This difference may reflect a nonlinear pattern, whereby some experience is positive, but too much experience represents outmoded strategies or provider fatigue. Also, it is possible that years in the field affects providers differently according to their institutional setting, or that working with different populations "matures" providers in different ways.

3) Participant-staff ratio

Non-experimental information suggests lower participant-staff ratios may be beneficial for children in center-based programs. However, provider wisdom suggests that for groups of older children, teens, or youth, too many staff present can be detrimental to the comfort level of the participants and the youth-centered dynamic. The correct ratio in this case may vary according to the specific setting, program type, and participant group.

4) Staff wages

Provider wisdom sources indicate that higher, more competitive, and/or more satisfying staff wages and benefits are beneficial to program outcomes. Provider wisdom sources indicate that this component is important for staff who serve participants of any age. Provider wisdom suggests that this component is important because staff who are satisfied with program compensation are more likely to stay with their programs and may also develop higher levels of engagement with the program.

5) Interactive approach vs. information-only approach

Experimental research, non-experimental research, and provider wisdom all find that interactive program approaches, especially as compared with didactic, information-only approaches, are relatively promising in producing positive outcomes in older children, teens, and youth. Programs that engage participants through activities such as role-playing and group discussions in center- or classroom-based sessions, for example, have been linked with positive behavioral outcomes (e.g., decreased substance use, deferred sexual initiation). In regard to non-interactive, information-only approaches, three sources find that they are not a promising approach for generating notable positive outcomes, yet one source finds that this approach may be promising when targeting specific genders. Furthermore, information-only approaches have occasionally been linked to changes in the level of participant knowledge about a targeted program topic, but not to changes in behaviors.

Many program approaches can be described as interactive or non-interactive. As evident in the two discussions that immediately follow, slight variations on the interactive program approach have been found to be promising as well.

6) Flexible, participant-oriented approach

Non-experimental research and provider wisdom sources indicate that dynamic, participant-oriented approaches are particularly positive. Specifically, this supports programs that change and adapt to accommodate participants' growth and development. Positive caregiver behaviors for this approach—for regular staff as well as for adult volunteers in mentoring programs—include being flexible in the interaction or teaching styles they employ with participants, and not imposing a preexisting, rigid "agenda" that they may have for participants. This approach is also closely tied to encouraging teen and youth participants to have a voice in program development and activity planning. Youth involvement is particularly important according to provider wisdom; this practice addresses youths' needs to feel useful and to find reason to engage in the program. However, it is important to note that, for younger children, this kind of flexibility has been linked in non-experimental studies with negative outcomes.

7) Multi-component and "whole person" approach

Both non-experimental research sources and provider wisdom identify multi-component and "whole person" approaches as particularly promising. These approaches are designed to address the varied needs of participants and the various contexts in which they exist (e.g., academic, family, social, residential/neighborhood-based), as well as the cross-cutting pressures endemic to playing multiple roles and having multiple needs (e.g., school-work conflict, sexual decision-making, pressure to join a local gang). The goal of the whole person approach is often the promotion of multiple positive skills in order to avoid or minimize negative behaviors, and healthy development for its own sake. The multi-component program typically effects change through addressing participant needs in various domains of his or her life, or through working with and involving people (e.g., parents, teachers) and institutions (e.g., schools, communities) from these other domains. Often, programs with these approaches teach a battery of social, problem-solving, and general life skills to enhance competency in these various domains. These approaches have been discussed mostly with regard to teens and youth, and are complementary to the

ideals set forth in youth development theory, but they are also in line with child development researchers as well (Shonkoff & Phillips, 2000).

8) Clear and consistent program goals

Both non-experimental research and provider wisdom sources provide information on the apparent effects of clear and consistent program goals. However, the information from the sources address different contexts in which to communicate these goals, and are—conceptually—slightly contradictory. Non-experimental sources suggest that in prevention programs, the practice of staff articulating to participants the program goals (e.g., to prevent drug use) or the programs' stance on certain behaviors (e.g., zero tolerance for drug use) facilitates achieving program goals for the participants (e.g., no drug use). However, provider wisdom asserts that it is occasionally detrimental to program success when behavioral goals, particularly those related to negative or stigmatized behaviors, are articulated. Participants and/or their families may be resistant to involvement in programs that address behaviors they feel do not apply to them (e.g., reproductive health programs, if parents do not believe their teen is, would be, or should be sexually active). On the other hand, however, provider wisdom suggests that it is appropriate and beneficial to articulate positive program goals (e.g., teaching problem solving, refusal skills) which often attract participation, and the achievement of which typically affects other (potentially stigmatized) behaviors.

9) Group size

Experimental research found that smaller group sizes produced more positive academic outcomes for school-age children. Non-experimental research indicates that, overall, smaller group sizes are a positive, though sometimes neutral, feature for all ages. For youth and adolescents, provider wisdom suggests, however, that the ideal group size may not necessarily be small. The ideal size would vary by type of program, participant needs, staff ability, and program resources; provider wisdom suggests that individual program providers can most likely identify for themselves what group size would be most optimal for their specific group.

10) Diversity of activities

Both non-experimental and provider wisdom suggest that it is beneficial to provide a variety of activities—activities that are interesting, engaging, and enjoyable—to help meet the needs of the various learning styles and interests of school-age, teen, and youth participants. There is non-experimental evidence, however, that for very young school-age children—first-grade boys, specifically—attending a program with "a larger number of different activities" was linked with negative outcomes. However, this finding may speak more to the fact that a greater number of different activities may not satisfy younger children's needs for significant structure in their programs than it speaks to problems with offering a variety of activities.

11) Incentives for participation

Experimental research and provider wisdom sources find that incentives, in general, are a positive program component. An experimental study found that cash incentives helped students to be more engaged in the academic programs in which they

were enrolled. Provider wisdom suggests that incentives and rewards (e.g., trips, snacks) encourage youth to participate more frequently and potentially increase enthusiasm for the program.

12) Parental involvement

Experimental research, non-experimental research, and provider wisdom find that parental involvement can be a positive component in programs; some experimental and non-experimental sources, however, also show that sometimes it appears not to matter. Overall, the effects of parental involvement appear to vary by level of participation, kind of participation and reasons for participation. Experimental studies have found that parental participation in a literacy program and two obesity-intervention programs produced significantly more positive impacts than involving the school-age children only. Regarding obesity, this may reflect the importance of involving the adults who buy and cook food. Parental involvement did not appear to matter in a social skills training program, however. Non-experimental reviews found similarly positive findings for parental interaction with school-age and teen participants' mentoring relationships and after-school programs (but little apparent effect in services for young children, perhaps because parental involvement is ubiquitous at these ages). Provider wisdom sources, however, suggest that engaging parents in programs for younger children clearly benefits participants.

13) Dosage and Duration

All three methodological approaches suggest that in certain contexts, higher intensities of involvement—that is, the amount or quantity ("dosage") of time in which a participant is involved—lead to more positive participant outcomes than lower intensities of involvement. This pattern was not significant for all ages or contexts, however. Experimental studies have found several examples of higher program dosage producing more positive and/or longer-lasting outcomes among teens and youth in mental health and family reunification programs and among younger children in education and low birth-weight intervention programs, for example, but not among certain risky behavior prevention or juvenile justice probation programs. Importantly, slight differences in frequency often did not produce significantly different outcomes. Non-experimental sources have found both higher dosage and longer duration to be positive; duration was identified as particularly important in the context of mentoring relationships. Provider wisdom identifies sufficient dosage and duration as one of the most important aspects of programming, asserting that a program's design must incorporate enough time and meetings for program providers to be able to address participant needs, and for participants to achieve program goals despite any unanticipated delays that may occur. Between experimental, non-experimental, and provider wisdom sources, all age ranges and program settings (see Table D2) are found to benefit from some combination of greater dosage and duration.

14) Overall program quality

It seems intuitive and self-evident that "overall quality"—often described in the context of a program meeting or surpassing a battery of established standards, or following various best practices—is a program characteristic that would be linked to

positive participant and program outcomes. Non-experimental research and provider wisdom sources confirm this expectation. Non-experimental sources find that quality—most often, environmental, as measured by several different established instruments—is related to positive child outcomes in the child care context even after controlling for various other potentially influential participant, center, and location-based factors. The provider wisdom discussion finds overall quality to be important, as well, describing it as the end product of different program practices and processes employed and the quality with which each of these individual components is implemented.

Conclusions that combine these 14 overarching findings with the roughly 35 other components we addressed are presented in the form of "ideal" programs. What would the perfect program look like according to all of these findings on what components appear to promote positive outcomes, to not matter one way or another, or to detract from program goals? Drawn from information across all three sources, a generic "ideal" program is presented immediately below. Descriptions of age-specific "ideal" programs follow, teasing apart differences in ideal program implementation according to the varied developmental needs of different age groups.

Generic "Ideal" Program

According to the information we have amassed from experimental, non-experimental, and provider wisdom sources, what would be an ideal program across different age groups and institutional settings? Answering this question is a bit like assembling an ancient vase from fragments found in an archeological dig. The pieces are not all there, and many of the pieces are chipped. In other words, the evidence is incomplete and variable in how definitive it is. Nevertheless, there are instances were the evidence from all three methodologies converges. Moreover, when evidence is available from multiple methodologies, it tends to point in the same direction in a number of instances.

An ideal program, according to all three methodologies, would benefit from having staff with training that is specific to the program and participant age group, and having more rather than less dosage and duration (i.e., greater quantity, intensity and length, though exactly what these might be in not yet clear). In addition, an ideal program would use an interactive approach to teach and communicate messages to children or youth. In most cases, didactic, information-only approaches would not be employed. Many publications reviewed in these three research approaches also suggest that an ideal program would incorporate parental involvement to some degree. Two studies (one experimental, one non-experimental) found that parental involvement appeared not to impact children's outcomes, but some type of parental involvement for younger children seems both very likely and essential; more research is needed. In most cases, the ideal program would enroll smaller rather than larger groups of participants, and have lower rather than higher participant-to-staff ratios.

The evidence base is even thinner for several additional program features, but these program components are still supported by two methodologies. Non-experimental

research and provider wisdom jointly support a number of program features. These include a flexible, participant-centered approach, which dictates that the program adjusts and refocuses as necessary to accommodate changing participant needs, and the multi-component and "whole person" approaches, which address the multiple domains of development and the multiple pressures that participants experience in the various contexts they inhabit (e.g., school, family, neighborhood). Communicating clear and consistent messages about positive program goals would also typically be a positive practice according to these two sources; provider wisdom suggests, however, that programs would do well to downplay addressing stigmatized behaviors, and highlight the development of positive behaviors.

Other positive practices include paying staff higher and more competitive salaries and offering diverse, engaging, and interesting program activities. Also, provider wisdom and non-experimental research suggest that the ideal program would have a high level of overall program quality—often described as the end result of a program meeting or surpassing a battery of established standards, or following various best practices. Finally, experimental research and provider wisdom support the practice of providing incentives for participants to achieve identified goals or attend the program.

Finally, a number of ideas come up in just one source, often from providers or non-experimental research. Some of the provider wisdom reflects common sense, while some of it may not be obvious or even salient for many adults. These program elements include program accessibility, developmentally appropriate behavior management and discipline, a safe and stable atmosphere, targeting the participation of hard-to-reach populations, staff retention (minimal staff turnover), program evaluation, and a positive, strength-based program philosophy. The provider perspective also notes that staff need a combination of commitment, positive personality traits, cultural competence, and experience with participant populations. These staff then need competent and consistent supervision, management, and support from directors, as well as opportunities and support for professional development. Finally, provider wisdom identifies collaboration with institutions and professionals in the community as a promising program practice. Non-experimental research identifies implementation fidelity to theory-based program design, and the screening, training, and supervising of adult program volunteers (e.g., mentors), as beneficial. The evidence with regard to the degree of flexibility or structure in programs seems to vary by age of child. Ideally, greater program flexibility would be featured in programs for older children and teens, while less flexibility and more structure would be featured in programs for younger children.

Several program elements consistently do not seem to matter. For example, in this ideal program, work with teens would not necessarily have to be delivered by adults; experimental findings show that trained peer leaders can work as well as (and sometimes better than) trained adults. Also, non-experimental research suggests that program leaders or mentors would not have to be of the same demographic characteristics as participants; it may be more important that the mentoring pair have the same interests, or that program providers are respectful of and sensitive to cultural differences.

Finally, this generic, ideal program would do well to avoid certain program practices. Non-experimental analyses have linked greater program flexibility (for younger boys only) with negative participant outcomes. Also, numerous experimental and non-experimental studies have found the use of "shock" approaches produces no or negative outcomes.

Age-Specific "Ideal" Programs

The characteristics of a given program obviously need to vary depending upon the age of the participant. The primary goal of this report was to identify program elements that work across age groups; however, most studies are conducted on narrow age groups, which makes it difficult and even somewhat risky to extrapolate the results to all ages. Accordingly, for certain components, it became important to discuss our findings by relevant age ranges: middle childhood (ages 6 to 11) and teens and youth (ages 12-13 and older). Information on differences by institutional settings is available along with these age-specific findings in Table D2 of the report.

In the interest of space and readability, sources for the following ideal program descriptions are expressed in abbreviations—Ex (experimental), N (non-experimental), and PW (provider wisdom).

Middle childhood

For participants in middle childhood—roughly ages 6 to 11 or 12—the ideal program would be staffed by caregivers who had received training specific to the program context and/or age group (Ex, N, PW), who receive regular and appropriate professional management, support, and development, and who had a combination of experience in the field and a positive outlook toward the program and participants (PW). Also, adult volunteers such as mentors would undergo comprehensive screening, significant program training, and staff supervision (N). In general, the program would have a lower, versus higher, participant-to-staff ratio (N, PW); however, provider wisdom asserts that if the goals or format of the program are considered by staff or program designers to benefit from fewer staff per children, the ratio should be determined according to these needs (PW). Staff would receive competitive salaries and benefits packages (PW), turnover would be minimal (PW), and staff would foster positive, warm relationships with the participants (N, PW). It would be important that providers be culturally competent, able to address and accommodate cultural differences (PW). It would not necessarily be important to match caregivers to participants based on race, ethnicity or gender (N).

The conceptual design of the program would be flexible and participant-centered, meaning the programs would adjust to and change with the participant's needs (N, PW). It would communicate clear and consistent messages about desired participant goals and program philosophy (N), it would be interactive in approach (rather than information-only (Ex, N, PW), and it would address the "whole person," teaching the skills needed to flourish in every arena of participants' lives (social, emotional, academic, neighborhood,

family, etc.; N, PW). The program would be implemented with fidelity to its research-based design (N).

Many different program practices would be best bets, as well. In general, the program would have a smaller, versus larger, group size (N, PW); however, if the goals or format of the program are considered by staff or program designers to benefit from a large number of participants, the group size would be determined according to these needs (PW). The ideal program, as a whole, would be of high quality (PW), and it would offer diverse activities to all but the youngest participants, who tend to need structured activities (N, PW); all activities would be designed to be interesting, engaging, and enjoyable to children in this age range (N, PW). Staff would employ developmentally appropriate behavior management and discipline (PW). The program might provide incentives for participation or for achieving certain goals (Ex). The program would be geographically accessible to its target participants, and its environment would be safe and stable (PW). If possible, the program would include parental involvement, though this may not be an essential component in every program (Ex, N); one of four experimental studies found that parental involvement appeared not to impact children's outcomes. In programs structured around caseworkers, participants and/or their families would receive intense case management or multisystemic therapy, rather than traditionally less intense services (Ex). In cases of child custody disputes, the ideal program would resolve the dispute with mediation instead of litigation (Ex). At predetermined intervals, the program would undergo assessment and evaluation, and implement changes accordingly (PW). In the community, program staff would have positive relationships with schools and other institutions in order to maximize various resources available to them (PW). Finally, the program design would involve higher, rather than lower, levels of dosage and duration (Ex, N, PW). Of course, the precise definition of these program elements, in general or for particular subgroups, cannot yet be specified.

Teens and youth

Surprisingly little experimental research informs program implementation for teens and youth. For teens and youth, the ideal program would be staffed by caregivers who have received training specific to the program context and/or participant age group (Ex, N, PW); who receive regular and appropriate professional management, support, and development; and who have a combination of experience in the field and a positive outlook toward the program and participants (PW). Also, adult volunteers such as mentors would undergo comprehensive screening, significant program training, and staff supervision (N). In most cases, the optimal participant-to-staff ratio would be determined by well-informed staff or program designers according to the particular program type and needs of the participant group (PW). Staff would receive competitive salaries and benefits packages (PW), turnover would be minimal (PW), and staff would foster positive, warm relationships with the participants (N, PW). It would be important that staff be culturally competent—able to address and accommodate cultural differences (PW). It would not necessarily be important to match caregivers to participants based on

race, ethnicity or gender (N). For prevention programs, it may be promising to involve peer leaders (Ex).

The conceptual design of the program would be flexible and participant-centered, meaning the programs would adjust to and change with the participant's needs (N, PW). It many cases, it would be a promising approach to communicate clear and consistent messages about desired participant goals and program philosophy (N, PW). However, for programs with goals addressing stigmatized behaviors (e.g., sexual or criminal activity), it may be more productive to emphasize positive secondary program goals (e.g., decision-making or job skills). This approach would help engage youth and families who otherwise may have felt alienated or offended by the focus on negative behaviors (PW). Similarly, the ideal program would embrace a "strength-based" philosophy, which emphasizes positive development and considers youth to be "at promise" rather than "at risk" (PW). The program would be interactive in approach, rather than information-only (N, PW), and it would address the "whole person," teaching the socio-emotional skills needed to flourish in every arena of participants' lives (social, academic, neighborhood, family, etc.; N, PW). The program would be implemented with fidelity to its research-based design (N). Finally, "shock" approaches—in which youth are made to visit prisons and observe the "horrors and difficulties of [incarcerated] life" there—would not be employed (N).

Many different program practices would be best bets, as well. In general, a program would have a smaller, versus larger, group size (N, PW); however, if the goals or format of the program are considered by staff or program designers to benefit from a large number of participants, the group size should be determined according to these needs (PW). The program would offer diverse activities (PW), all of them designed to be interesting, engaging, and enjoyable to participants in this age range (N, PW). Staff would employ behavior management and discipline based on best practices and research (PW). The program might provide incentives for participation or for achieving certain goals (PW), though experimental research indicates that case management or multisystemic therapy may be even more successful at promoting positive outcomes than incentives (Ex). The program would provide leadership opportunities for youth to further enhance participation and positive outcomes (PW). The program would be geographically accessible to its target participants, and its environment would be safe and stable (PW). Special efforts would be made to attract participants in this age group, as program participation must compete with myriad other potential distractions and activities that the youth can choose to do instead (PW). If possible, the program would incorporate collaboration with the community (e.g., financial supporters, informal participant recruiters; PW) and parental involvement (N). If employing videotape instruction, the ideal program would supplement it with skills training from and interaction with a program leader (Ex). At predetermined intervals, the program would undergo assessment and evaluation, and make changes accordingly (PW). The program design would involve higher, rather than lower, levels of dosage and duration (Ex, N, PW). Finally, the overall quality of the program—the end result of the various practices and processes it implements—would be of high quality (PW). Again, the details for who

much and form whom are unclear, though the general outlines of good program elements take shape across the available studies.

Conclusion

At this point in time, what do we know about program implementation? We have examined experimental studies (Chapter II), quasi-experimental and non-experimental research (Chapter III), and provider wisdom (Chapter IV) to inform this critical question. The field benefits from having some high-quality experimental studies, but they are so few in number and spread across such disparate domains that it is difficult to reach definitive conclusions based solely on the "gold standard" evaluations. Moreover, very little guidance exists on the cost-effectiveness of various implementation elements.

On the other hand, there is a great deal of non-experimental research; however, this type of information is almost always accompanied by significant caveats, most often because the study design obscures causal relationships.

Finally, wisdom from the field is easily accessible and grounded in experience, but it cannot inform the field with the certainty inspired by controlled studies, and it tends to reflect the unique values, experiences and work view of the person providing the opinion.

Taken together, however, information from these sources can be greater than their parts. By spanning a variety of sources, program types and populations, preliminary conclusions on "what works" can be made with more certainty than they could when based on a single study or type of information. In addition, many hypotheses are provided for much-needed research on the elements of program design.

This is a field in great need of more, and higher-quality, experimental studies. Many program components would benefit from rigorous research conducted on them. Studies of program component effects on historically understudied groups (e.g., Asian American, Latino American, immigrant, sexual minority populations) would also advance the field, as disparate population effects are likely. Examination of varied elements for males and females warrants study, as does examination of children from low-income versus higher-income families and children of varied ages. A better understanding of what works—and how it works for whom— should enhance the cost-effectiveness of programs and should improve program outcomes, as well.

I. INTRODUCTION TO THE REPORT

When funders and policy-makers think of establishing a program, they tend to think primarily about what services or instruction to provide, not about how best to provide them. The majority of evaluations have a similar focus, assessing the services or instruction that programs provide more than the ways in which the programs went about providing them. This practice informs the field of a very limited type of conclusion: whether, and to what degree, a program "works" to achieve its specific outcomes. Program effectiveness, however, is a sum of many parts, including how and how well programs are implemented. Unfortunately, there is surprisingly little high-quality research on whether, when, and for whom these various parts "work," themselves, especially relative to the myriad programs under way and the large sums spent on varied prevention, service, and intervention approaches. Thus, the components that account for program effectiveness largely remain a "black box."

The goal of this report is to identify and synthesize what is known about specific program components, such as staff wages, group size, and theoretical approach. We seek to provide an overview of the state of the field on individual factors that make up the "how" of program implementation. The way in which a program is implemented affects take-up, participation levels, and impacts; unfortunately, we have found that the field of implementation evaluation is very underdeveloped. Given this dearth of evidence (National Research Council & Institute of Medicine, 2000), researchers, funders, and practitioners do not enjoy much certainty about which program implementation practices are most successful.

Given this dearth of information, combined with the need to know what works when and for whom, we have employed three approaches to identify elements of implementation that enhance (or undermine) program effectiveness. First, we examined evaluations that lived up to methodological "gold standard;" specifically, we have reviewed studies that employed an experimental design to investigate how and whether outcomes varied with the manipulation of a regulatable or manipulable feature of the program (e.g., staff-participant ratio, parent involvement, frequency of contact, etc.). Unfortunately, there are few studies that manipulate such program components deliberately in a rigorous experimental study. Repeated searches have yielded only 35 random assignment studies of social intervention for school-aged children and youth. Findings based on experimental studies are summarized in Chapter II.

Therefore, we also pursue a second approach, presented in Chapter III, a review of results from non-experimental studies. Many evaluation findings that can inform this discussion come from studies where variations in manipulable features were varied within quasi-experimental conditions, or where variations occurred accidentally in otherwise experimental studies. Additionally, this section discusses findings proffered in meta-analyses and reviews of program studies, which typically combine across methodological approaches.

Third, we sought information from the perspective of those who implement programs, themselves—program providers. These perspectives, presented in Chapter IV, are meant to elucidate what practices and approaches are found to be most successful by those who have "hands on" experience with programs. What components do they believe lead to positive outcomes? Through descriptions and anecdotes, this chapter provides valuable insights and helps to extend, broaden, and provide additional meaning to the constructs discussed in the preceding research-based chapters.

Finally, in Chapter V, we attempt to synthesize findings from these three sources in order to draw several general conclusions. Conclusions for Chapters II, III, and IV are presented in their respective chapters in the form of a description of an ideal program, based on information from that particular section. Conclusions based on information from the entire paper—including those specific to programs for school-age children, and teens and youth—are also presented in Chapter V.

Throughout the document, we identify specific program features as "positive," "neutral," or "negative," in the context of producing positive participant or program outcomes. The meaning of positive and negative assessments correlates with whether the component appears to benefit or detract from positive participant outcomes, respectively. In general, an assessment of neutral indicates that a particular component does not appear to produce a measured program or participant outcome; in lay terms, it appears not to "matter." It is important to note that the definition of these terms varies slightly between chapters in order to address the dissimilar types of information specific to experimental, non-experimental, and provider wisdom sources. Please see each chapter's methodology section for its distinct definition of these terms.

It is also important to note that our definition of "program" is quite broad. It includes organizations and strategies that fit the more traditional notions of what a program is, such as after-school recreation centers and school-based substance abuse prevention workshops. It also includes child-, youth-, and family-centered services that do not function in as formal or structured contexts, such as mentoring programs and home-based parenting lessons. Our definition, however, does distinguish these programs from child-, youth, or family-related policies, laws, regulations, and funding streams, which have not been included in the discussion. Also, biomedical interventions are not included.

Information on program features that can be regulated or manipulated has implications for program design, development, implementation, and funding. For example, if several methodologically-sound studies find that a program message is accepted by third-graders at the same level whether it is delivered by an expensive project consultant or by their current teacher trained with appropriate materials, the financial implication is clear. On the other hand, schools or funders would have much more confidence in hiring a project consultant if children were found to experience a significantly higher understanding of the message when taught by a consultant than by a trained teacher.

As others have stated, there is "no straightforward answer to the question of what works best" in programs (Fashola, 1998, p. 54). Programs have different goals, designs, participant characteristics, and cultures and communities in which they are housed. Findings from the research on them are understandably dissimilar and occasionally contradictory. Nevertheless, some patterns are beginning to emerge and these patterns can help identify implementation features that enhance program effectiveness, as well as program features that do not work and a few that appear to be harmful. Brought together, these disparate findings can help inform program designers and funders of promising practices and approaches. Moreover, this review can suggest valuable directions for future research.

II. EXPERIMENTAL FINDINGS

By Akemi Kinukawa, Zakia Redd, and Brooke Wilson

A. Introduction

As noted in the introduction, the field of program design needs more experimental studies on program practices to understand definitively how different regulatable or manipulable features affect child and family outcomes (Hollister & Hill, 1995). High-quality research on specific program components is extremely useful for service providers in designing programs and making programming decisions, such as staffing and how long to run a program.

As seen in the following chapter, many studies have addressed the potential effects of particular program components. However, none of the studies discussed in Chapter III have the distinct advantage of an experimental design. Without this "gold standard" design, we are unable to know with certainty whether better outcomes are attributable to a particular regulatable or manipulable feature, or to other influential program characteristics (Campbell & Stanley, 1963). To provide hard evidence on the impacts of program components, we have dedicated this chapter to a synthesis of findings from rigorous experimental studies. In this chapter, we reviewed only studies that employed the following design: participants were randomly assigned to different groups, and between these groups, the program experience differed only in regard to variations in one (or more) manipulable feature(s). In such studies, statistically significant outcome differences between the groups can be understood as the effects of the manipulated implementation feature(s). More detailed information on the methodological approaches used in these experimental studies is included throughout the chapter.

Thorough literature searches generated a very limited number of studies that experimentally manipulated program components and examined their particular effects. Only 35 studies met our criteria for inclusion, despite a wide range of acceptable target populations, program types, and years of publication. Programs from diverse substantive areas, such as substance use prevention, maternal and child health, juvenile justice, and child welfare are represented. Due to the very limited number of studies—of which many are on different substantive areas and serve very different populations—it is difficult to draw definitive and overarching conclusions. However, information from experimental studies offers the best available testimonies for effective program features and leads us one step closer to finding the most effective program practices and features. More detailed information about the studies described in this chapter is presented in Appendix A.

This chapter is organized by areas of program implementation: staff and staffing, program practices, and dosage of program services. The section on staffing reviews eight studies related to issues such as staff-child ratio, service providers' background and

training, and type of instructors (e.g., adult vs. peers, or professionals vs. paraprofessionals). The section on program practices reviews sixteen studies on program practices employed to administer or enhance program services, such as parental involvement, provision of incentives, or video instruction. The last section is on dosage of services, and reviews twelve studies that examined whether different levels of dosage and duration affected program or participant outcomes. For clarity of discussion, some sections have been sub-divided by program type.

B. Methodological Approach

1. Literature search

We conducted an extensive literature search in EBSCO Host, a research database that provides access to four major multidisciplinary databases relevant to our work: PsycINFO (abstracts and articles), Psychological and Behavioral Sciences Collection, Sociological Collection, and Social Science Abstracts. We employed over 30 search terms from the following categories: (a) regulatable or manipulable program features such as peer-led classes, dosage, or child-staff ratio; (b) study design, including experimental or random assignment; (c) population, such as child, teen, or adolescent; and (d) program field, such as youth development or pregnancy prevention. Peer-reviewed journals were searched using various combinations of our search terms. Although priority was given to the most recent articles, we extended our search to as early as the 1980s in an effort to augment the small number of later experimental studies. We also reviewed a handful of articles published in the 1970s that were repeatedly cited in more recent articles as still relevant and valid.

When a study appeared to be a true experiment, relevant to this report and sufficiently rigorous, the article was downloaded or retrieved from a library. For example, for those articles published since the early 1980s with "experimental," "program," and "children" in their titles or abstracts, Child Trends staff reviewed hundreds of abstracts. The combination of the search terms "experimental," "program," and "adolescent" produced far fewer results, leading to reviews of about 200 abstracts of articles published since 1980.

2. Selection of studies

The 35 studies that met all of the following criteria are reviewed in this chapter:

1. Examined the impacts of social program implementation on child and/or family outcomes, or outcomes related to program or provider performance, by deliberately varying at least one aspect of program implementation using a random assignment/experimental methodology.
2. Employed statistical tests to examine whether outcomes were different by any aspect(s) of program implementation.
3. Employed random assignment of experimental and control groups (such as program providers, children, classrooms, and/or sites).

4. Achieved a high level of research quality, determined by examining explanations of study design, sample size, attrition, methodologies for the analyses, and statistical tests. If an article did not articulate how the study was conducted or did not provide sufficient information and evidence that the study was experimentally manipulated and results statistically tested[1], the study was not reviewed in this chapter.
5. Retained a post-program sample size of 30 or more study participants.
6. Provided services for children, adolescents, youth, and/or families with children.
7. Published in the past 25 years.[2]
8. Conducted in the United States.

In the interest of the integrity of this chapter, we were extremely stringent about upholding these selection criteria. Studies had to have met all eight of the criteria listed to be included in this discussion. As a result, all quasi-experimental studies that were found in Chapter II's literature search, as well as some experimental studies that did not deliberately manipulate program features, were excluded from this chapter and reviewed in Chapter III.

3. Limitations of studies

All studies, including rigorous experimental studies, can suffer from methodological flaws. For instance, even when random assignment occurs, it doesn't always result in control and treatment groups whose backgrounds are fully equivalent, especially when sample sizes are small (Dynarski et al., 2003). It is important for studies to control for any baseline differences between groups. When interpreting findings, it is also important to consider the services received by the control group; also, contact with program staff that control participants sometimes receive during data collection could in theory result in accidental, but meaningful, intervention effects. It is important to note that experimental studies reviewed here are not based on nationally representative samples; in fact, the studies are often based on limited populations in specific cities or neighborhoods. Study findings, therefore, cannot be generalized across populations for which they may not apply. For this reason, we are careful to say that findings may be specific to the program being evaluated and may only be generalized to similar programs serving similar populations.

Examples of limitations repeatedly found in these studies are listed below:

- The effects of multiple implementation features were tested in a single analysis without distinguishing one specific feature from others. For instance, experimental and control groups in some studies received different services not only in terms of dosage but also in terms of the timing of the program. With these studies, we would not know which implementation feature specifically

[1] In this review, only results tested at the significance level of 0.05 or better are reported as statistically significant findings.
[2] This chapter was more inclusive than other chapters in terms of acceptable years of publications in order to increase the number of studies available to review.

contributed to outcomes. However, we included information from these studies in this chapter because it identifies a certain "package" of implementation features that may be effective in improving outcomes.

- Although some studies manipulated program components by requiring providers to follow certain rules, they did not necessarily regulate other aspects, such as the content of the program. Because of this, we could not discern with certainty what contributed to differences in the outcomes.
- Although program participants were randomly assigned to different conditions, service providers were self-selected and not randomly assigned in some studies. In such cases, differences in outcomes may have been attributable to the differences in the characteristics of service providers.
- As described above, many studies targeted specialized populations, which may limit the extent to which findings can be generalized. Indeed, several such studies suggest that certain program components may have different impacts across different populations.

4. Methodologies used by studies

We found two major approaches employed by the studies reviewed in this chapter. The first one is a **2-arm study approach** using one experimental or treatment group (E) and one control group (C) (thus it can be expressed as E-C). Under this approach, studies compared a control group and a treatment or experimental group, and the treatment group differed from the control group only by one aspect of implementation and that aspect was tested through randomized manipulation. Control and treatment groups had similar program conditions with one implementation aspect differing.

The second one is a **3-arm approach**. In this type of study, participants were randomly assigned to one of three conditions: two different experimental groups (E1 and E2) and one control group (C) (and thus can be expressed as E1-E2-C). The control group usually received no services while the treatment groups received nearly identical services with the exception of how the programs were implemented. Some studies first looked at differences between one of the experimental groups and the control group (E1-C) and between the other experimental group and the control group (E2-C) separately, and then tested differences between (E1-C) and (E2-C). Other studies included no information on how E1 performed relative to E2. But some authors suggested differences between the two treatment groups if one differed significantly from the control group while the other did not. In cases in which the comparative findings between two treatment groups were described, but not based on experimental analyses, we noted this explicitly.

Only analysis results with a significance level of .05 or better are reported as statistically significant findings.

5. Identifying "positive," "negative," and "neutral" program components

In the following discussion, assessments of program components as "positive," "negative," "neutral," and any combinations thereof are defined as the following:

1. A program component is **positive** if it has a favorable, statistically significant impact on the participants, program, or provider, compared to the alternate program condition.

2. A program component is **negative** if it has an unfavorable, statistically significant impact on the participants, program, or provider, compared to the alternate program condition. Negative findings are generally considered to be those that are undesirable and/or in opposition to the intended program goals.

3. A program component is **neutral** if it has no statistically significant impacts on participant, program, or provider outcomes; that is, it doesn't have any positive or negative impacts, compared to the alternate program condition.

4. A program component is positive to neutral (**positive-neutral**) if findings on it range from positive to neutral, as defined above. In other words, when findings from a set of studies examining a particular program component are divided between the neutral and positive categories, then the program feature is identified as positive to neutral. As there are few experimental studies that manipulate a given program feature, we used this terminology to try to be inclusive of findings from multiple studies.

5. A program component is negative to neutral (**negative-neutral**) if findings on it range from negative to neutral, as defined above. In other words, when findings from a set of studies examining a particular program component are divided between the neutral and negative categories, then the program feature is identified as negative to neutral. As there are few experimental studies that manipulate a given program feature, we used this terminology to try to be inclusive of findings from multiple studies.

C. Summary of Studies by Regulatable or Manipulable Feature

1. Staff and staffing

This section comprises a review of eight studies that experimentally manipulated staff-related practices. Information from these studies can inform program leaders in their staffing decisions—from determining how many people to hire, what educational background or experience to demand, or what training to provide. Information from these studies may also be helpful for directors who are wondering how to structure classes in regard to the number of participants per staff, how many children or youth to enroll in a given group, or what type of facilitators to employ—trained professionals or peer leaders, for example. Although this kind of information is quite important and may

affect how well a program meets its intended goals (of improving child and family well-being), very few high-quality experiments have been published on these staff-related subjects that reveal what staff characteristics and staff-child composition best promote child well-being.

Staffing of program: Staff-child ratio and group size

We found and reviewed one high-quality experimental study that examined the impacts of staff-child ratio; it focused on elementary school-aged children in the Tennessee Project STAR class size experiment (see Mostellar, 1995 for a review of short-term and long-term evaluations of this program).

Tennessee's Project STAR class size experiment was implemented in a school setting, but may have implications for student learning in other settings, as well. Specifically, Project STAR evaluations have shown that children who were in small classes (13-17 students per class) performed better academically on reading and math tests than control group students who were assigned to regular-sized classes of 22-25 students in kindergarten through third grade (Mostellar, 1995). The impacts of the class size experiment were sizeable, with positive impacts found through the seventh grade follow-up (Nye, Hedges and Konstantopoulos, (1999). The follow-up data were based on 4,944 of the original 6,572 students who were randomly assigned to 331 classes in 76 schools.

Formal education and training of staff

With regard to staff training, a few studies of a health promotion and health risk behavior prevention program weakly suggested that trained leaders may be more effective than untrained leaders (Connell, Turner, & Mason, 1985; Hall & Hord, 1987; Nelson, Poehler, & Johnson, 1988). The gains experienced by students of trained leaders were slightly larger than those of students of untrained leaders. The study showed that staff training promoted better fidelity to program guidelines in a substance prevention use program (Ross, Nelson, & Kolbe, 1991).

Type of instructor

It is intuitive—and founded on research (Telch, Miller, Killen, Cooke, & Maccoby, 1990)—that programs benefit from having a human component there to deliver and interpret program information, as opposed to having a video screening alone. Although it must be noted that a gender-specific computer intervention has shown positive impacts for girls, most of the studies reviewed supported the positive impacts of having an instructor to personalize information. Six studies that met our criteria for inclusion used experimental design to study the effects of different instructor types (e.g., peer vs. adult) on child and youth outcomes. These studies covered a variety of substantive areas, targeted children, youth and families, and generally fell in the domain of health and safety. Namely, the studies evaluated programs focusing on decreasing substance use and sexual risk-taking behaviors, and increasing maternal and child health.

One study found evidence that having a health educator was most effective at promoting substance use prevention behaviors (Sussman, Sun, McCuller, & Dent, 2003). Students participating in the program were assigned to one of the following conditions: a group lead by a health educator, a self-instruction group, and a standard care group (which received none of the substance use prevention curricula used with the health educator led group). Results showed that the self-instruction group did not differ from the standard care control group, but students in the health educator led group had fewer instances of drug or tobacco use at the two-year follow-up.

Studies find that peer leaders can be as effective as adult leaders in substance abuse and sexual risk prevention (Jemmott, Jemmott, & Fong, 1998; Perry, Telch, Killen, Burke, & Maccoby, 1983). In a few cases, peer leaders were found to be more effective than adult program facilitators, particularly with certain substance use prevention strategies requiring teaching students skills about dealing with negative social influences (Murray, Richards, Luepker, & Johnson, 1987; Perry et al., 1983). Philliber (1999) reached a similar conclusion in her review of a set of studies looking at peer intervention programs for health risk prevention.

This research was taken a step further and showed evidence that having student groups run by a nominated peer leader, as opposed to a teacher-chosen peer leader or random assignment peer leader, was an effective way to structure a substance use prevention program (Valente, Hoffman, Ritt-Olson, Lichtman, & Johnson 2003). A post-test showed that students assigned to their group of choice led by their nominated peer leader showed improved attitudes and self-efficacy, including a decreased intention to smoke when compared to the group led by the teacher-chosen peer or the group led by a randomly-assigned peer leader. More research is needed to understand how young people can best be used as leaders, as some research indicates that their talents may differ from those of adults (Murray et al., 1987; Perry et al., 1983). Also, research is needed to better understand what peer leaders themselves gain from their leadership experiences.

An experimental study of a home visiting program for low-income women and infants found that when the program was implemented by nurses, it had positive impacts across a number of measures; it had few impacts when it was implemented by paraprofessionals, suggesting that the type of provider who implements the program matters (Olds et al., 2002). Paraprofessionals' impacts on the outcomes were half the size in standard deviation units compared with those of the nurse program, though the differences between the outcomes of the two treatment groups were not statistically significant except for those for children's language development. The study also found that paraprofessionals were likely to complete fewer visits than nurses (6 visits vs. 7 visits during pregnancy and 16 visits vs. 21 visits during infancy, respectively) since the number of visits were not controlled, and families in the paraprofessional program were more likely to discontinue the program than those in the nurse program (48% vs. 38%). Overall, the study showed that both nurse and paraprofessional home visiting treatment conditions resulted in positive impacts in mothers' behaviors and children's development outcomes when compared with women and infants assigned to the control condition who

did not receive home visiting services (Olds et al., 2002). The program was provided during pregnancy and for two years after childbirth, and outcomes were assessed upon completion of the program.

A follow-up study found that the paraprofessional-visited mothers began to experience some, although limited, benefits from the program two years after its end, which included working more, having a greater sense of mastery, and better mental health as compared to the control group (Olds et al., 2004). It must be noted that there were no significant effects for their child's outcomes. This lends support to the delayed positive effects of paraprofessionals on maternal outcomes. Although these impacts for paraprofessionals were found at the follow-up, the maternal and child outcomes for the nurse-visited mothers were significant. Outcomes included greater intervals between the births of their first and second children and less domestic violence from partners. These findings are significant in view of related experimental research indicating that nurse home visiting improves child outcomes into childhood and adolescence (Olds et al., 2004).

Summary of staff and staffings

In sum, experimental research provides some evidence that child-staff ratios and smaller group sizes produce positive child outcomes. Studies also find that peer leaders can be as or more effective than adult leaders in substance use and sexual risk-taking behavior prevention. Furthermore, peer leaders may be even more effective when leaders are chosen rather than assigned to be a leader. Finally, one study suggested that nurses are more effective in implementing home visiting programs than paraprofessionals. Overall, smaller group size and nurses instead of paraprofessionals appear to be positive program components. Also peer program leaders, compared to adult leaders, appear to be positive-neutral components.

Again, it is important to note that these studies looked at programs focusing on very specific groups (by location, sample size, and program type). As such, it is difficult to know if these findings can be broadly applied to all child and youth development programs.

2. Program practices

It is also important for program providers to know if a given program practice—a strategy or method used to administer or enhance program services—has been found to be a positive component, particularly for its particular population or service. For instance, many program designers and directors employ specific program practices—such as video training, case management, cash incentives or parental involvement—in an effort to achieve program goals. Some of these practices, as discussed below, have been found to be more effective than others.

Interactive Program Components

The literature found on the effectiveness of incorporating interactive components into substance use prevention and health programs has yielded mixed results. Four of the six studies reviewed found that having an interactive program component, whether it be a role playing, parent-child communication, peer leaders, etc. had significant positive impacts on program outcomes. Contradictory findings from two of the studies showed that having an interactive component was not always effective at obtaining positive outcomes.

As discussed in the staffing section, Sussman et al (2003) found evidence that having a health educator was most effective at promoting substance use prevention behaviors. Relevant to this section, the health-educator group had a high level of interactive components such as relationship development and social interaction, which was found to have significant effects of substance use prevention at the two-year follow-up.

Another related study found a program that was focused heavily on interactive components, such as parent-teen communication skills and family interaction, had a positive impact on adolescents' intentions to postpone sex in comparison to traditional didactic teaching (Lederman and Mian, 2003). Students were randomly assigned to either the social learning education group (which included interactional components) or the traditional didactic teaching group in order to promote prevention of STDs and pregnancy in middle school youth. It is important to note that no significant impacts were found on other outcomes.

An additional study addresses preventing high-risk behaviors among middle school youth (Flay, Graumlick, Segawa, Burns, and Holliday, 2004). Three separate programs were randomly assigned to African American youth: the social development curriculum, the school/community intervention, and the control group with the health enhancement curriculum. Reduced rates of violent behavior, school delinquency, drug use, and recent sexual intercourse, and the rate of increase in condom use were all significant findings for both the social development curriculum and the school/community intervention when compared to the health enhancement curriculum. For the social development curriculum, an even stronger impact was found on the combined behavioral measure. Although these findings are promising, there were no significant impacts on these outcomes for girls.

A handful of studies have examined the effectiveness of using videos as compared to, or in addition to, other teaching activities. One study on a smoking prevention program showed that the use of peer leaders in combination with video tape training was more effective in preventing the onset of student smoking than receiving video tape training only (Telch et al., 1990). Both of these treatment groups had lower smoking rates than control group youth. Similar findings were found for an HIV prevention program, in which groups receiving an HIV informational instruction video plus a skills training were compared to a group receiving video instruction alone. The video-only approach was found to be as effective in increasing knowledge as the video and skills-training combination (Winett et al., 1993). However, families in the combination group demonstrated an additional increase in family problem-solving behaviors, while the

video-only group did not. Unfortunately, the study did not appear to test for differences between the two treatment conditions.

Findings from a separate study investigating teen smoking cessation (Lipkus, McBride, Pollak, Schwartz-Bloom, Tilson, & Bloom, 2004) yielded different results. This investigation found that providing students with phone counseling in combination with video instruction and written self-help materials did not increase smoking cessation as compared to students receiving video instruction and written self-help materials alone without phone counseling. Although the telephone counseling group did not differ in cessation rates, teens reported that the self-help materials (including the video) and the counseling improved attitudes toward smoking and quitting.

This investigation concluded that a gender-specific computer intervention compared with a conventional intervention (not computer-based) produced better attitudinal outcomes for girls regarding substance use prevention (Schinke and Schwinn, 2005). A total of 91 seventh-grade girls participated and were randomly assigned to one of the two groups: a computer-based program on a CD-ROM, which included video images of adolescent girls sharing their perceptions and experiences with stress and drug use or the conventional intervention which was delivered by teachers and included a session of didactic information, class discussion, and student-volunteered examples of the material's applications to everyday situation. This study showed evidence that providing technology-based programs tailored to girls, may have more positive impacts than the conventional teacher-led programs. This is evidence that interactive program components do not always produce the most effective outcomes, and tailoring interventions based on gender maybe a valuable tool in substance use prevention in adolescents.

Multisystemic Therapy (MST)

Multisystemic therapy has shown promising results in the areas of mental health and juvenile justice. It is an action-oriented intervention that directly addresses intrapersonal and systemic factors (family, peer, and school) and is individualized for each person according to their needs. Three studies reviewed in this section examined the impacts of multisystem therapy versus various other types of intervention options. The studies provide evidence as to the positive impacts that multisystemic therapy has on certain youth populations. There is growing evidence that multisystemic therapy is an effective program practice because of its individualized focus as well as its flexibility to incorporate a wide array of the needs of youth and their families into treatment.

In a study looking at the long term effects on recidivism rates for juvenile offenders between 10 and 15 years later, results found that MST participants had significantly lower recidivism rates at follow-up than did their counterparts who participated in individual therapy and had 54% fewer arrests and 57% fewer days of confinement during the 10 to 15 years following their therapy (Schaeffer and Borduin, 2005).

An additional study found evidence that multisystemic therapy foster care (MTFC) was more effective than group care in reducing delinquency in adolescent girls (Leve, Chamberlain, and Reid, 2005). A total of 81 girls were randomly assigned to one of two groups: MTFC where girls were placed with highly trained foster parents and received individualized treatments, or group care where girls were placed in community-based group care programs. The focus of MTFC was one-to-one adult mentoring and having girls live in a family setting away from delinquent peers whereas the group care programs stressed peer-focused interventions with shift staff. Results at the one year follow up showed that MTFC girls spent significantly fewer days in locked setting and had a greater reduction in criminal referrals compared to the youth in group care. Moreover, this study concurs with Dodge, Dishion, & Lansford (2006)'s Social Policy Report, which examines the ineffectiveness of some delinquency prevention programs due to the negative peer influences that occur when youths are placed in group treatments.

The third study reviewed found evidence that multisystemic therapy improves mental health and social support for youth with serious emotional disturbance (SED) (Rowland, Boykins, Henggeler, Cunningham, Lee, Kruesi, and Shapiro, 2005). This study examined two different programs for 55 youth with serious emotional disturbance and their families: MST adapted for SED populations and usual state services. Although having a small sample size, the study found that youth in the MST condition reported significant reductions in externalizing symptoms, internalizing symptoms, and minor criminal activity. In addition, their caregivers reported near significant increases in social support, and archival records showed that MST youths experienced significantly fewer days in out-of-home placement. It must be noted that there was no significant findings for improved family functioning or a reduction in index offenses or substance use.

Case management vs. financial incentives

An evaluation of a school failure prevention program compared the impacts of no intervention to two treatment conditions—one in which students received cash and another through which students were provided intensive case management (Reid & Bailey-Dempsey, 1995; Reid, Bailey-Dempsey, Cain, Cook, & Burchard, 1994). Both treatment conditions were superior to the control condition. As a program practice to decrease student failure, the case management treatment was found to be more effective than the cash incentive treatment. The study included 112 female students considered to be at risk of school failure. It is important to note that the students who participated in the cash incentive program experienced small declines in self-esteem. The evaluators hypothesized that this was because many of the students improved their grades or attendance, but were not able to meet the stringent criteria of increasing their school attendance and grades by increments of 15 percent or by maintaining above average attendance and academic levels.

Mediation vs. litigation

Mediation is another intervention approach—like case management and program leader involvement, described above—that shows some promise of improving child and

family outcomes. For one study, recently divorced couples with child custody disputes were randomly assigned to resolve the disputes with litigation or mediation (Emery, Laumann-Billings, Waldron, Sbarra, & Dillon, 2001). Follow-up data from both mothers and fathers showed that family functioning outcomes, such as involvement of the non-custodial parent with the child, were favorably impacted when mediation had been used rather than litigation. Fifty-two of the original 71 families (73%) were included in this 12-year follow-up, with nearly equal numbers lost from the original mediation and litigation sample groups.

Parental involvement

Five studies reviewed in this section examined whether involving parents in programs for children was an effective way to improve program outcomes. Given the critical role parents play in children's development and behaviors, parental effects in the program context are a potentially important factor to investigate. Across the different possible types and levels of parental involvement, the most effective way of involving parents may differ by target population or program goal. The size of impacts and cost-effectiveness are of particular importance to service providers and policy-makers given the additional resources required to involve parents in a program. These experimental studies, below, are based in very different substantive areas, with programs addressing literacy, obesity, and attention deficit hyperactivity disorder (ADHD). Four out of the five studies found parental involvement to have positive impacts on youth outcomes; one study found no statistically significant impacts.

An experimental study of a school-based literacy program for 1st-3rd grade children found that involving parents was highly effective in improving children's literacy skills (Morrow & Young, 1997). In this program, parents in the experimental group learned how to conduct literacy activities with their children. The 28 children in the experimental group had much higher post-program scores in five of the six literacy measures than the 28 children in the control group (without the family component). The differences in the scores were large, ranging from 56% to 290% of standard deviations.

Several studies on parental involvement were found in the field of child obesity programs; only two of them met our selection criteria. Unlike literature in other fields, these studies generally did not examine the impacts of adding parental components to a program, but discussed whether a program should focus on children (and their self-regulation skills and knowledge about obesity) or parents. The rationale for focusing on parents differed by study. Some programs helped parents acquire knowledge and change behaviors related to child obesity, or they treated parents' obesity along with children's in order for them to become role models and to change food environments for children. Other programs helped parents acquire control and monitoring skills. These studies, described below, provide useful insights about the way parental involvement may influence children's outcomes in various contexts and in regard to diverse objectives.

One study of a program for obese children (ages 6 to 11) found that a parent-focused approach, in which only parents participated in sessions on lifestyles, diet, food stimuli,

parenting skills and role modeling, had much larger impacts on children's behaviors and weight loss than a child-focused approach in which children were solely responsible for losing weight (Golan, Fainaru, & Weizman, 1998). The differences were fairly large: by the end of the one-year treatment, the average percentage overweight[3] decreased by about 15% for the 30 children in the parent-focused group, and by 8% for the 30 children in the child-focused group.

Similarly, another study on a children's (ages 6-12) obesity-treatment program found that treating obese parents along with children was more effective than focusing on children's behaviors alone (Epstein, Valoski, Wing, & McCurley, 1990). Sixty-seven children were assigned to one of three groups: the first group was "child and parent target," in which obesity of both parents and children was addressed; the second group was "child target," in which the goals and focus of modules were in regard to children, though parents participated in the activities; and the third group was the control group, which had no specific target. The first group was the only group that maintained a significant long-term loss of percentage overweight. At the 5- and 10-year follow-ups, the average percentage overweight of child participants was significantly lower for those who had participated in the joint parent-child target group than for those from the other two groups.

Unlike the previous four examples, a study on a social-skills training program for elementary school children—including those with ADHD—found that the program significantly improved children's social behaviors regardless of parental involvement (Burrow, 2001). Furthermore, results showed that adding a parental training component did not make any significant differences in children's behavioral outcomes. It should be noted that a total of 42 children (23 with ADHD, 19 without) were randomly assigned to the experimental and control conditions, which resulted in very small sample sizes in each group. Small sample sizes may have played a role in producing nonsignificant results.

A separate study on delinquency prevention in aggressive preadolescent boys was conducted. Boys were randomly assigned to the child only component, child and parent component, or the control condition. Significant impacts were found for covert delinquency, parent-reported substance use, and teacher-rated behavioral improvement for both the Child and the Child/Parent interventions. In addition, the strongest impacts were found for the Child/Parent intervention further lending support to the positive practice of incorporating parents into substance use and delinquency prevention programs.

Although the effects of parental involvement may vary by type and level of parental involvement, target population, and program field, four of the five studies discussed indicate that involving parents generates more positive outcomes than not involving parents; furthermore the effect sizes have the potential to be fairly large. The fifth study

[3] "Percentage overweight" refers to the percentage by which a child was calculated to be over the average weight for children of his or her age, gender, and height. Decreasing in percentage overweight does not necessarily correlate with weight loss, as children typically increase their weight as they age.

found parental involvement to have no significant effects, but was undermined by a small sample size. In all, parental participation appears to be a positive to a possibly neutral program practice.

It is important to keep in mind, however, that the conclusion and findings are based on a limited number of experimental studies that originate from very different program types and serve very different specialized populations (e.g., obese children and children with ADHD). Also, it is unknown how the influence of parental involvement may differ for different participant age groups. Because of this, these findings may not be generalizable to other populations or topics.

Summary of program practices

In general, studies suggest that providing group discussions in combination with video instruction seems to be more effective in decreasing problem behaviors in teens and youth, as opposed to video instruction alone. A few studies suggest that case management, compared to financial incentives, can promote more positive academic and behavioral outcomes for teens and youth. Mediation, when used in child custody disputes, was found to produce more favorable parenting and child impacts than litigation. It is interesting to note that all but one of the studies lend support to interactive (rather than non-interactive) program practices. Moreover, interventions incorporating multisystemic therapy show great promise in the areas of mental health and juvenile justice. Finally, all but one program with parental involvement produced more positive outcomes than programs without, and—when produced—impacts were of considerable size.

3. Dosage of program services

This section investigates whether differences in the level of program dosage correspond with differences in participant or program outcomes. The term "dosage" typically refers to the amount of services participants receive. Dosage can vary by, for example, hours of services, frequency of services, and frequency of contact with service providers. Studies used a variety of measures to investigate the dosage of services children and families received. Examples included different methods of contact (e.g., face-to-fact vs. phone contact), different levels of contact with program content and staff, and/or exposure to extra program components.

Twelve experimental studies that investigated dosage were identified in our literature search. These studies were conducted in various different program contexts, including those that address juvenile justice and health education (HIV and substance use prevention). In these studies—discussed below according to type of program or service—control and treatment groups received different dosages of the intervention.

These types of studies may be useful for service providers in deciding how many hours or days of services they should provide for children and their families in order to

maximize the effectiveness of their program. Given the higher costs associated with providing programs with higher dosage and longer duration, the cost-effectiveness of different approaches concerns both service providers and policy-makers. Does a higher dosage of program services lead to more positive outcomes, or can programs designed with lower dosage levels be as effective?

Juvenile justice probation programs

Two studies on juvenile justice probation programs showed limited positive impacts of higher dosage on reducing recidivism; four studies found no significant impacts. It is important to note that participation in most juvenile justice programs is mandatory as part of probation requirements, not voluntary like participation in the majority of other programs discussed in this report. Findings may not necessarily be generalizable to other populations due to the challenging nature of this population.

Sontheimer and Goodstein (1993) conducted an experimental study on a program that provided social casework and surveillance for 90 juveniles; participants had at least one preceding adjudication for serious crimes after their release from a correctional facility. Probation officers in the experimental group were required to have smaller caseloads and more frequent contact with offenders than the ones in traditional aftercare programs. Higher program dosage made no significant differences in the percentage of juveniles who were arrested again, even after controlling for the time youth were observed; however, the average number of re-arrests was significantly lower for the experimental group (1.02) than for the control group (2.07). This means that similar percentages of juveniles would have been arrested regardless of dosage, but those with more contact with probation officers would have been arrested fewer times. It should be noted that the program did not have many requirements dictating its content—for example, there were very few instructions provided on program structure, emphasis, or mission—and these aspects were not controlled in the analyses. Some differences in outcomes, then, may have been due to variations in program content.

A study on a probation program for 14- and 15-year-old females found that higher program dosage was effective for those who were on probation for their first offense, but not among those with prior offenses (Land, McCall, & Williams, 1990). Among those with first-time offenses, the program version with more frequent staff-offender contact was more effective in reducing the probability of new delinquent offences during the two-year period of supervision (12% out of 42 females) than was the traditional care (28% out of 51 females). When only participants with prior offenses were investigated, however, no statistical differences were found between the traditional and higher-dosage groups in the probability of new offenses. It is important to note that the lack of significance may be attributable to an extremely small sample size of participants with prior offenses (total n = 13).

Four other studies, discussed below, also found variations in dosage to have no effect on recidivism among habitual delinquents. One experimental study (conducted at two sites) found that dosage made no significant difference in the percentage of juveniles who

were arrested again or who self-reported drug use during a 12-month follow-up period (Greenwood, Deschenes, & Adams, 1993). The study included more than 50 youth who had an average age of 17, were predominantly black, male, and had prior arrests in each site. Similarly, a study with more than 500 male adolescents—mean age of 15, most with prior charges for serious crimes—found that higher program dosage did not make any difference in the number of new criminal charges at the end of a two-year follow up period (Barton & Butts, 1990). In the experimental group, youth received extra educational and recreational services in addition to regular supervision, and caseworkers had a limited number of youth. At the end of a two-year follow-up period, the average number of criminal charges was in fact higher for the experimental groups than the control groups (1.85 vs. 1.17). However, youth in the control group committed more serious crimes and were incarcerated for a longer period of time than those in the experimental group, which limited opportunities for youth to commit more crimes. Therefore, after the time youth were not incarcerated was controlled, the differences in the average number of charges between the control and experimental groups became non-significant. Two other studies, reviewed in an article by Sontheimer and Goodstein (1993), also found no impacts of higher program dosage (e.g., increased face-to-face or phone contact with program staff) on the percentage of juveniles who were re-arrested (National Council on Crime and Delinquency, 1987; Wiebush, 1991).

Health education

The three studies described below are based on HIV and substance use intervention programs. Two of the three studies found that both enhanced (e.g., intensive case management, high interaction with program) and standard programs had more positive impacts on behavioral outcomes than the control groups, which had minimal or no intervention; however, the size of the impacts did not differ significantly by dosage. The third study found small, yet statistically significant, differences in impacts between an enhanced program and a standard program. It should be noted that in all three of these studies, the participant samples comprised a combination of children, adolescents, and adults.

First, an experimental study of a counseling program for HIV and sexually transmitted disease (STD) prevention had more than 5700 males and females (including adults) who were randomly assigned to: (a) a high dosage program consisting of enhanced counseling and four interactive, theory-based sessions (200 minutes total) provided over three to four weeks; (b) a lower-dosage program with brief counseling and two 20-minute interactive sessions provided over 7 to 10 days; or (c) a control group that received two brief clinician-administered educational messages about HIV and STD prevention that are typically seen in current clinical settings. The program was provided at inner-city STD clinics in six major cities for HIV-negative patients ages 14 and older. found that a program with different dosage levels, the enhanced counseling group and the brief counseling group, produced positive impacts on sexual behaviors and HIV/STD contraction rates throughout the 12-month follow-up period, but the differences in the size of impacts on most of the outcomes were not statistically significant between the two groups (Kamb et al., 1998). The only exception was in condom use where the enhanced

counseling group showed significant increases in condom use at the 3-month and the 6-month assessment compared with the brief counseling group, but the significant differences disappeared after the 6-month follow up. The second study—based on an HIV intervention program for drug users—found that when extra service components such as individual counseling and behavioral skills training were added to a standard two-session intervention, it produced small but statistically significant decreases in HIV risk behaviors and drug use behaviors (Rhodes & Malotte, 1996).

The third study examined dosage-based differences in a three-month-long alcoholism treatment program, in which one group received four sessions and another received 12 (Project MATCH Research Group, 1998). Results show that both low- and high-dosage interventions produced positive but nonsignificantly different impacts on risk behaviors, suggesting that the difference in dosage did not appear to matter. Importantly, study authors speculate that follow-up participant interviews at 3, 4, 9, 12 and 15 months—which were conducted with all of the treatment groups—may have diminished the dosage impacts; participants in the "lower dosage" group were, in fact, receiving the same amount of contact from individuals associated with the program as the "higher dosage" group for one year after program completion.

Mental health

One study explored an intensive case management strategy for children with serious emotional disturbance (SED). The evaluators found assignment to intensive case management, as compared with assignment to a regular mental health clinician with larger caseloads, was related to better program implementation and outcomes, as indicated by longer participation, wider use of community-based services, and use of fewer inpatient days (Burns, Farmer, Angold, & Costello, 1996).

Family reunification

Two experimental studies on one family reunification program found that the program with higher dosage produced more positive program impacts in reunifying foster children in foster care with their biological parents than the program with lower dosage (Fraser, Walton, Lewis, Pecora, & Walton, 1996; Walton, Fraser, Lewis, Pecora, & Walton, 1993). Home-based support services were provided over 90 days with different dosages (e.g., time spent with each family and frequency of visits). Caseloads per caseworker were also smaller for the experimental group, permitting more dosage per client. The differences in the outcomes were large: the first study showed that at the end of the 90-day program period, 93% of the families in the experimental group, compared with 28% of those in the control group, were reunited (Walton et al., 1993). The second study showed a similar pattern of dosage impacts (Fraser et al., 1996). It should be noted, however, although families were randomly assigned to experimental and control groups, caseworkers were self-selected. Therefore, the differences in the outcomes could possibly be attributed to the differences in the characteristics of the caseworkers.

Summary of dosage

On the other hand, six of the twelve experimental studies we reviewed found positive impacts associated with higher levels of program dosage. Studies on mental health programs and family reunification programs showed that programs with higher dosage had more positive impacts than programs with lower dosage.

Six of the studies we reviewed found no or very limited impacts relative to variations in dosage. The six studies on juvenile justice probation suggested there can be positive impacts of higher dosage on reducing recidivism, but the impacts are limited. Some of the study authors argue that the dosage of services was simply not high enough to generate results significantly different from the control groups' (Sontheimer and Goodstein, 1993). Others argue that higher dosage alone is not enough to make changes, but that the nature of services provided during the increased contact also needs to be addressed in order to have a significant impact on habitual behaviors. For example, Greenwood et al. (1993) suggest that the program needs to provide services that specifically address delinquent behaviors such as anger management and substance abuse treatment in order to make differences in chronic delinquent behaviors. Similarly, studies in the HIV and substance use intervention fields found very little or no differences between the impacts of standard and higher-dosage interventions, but enhanced interventions were typically short-lived and dosages did not vary greatly between the intervention types.

Overall, studies appear to suggest that higher dosage often does not always appear to make a significant difference in outcomes; however, when it does have an impact it is always positive, never negative. Higher dosage appears to be a positive to neutral program component. It is important to note, however, that findings differed greatly by substantive field, and that it may be the combination of higher dosage or longer duration along with the nature of services provided during this increased contact that determines effects. More research is needed in different program contexts, on different populations, and in regard to which program content is manipulated as rigorously as duration and dosage.

D. Chapter Summary

It is difficult to draw firm conclusions based on the limited number of available experimental studies, which were set in a variety of very different substantive areas and designed to serve very different populations in terms of risk levels, age groups and gender. The findings from this assortment of studies may not be applicable in different contexts.

However, these studies provide some promising evidence that certain manipulable features can favorably impact child and family well-being. In the area of staffing, experimental research shows that smaller group sizes, and staff with child-related training are **positive** components in programs serving young children. In the field of risky

behaviors prevention for teens, trained peer leaders appear to be as or more effective leaders than trained adult leaders (**positive-neutral**). Video instruction with an interactive component (e.g., an instructor, skills-building training) appears to be a more promising mode of instruction in positively affecting a variety of outcomes than video instruction alone with the exception of one study; however, video instruction is able to impart informational advice even when used alone (**positive-neutral**). However, a technology-based program (without an interactive component) has promise when tailoring the program specifically to girls (**positive**). Additionally, interpersonal interventions, such as case management and mediation, appear to be more effective than cash incentives and litigation, respectively, in impacting child and family functioning outcomes (**positive**). Furthermore, multisystemic therapy was more effective than other standard community treatments and/or therapies when treating specific populations (**positive**). Research also provides evidence that programs that involve parents can positively impact child outcomes; however, one study found that it does not make a difference (**positive-neutral**).

Findings about the effects of dosage of services varied from **neutral to positive** across the substantive fields and target populations. It may be that the combination of higher dosages with a specific type or quality of services provided during the extended period that is important to produce favorable program outcomes.

1. The ideal program, according to experimental studies

Based on experimental evidence, what can we conclude about how to ideally implement an intervention program for children or adolescents? Most of the questions that a program provider would want to have answered, unfortunately, cannot be answered on the basis of the literature currently available, because the effectiveness of many manipulable program features has not been experimentally tested. Furthermore, many of the studies that have been conducted were for very specific types of programs (e.g., literacy programs, adolescent health risk prevention programs) or for programs serving specific populations; and findings therefore, cannot be generalized to the broader population. However, some conclusions may still be drawn (keeping the aforementioned caveats in mind).

An ideal program would have a smaller group size, and the staff would have training that is specific to that program and to child and youth development. In addition, the program would not use a solely didactic approach, but would involve participants in interactive activities. In addition, discussion or activities and not just presenting a video seem to be valuable. An exception might be computerized programming intended for a specific gender, particularly girls. Also, intensive case management or multisystemic therapy seems effective, and the program would seek to involve parents. On the other hand, there is some evidence that it is alright to use trained peers to deliver a program activity, not just trained adult leaders.

2. Future research

Extant experimental research seems to be clustered in very specific problem prevention areas, particularly in areas of health risk behavior prevention. Studies are needed not only for different populations and different formats but also for diverse fields. For example, we were unable to identify any experimental studies that examined the impacts of parent involvement in pregnancy prevention programs. Such studies may provide important insight for the field given the implications of parental influences on teenagers' sexual behaviors evidenced in recent literature (Kirby, 2001): positive relationships with parents, "appropriate supervision and monitoring," and parental conservative attitudes about sexual behaviors were found to have significant influence on teenagers' sexual behaviors. Parental involvement studies in teenage pregnancy prevention field may reveal more complicated pictures due to the confidential and personal nature of the subject.

Experimental studies are needed to understand with certainty what aspects of program implementation are effective or not. Hundreds of articles were found for the following chapters, which are based on non-experimental research and practitioner's opinions, while this chapter is based on 35 studies. It is troubling that so little experimental evidence is available to programs as they make staffing, duration, and other important decisions that could impact their level of effectiveness. While it is not possible to do experimental research on every manipulable program feature, such as safety, much more research is needed. Further research on a more broad set of programs, especially positive development programs, examining a more diverse set of outcomes, is necessary to draw firm conclusions about how aspects of program implementation affect child and family outcomes.

III. FINDINGS FROM NON-EXPERIMENTAL RESEARCH

By Sarah B. Garrett and Lillian Bowie

A. Introduction

Chapter III draws from non-experimental research studies to address our questions about program implementation. Significantly more material falls into this category than that of the more rigorous standards of Chapter II; there is a great deal of information to share from this type of research. However, from the perspective of asserting causality, the sources in this section, are less rigorous and definitive. Therefore, findings should be viewed as promising practices or approaches, not as definitive proof that certain regulatable or manipulable features do or do not "work." This discussion is meant to complement and not overshadow, despite the greater volume of sources, the causal findings identified in Chapter II.

This chapter is divided into the following categories: 1) Staff Demographics and Professional Support staff; 2) Program Design; 3) Program Practices, and 4) Conceptual Approaches. Related subtopics of each of these categories feature discussions on findings, relevant caveats, and conclusions.

B. Methodological Approach

1. Literature search

As described in Chapter II, we conducted an extensive literature search, using EBSCO Host to access studies from peer-reviewed journals published between the early 1980s and 2006. If any title looked promising, the abstract was reviewed and the article was downloaded or collected from local libraries. Please see Chapter II for a more detailed discussion of this process.

2. Selection of studies

The publications that inform this section are based on programs that provided services for children, adolescents, and/or families with children, and that fit at least one of the following criteria:

1. Quasi-experimental studies that manipulated at least one aspect of program implementation and compared the resulting effects (e.g., child and youth outcomes, outcomes related to child or youth well-being, program performance) in similar, but not randomly assigned, groups. Statistical tests were typically conducted to examine whether outcomes were different by one or more aspects of program implementation.

2. Experimental and quasi-experimental studies that were not designed to investigate or manipulate variations in regulatable features, but which conducted analyses relevant to this discussion. For the most part, these studies were designed to investigate general program effects, but, in later stages, revealed variations in program implementation and accompanying variations in outcomes.

3. Studies that investigated the effects of particular regulatable and manipulable features through multivariate regressions. Findings from these studies show which factors, if any, significantly predicted certain outcomes when potentially influential factors (e.g., child or family background, program site characteristics) were controlled.

4. Meta-analyses. Meta-analysis has been defined as a "specific methodologic and statistical technique for combining quantitative data' that produces a statistical inferential statement such as a confidence interval or a statistical test" (Mulrow, Cook, & Davidoff, 1998, as quoted and discussed in Beckett, Hawken, & Jacknowitz, 2001).[4] Meta-analyses typically standardize effect sizes—or, occasionally, a different statistic—that are reported in multiple studies on the same type of program (e.g., a collection of different studies on drug prevention programs that were published between 1990 and 2000). Standardization is done in order to produce a single statistical statement about how large an effect a certain type of program has on select participant outcomes. While statistical significance is important, it does not provide any information about the size of the effects; meta-analyses can indicate that a program or component has a statistically significant effect on a particular outcome, but the effect size (or magnitude of the effect) may be minimal. Effect sizes are important in that they are able to suggest whether the size of the effects for a program are large enough to result in long-lasting effects on program participants' lives. In cases where effect sizes are large enough to have a meaningful effect on children's lives, they can be said to be of practical importance.

5. Formal reviews of program evaluations and studies were also included. Certain procedures were established in order to avoid over-representing findings; in particular, all or some conclusions of certain reviews were excluded from this discussion if their findings were based heavily on studies that we had reviewed and chosen not to use, or that we had included independently in our discussion (e.g., Beckett et al., 2001; Dungan-Seaver, 1999; Fashola, 1998; Vandell & Wolfe, 2000).

6. A handful of earlier studies are referred made reference to in the discussion, but, in the interest of time, we prioritized reviewing and discussing studies from 1990 to the present.

Studies determined to be appropriate for this section were reviewed and organized according to the regulatable or manipulable feature by which outcomes differed (see

[4] This definition may have some room for flexibility, however, as certain studies have used this approach without abiding by all aspects of its formal definition (e.g., Beckett et al., 2001, included studies that did not meet the standards for traditional meta-analysis because the number of those that did was too small on which to conduct analyses).

Appendix B). In certain cases, findings that fit our criteria were not included if the variation they addressed was not addressed by any other selected studies. These "lone" studies" are valuable in that they provide interesting ideas on which to base future research; these will be discussed at the end of this chapter.

Substantively, there are important differences between the previous chapter and this one. The experimental design of studies in Chapter II allows us to identify causal links between variations in implementation and variations in outcomes. This cannot be done with non-experimental findings. At best, sources in this chapter provide information on the way a certain variation affected two similar, but not randomly-selected, groups. For example, two classrooms of roughly the same size and demographic makeup may receive the same anti-violence program, one for two hours a week, the other for one hour a week. Differences in outcomes could be due to the different dosage of the program, or they could be due to characteristics of the particular class—the teacher's management style, for example—that may have made it more or less receptive to the program. Sources in this chapter also provide information at the level of associations and correlations, which may say as much or more about selection issues as it does about implementation effects. In the evaluation stages of a family therapy program, for example, researchers may find that parent-adolescent pairs that attended the program more frequently had more positive outcomes. This association could mean that greater dosage led to more positive outcomes, or the relationship could be spurious; that is, the characteristics that led those pairs to attend more frequently than others (e.g., strong commitment to developing a positive family dynamic) may also work to improve the parent-child relationship, independent of the program. Accordingly, caveats will be cited throughout the following discussion.

3. Identifying "positive," "negative," and "neutral" program components

In the following discussion, assessments of program components as "positive," "negative," "neutral," and any combinations thereof, are defined as the following:

1. A program component is **positive** if it is significantly associated with favorable program, participant, or provider outcomes in studies or meta-analyses *after at least some potentially influential participant, program, or location-based characteristics are statistically controlled*. Controlling for other factors that may influence study outcomes as much or more than the component in question—such as participant gender or family SES, or program center characteristics such as location, cost of services, or overall quality—reveals whether the component has a distinct, independent influence on program outcomes. In the few cases where one study controls for certain characteristics and finds the component-outcome pattern(s) to remain significant, while another controls for the same or different characteristics and finds the association to be nonsignificant, the assessment will reflect the possibility that other influential factors drive the association by being linked with "-neutral" (see #4).

2. A program component is **negative** if it is significantly associated with undesirable[5] program, participant, or provider outcomes in studies or meta-analyses *after at least some potentially influential participant, program, or location-based characteristics are statistically controlled*. A program component is also considered (implicitly) negative if a component to which it is the diametric opposite has been identified as positive (e.g., lower levels of caregiver experience are associated with higher classroom quality ratings, therefore, in this case, higher levels of caregiver experience are implicitly negative);

3. A program component would be **neutral** if it were not significantly associated with any outcomes—positive or negative—in studies or meta-analyses. However, no components have been identified as completely neutral in this chapter;

4. A program component is positive to neutral (**positive-neutral**) or negative to neutral (**negative-neutral**) if it is significantly associated with positive or negative outcomes—or identified as positive or negative in reviews—but loses this association when other influential participant, program, or location-based characteristics are statistically controlled. Or, an assessment of positive- or negative-neutral will be given if one study controls for certain characteristics and finds the component-outcome association(s) to remain significant, yet another controls for the same or different characteristics and finds the association to be nonsignificant. In both of these cases, components identified as positive-neutral or negative-neutral are associated with positive or negative effects, respectively, but these associations may be driven more by other factors that accompany the particular components than the components themselves. In essence, the effect of a component is neutralized when other factors are taken into account. For example, a study may find that higher staff wages are associated with positive program outcomes; however, when unobserved child care center-based characteristics are statistically controlled, the association between staff wages and program outcomes is no longer significant (Blau, 2000). Because these positive-neutral and negative-neutral components are associated with positive or negative outcomes to some degree, and are most likely an indicator of and linked to positive or negative components, they are considered promising or harmful to the program, respectively. However, they carry less weight than the positive or negative components (described in 1 and 2, above) whose effects are not tempered by other factors.

   Additionally, studies that did not control for other potentially confounding factors, as well as literature and evaluation reviews, contribute to our assessments. However, if their methodology is flawed (e.g., self-selection effects, spurious relationships), their findings are categorized only as positive-neutral or negative-neutral (not positive or negative). Similarly, if the studies reviewed vary in quality, even if authors literature review identify a given component as positive or negative based on a review of important characteristics of the studies on which the review is based (e.g., quality of the research study designs, appropriateness of measures used, types of statistical

[5] Undesirable, in this case, could refer to unwanted outcomes as perceived by social norms, and/or relative to identified program goals.

analyses conducted, consistency and/or magnitude of findings, etc.), the component is considered here as positive-neutral or negative-neutral, respectively.

C. Summary of Studies by Regulatable or Manipulable Feature

1. Staff Demographics and Professional Support

Staff demographics and professional support variables are salient to virtually every program. As evident in this and the previous chapter, staff and staff-related issues are perhaps the most studied elements of program implementation. Specifically, this section explores the relationship between program staff demographics characteristics including age, gender, and racial/ethnic identity and that of program participants. This section also examines several non-experimentally manipulated studies on issues of staff training, screening, and supervision.

Staff Matching: "age, race, and gender"

There has been great debate on whether program messages are more convincing or relevant if delivered by individuals similar to the target audience. Years ago, race-, gender- or age-based matching of program leader to participants may have led to more positive outcomes (e.g., Dembrowski, Lasater, & Ramirez, 1978, cited in Jemmott, Jemmott, Fong, & McCaffree, 1999), but, if so, this pattern does not appears to be as strong today.

Several studies suggest that matching based on demographic characteristics does not appear to affect program outcomes consistently or at all. In a study of 496 inner-city African American adolescents in HIV prevention interventions, program effects "did not vary as a function of the facilitator's race or gender, participant's gender, or the gender composition of the intervention group" (Jemmott et al., 1999). The matching hypothesis was not supported, and "there was even a moderate negative correlation between the number of matching factors and participants' reactions." Similarly, telephone interviews of 669 mentors revealed that cross-ethnic and cross-gender matches were considered just as close and supportive as matches of the same ethnicity or gender (Herrera, Sipe, McClanahan, Arbreton, & Pepper, 2000).

Similarly, with regard to gender, a study involving 157 African American male adolescents in HIV/AIDS prevention programs found that youth did not experience substantively different outcomes whether they were randomly assigned to male or female program facilitators (Jemmott, Jemmott, & Fong, 1992). Outcomes by gender differed somewhat—youth with male facilitators experienced short-term increases in knowledge, and those with female facilitators reported more positive attitudinal and behavior outcomes at a three-month follow up—but study authors concluded that the results "provide scant support for the view that matching [by gender]… enhances the effectiveness" of the program (Jemmott et al., 1992).

Recent research suggests that matching based on age or ethnicity may in some instances benefit certain demographic groups, but not others (McCurdy, Gannon, & Daro, 2003). For example, ethnicity-based matching with a family support worker was associated with higher participation in a home-based family support program for African American mothers, but not for Latino[6] or Caucasian mothers. Similarly, closeness in age between the mother and the provider was important for participation of Latino mothers only. Provider-mother similarity (with regard to race, age, education, marital status, residential location, etc.) was not associated with any differences in participation by Caucasian mothers (McCurdy et al., 2003).

Finally, program reviews, largely based on mentoring relationships, find little effect attributable to participant-provider matching. A 2001 review of mentoring program evaluations found "no significant differences… in the rates of interaction (i.e., longevity of the Big Brothers/Big Sisters [BBBS] match and rate of interaction between volunteers and youth) occurring in the same-race and cross-race matches" (McGill, Mihalic, & Grotpeter, 2001). A 2002 review of a handful of mentoring programs concluded that, taken together, the studies on which it was based did not suggest that racial similarity determines the strength or potential effects of the mentoring relationship (Jucovy, 2002). Based on the typical mentoring program where there are fewer minority mentors than minority participants, the authors posited that the slight differences they encountered might have been due to parents or guardians who are "most worried about their child" hurrying their child into matches regardless of race, while parents less worried about their child may have felt comfortable waiting for a racially-similar mentor (Jucovy, 2002). Another review of mentoring program finds that cross-race mentor-mentee relationships appear to be just as successful as same-race matches for a variety of academic outcomes (Jekielek, Moore, & Hair, 2002). Finally, a review based on ten Public/Private Venture mentoring reports from 1998 to 1995 found that no "objective factors," in this case described as age, race, or gender, "correlate[d] very strongly with either frequency of meeting, length of match or effectiveness" (Sipe, 1996).

Professional Support: training, screening and supervising

The training of staff in skills and information specific to the programs they will deliver and populations they will serve—through periodic staff development workshops or course of study (e.g., B.A. in Early Childhood Education), for example—may be just as, or more, beneficial as attainment of formal general education. A study based on over 3,500 school-based prevention programs—provides information on the importance of staff training in programs for older children. Analyses from this study show that both the quality (e.g., clear presentation, use of examples in discussion) and quantity of program-specific provider training were significantly correlated to the proportion of best practices providers used in the implementation of various prevention programs (Gottfredson & Gottfredson, 2002). Best practices were in regard to methods employed in, and the substantive content of, the programs (see Appendix B for a detailed description).

[6] Some evidence suggests that Latinos for whom English is not their primary language may benefit from having Latino providers.

As certain programs engage adults who are untrained and inexperienced in child- or youth-related activities, the selection, training, and supervision of these adults emerges as a necessary program component. The selection of adult volunteers is the initial step in the engagement. Studies and reviews of mentoring literature have found that programs that employ comprehensive participant screening (e.g., regarding schedules or ideological compatibility to program goals) to ensure the compatibility of the potential mentor and mentee are associated with more positive relationships than those that do not employ similarly detailed screening (Herrera et al., 2000; Sipe, 1996). After selection, volunteers in mentoring programs appear to benefit from initial and continued training and supervision. Programs that offer support and training before and after the match are associated with more positive mentor-mentee relationships than programs that engage in minimal contact with the mentor (Sipe, 1996).

A review of the Big Brothers/Big Sisters programs identified supervision of mentoring relationships as the program practice "most associated with positive [mentor-mentee] match outcomes;" in fact, the authors of the review found that "those sites following national procedures for regular supervision had matches that were meeting at the highest rates" (McGill et al., 2001). Another review concluded that the frequency of contact from professional program staff was associated with more regular mentoring and higher quality mentor-mentee relationships (Sipe, 1996). Furthermore, the lack of regular contact by program staff was associated with "failed" matches ("those that did not meet consistently and, thus, never developed into relationships"). A study of almost 700 mentors found that the "critical level" for program preparation in order to achieve the most close and supportive relationships was either more than six hours of pre-match orientation and training, and "at least two hours of post-match training, or a minimum of monthly contact with program staff" for new and aspiring mentors (Herrera et al., 2000). The authors point out that it is not known what type of training, or what kind of topics to be covered in it, are most important to the mentoring relationship.

Summary of Staff Demographics and Professional Support

Overall, and particularly in the realm of mentoring programs, it appears that matching based on gender and race does not appear to produce more positive participant outcomes; it appears to be a **neutral** program practice.[7] Indeed, a study based on telephone interviews of 669 mentors suggests that, of all of their measured characteristics, the only similarity that mattered was in identified interests between the mentor and mentee (Herrera et al., 2000).

With regard to professional support, the findings are correlational, however, it is possible that the relationship between the screening/training practices and positive relationships may be spurious. Programs that have sufficient resources to conduct thorough screening and training exercises, or which have prioritized these exercises, may be more likely to generate more successful relationships; in this case, it could be these

[7] This conclusion may be particularly relevant to the ongoing debate in the field of nurse home visiting programs, many of which have invested in matching nurses with clients based on demographic characteristics.

characteristics, and not the preparation of the adult participant, that most strongly influences program outcomes. In addition, the screening, training, and supervision of adult volunteers have been positively related to more successful relationships with program participants in two reviews and one study. We conclude that the informed selection, comprehensive preparation, and ongoing staff support of adult mentors are **positive to neutral** program practices.

2. Program Design

Organizational Capacity

Organizational capacities of programs and agencies have been found to influence the implementation and success of programs though research is limited. For example, behavior change was associated with the number of home visits delivered to participating families. Specifically, those sites delivering fewer visits also demonstrated the lowest levels of behavior change. The study found that both process and structural organizational capacities were also related to behavior change. For instance, when the capacities of the program were low, this was evident in behavior change. Additionally, high structural capacities were positively associated with the number of modules delivered by the parent educators (Kelly, et al., 2004).

Group size

Group size—the number of children or youth who participate in the program or service—is one of the few program components that is relatively easy to manipulate and, thus, this is the subject of two studies.

Larger group sizes have been associated with less positive outcomes than smaller group sizes. A study of 30 mostly-Caucasian afterschool programs in and around Madison, WI, found that participant reports of "poorer program climate" were significantly associated with larger enrollment sizes (Rosenthal & Vandell, 1996). A study of youth and young adults in the Youth Corps community service programs found that sites with 50 or more participants "tended to have lower completion rates than their smaller counterparts" (Jastrzab, Blomquist, Masker, & Orr, 1997). Finally, one study found that after the addition of two school-age children to a family child care home, ratings of provider sensitivity were lower than they had been before the addition; however, it is important to note that this change is conflated with change in child-staff ratio, and that environmental ratings and activities were similar before and after the addition (Howes & Norris, 1997).

There are several caveats to consider in regard to findings about group size. Howes et al. (1992) have observed that group size may be "a marker variable for other unmeasured influences including teacher training and commitment of the child care center's sponsoring organization to provide quality care." This speaks to the worry that group size has a spurious relationship with positive outcomes, and predicts the findings that controlling for certain characteristics may neutralize effects (e.g., Blau, 2000). It

may be that center-based characteristics, such as those cited by Howes et al., determine group size as well as child or youth outcomes, and that the two do not influence each other. Also, studies illustrate that group size may affect some outcomes (e.g., developmentally appropriate activities) but also found that it did not affect others (e.g., appropriate caregiving). Furthermore, Jastrzab, Blomquist, Masker, and Orr (1997) posited that in elective youth programs, the link between smaller group size and positive outcomes could reflect the effects of other unmeasured or unnoticed factors, such as more selective screening practices employed in the case that the program's enrollment targets are lower (e.g., if the program's proscribed enrollment to secure funding is low).

These caveats raise doubts about the specificity of the conclusions that can be drawn regarding the group size, particularly when other typically unmeasured factors such as community and center characteristics are not accounted for. Nevertheless, based on the available research evidence, we find group size to be a generally **positive but sometimes neutral** feature of successful program implementation. This conclusion draws most from studies of young children, but also speaks to programs for school-age children and youth. Work on group size among older children and youth is sparse, excepting the experimental Stars program reviewed in Chapter II.

Degree of Structure vs. Flexibility in Programs

In the field of program evaluation, the meaning of both "structure" and "flexibility" tends to be ambiguous and differs by study. Flexibility has been described as "the freedom of children to choose among an array of interesting activities (or to select being alone if desired)" (Beckett et al., 2001). Dungan-Seaver (1999) concludes that "structure should mean identifiable activities, progress and boundaries, but should not mean regimented participation or outcomes." It is possible that structure and flexibility are not necessarily oppositional qualities. Furthermore, the importance of structure may be moderated by age: a review of several studies of after-school programs concluded that, in regard to structure, "younger children appear to benefit from more [structure], older children and adolescents from less" structure (Dungan-Seaver, 1999). Our assembled research appears to reflect this pattern.

Flexibility appears to be associated with positive outcomes among school-age children. Pierce, Hamm, and Vandell's (1999) study of 150 after-school programs found program flexibility to be significantly associated with better social skills for first-grade boys, net of family structure. In a meta-analysis of after-school programs, social skills were significantly higher for those involved in more flexible programs; however, written language grades were significantly lower for this group (Fashola, 1998).

In sum, the extant research indicates that **flexibility is a positive-neutral feature in programs for school-age children**, particularly, and possibly exclusively, with regard to social skill development. **Structure appears to be a positive-neutral feature in programs for younger children**, particularly, and possibly exclusively, in regard to cognitive development.

Information-only vs. more interactive

Research finds information-only programs to be largely ineffective. The information-only approach, as described in the following studies, is didactic and minimally interactive, if at all. Approaches that are designed to be interactive are not necessarily the diametric opposite of information-only programs; however, the texts reviewed below compare the two and therefore inform the comparison. One description of an interactive program is that they "provide contact and communication opportunities for the exchange of ideas among participants and encourage learning [program-specific] skills" (Tobler et al., 2000, p. 287).

Multiple meta-analyses by Nancy Tobler show that, generally, the more a program allows for participant interaction, the more successful outcomes result. Comparison of the weighted effect sizes in Tobler's (2000) most recent meta-analysis, (also discussed above) shows non-interactive lecture-oriented programs to have minimal impact compared to interactive programs. When examined by specific prevention focus, programs with interactive approaches were significantly more effective than those with non-interactive approaches in decreasing both tobacco and general substance use. There was no significant difference, though, between the two approaches for programs that targeted alcohol use, which elicited smaller overall effects than the tobacco and general substance use programs. Regardless of the number of participants in the program, or whether the program school was "less than half white and/or [had] problems (e.g., high absenteeism or dropout rates, poor academic records," programs with interactive approaches were associated with significantly greater effectiveness than those with non-interactive approaches (p. 289). The same pattern was found regardless of degree of program attrition or the quality of research design, as well (Tobler et al., 2000). It should be noted that at the point at which programs' group sizes reached "a few thousand students," the effectiveness of interactive and non-interactive approaches was not significantly different (p. 315).

Tobler's 1992 meta-analysis of 97 drug prevention programs reveals complementary findings. Programs with a knowledge-only approach had a substantially lower effect size than peer-led programs, which typically feature high levels of participant interaction (Tobler, 1992). The other approaches that were similarly ineffective were an approach that offered no information on drugs or drug use and instead focused on psychological risk factors (the "affective approach"), and a combination of the knowledge-only and affective approach (Tobler, 1992). Earlier work by Tobler (1986) found that knowledge-only programs had positive effects on knowledge outcomes, but "negligible change" for attitudinal or behavioral ones.

A review of findings from sexual health programs also concludes that education-only approaches were less effective than behavior skills training at reducing unprotected sexual activity and deferring sexual initiation (Leigh & Andrews, 2002). Similarly, an overview of research on pregnancy prevention programs from the last 20 years concludes that knowledge-only approaches work to increase knowledge, but "they do not significantly change sexual or contraceptive behavior" (Kirby, 1999). Instead, the review

showed that effective curricula "employed a variety of teaching methods designed to involve the participants and have them personalize the information."

In sum—even though they succeed in imparting information—we find that **information-only approaches appear to be a neutral program component** when compared to the relative effectiveness of more interactive approaches. Conversely, **interactive approaches appear to be a positive program component**, even net of characteristics such as school size and population group characteristics, compared with non-interactive approaches.

3. Program Practices

Dosage : Variation of Time, Curriculum, and Attendance

Dosage can be understood in a variety of ways. For the most part, the following discussion considers dosage as either a variation of quantity of time (e.g., 30-minute versus hour-long sessions) or frequency of attendance or participation (e.g., one versus three meetings per week). Dosage is distinct from duration, which means the length of time one participates in a program (e.g., as measured on a calendar). Duration will be discussed in the subsequent section.

Variation of quantity of time: In a study of a gang prevention program, more frequent attendance was significantly associated with a variety of positive outcomes: less likelihood of starting to wear gang colors and "increased levels of positive peer and family relationships" for those not in gangs; less stealing with gang members, less engagement with most measures of "gang-associated behaviors and peers," "lower incidence of being sent away by the court," and "greater expectations of graduating from high school or receiving a GED" among those already in gangs (Arbreton & McClanahan, 2002). In a study of a gang involvement intervention program, in which youth were already engaged in gang activity, more frequent attendance was associated with several positive outcomes as well: disengagement from gang-associated behaviors and peers; less contact with the juvenile justice system; and greater expectations of graduating from high school or receiving a GED (Arbreton & McClanahan, 2002).

More recently, Harnett and Dadds (2004) evaluated the implementation of a Universal Prevention of Depression program conducted in two independent girl schools in Brisbane, Australia. A 27-item self- report questionnaire was used to evaluate the program. Training was delivered to participants over a 1-day period that was essentially supposed to be a two-day training session. In evaluating the relationship between the quality of program implementation and student outcomes, results showed that students exposed to more key concepts responded better to the intervention than students exposed to a lower percentage of key concepts during the 1-day session (Harnett et al., 2004).

A study conducted by Cervantes and his colleagues (Cervantes et al., 2004), tested the effectiveness of a culturally focused juvenile and substance abuse intervention

program for first time Hispanic youth offenders. The analysis found a significant interaction between length of intervention and legal knowledge. Participants in the four session intervention had a larger increase in legal knowledge in comparison to the three-session group.

Dosage of time in a program proved to be important in mentoring programs as well. Michael Karcher (2005) conducted a study on the effects of development mentoring and high school mentors' attendance on their younger mentees' self-esteem, social skills, and connectedness. Using a pre-post randomized experimental study design, the study showed that the relationship between mentors; inconsistent attendance and mentees' decline in self-esteem and behavioral competence suggests that absent mentors may do more harm than good. Mentors' attendance predicted changes in mentees' social skills and self-esteem. Mentor attendance was a better predictor of mentee change than was mentee attendance, suggesting it was the experience, or length of time spent with the mentor that best accounted for changes in the areas of self-management, self-esteem, and social skills for the young person. The study showed that when mentors were inconsistent, it had a negative effect on their mentees (Karcher, 2005).

Variation of curriculum: A greater number of lesson hours across multiple studies was associated with more positive outcomes for interactive adolescent drug prevention programs (Tobler et al., 2000). The meta-analysis showed that programs found which were most effective in reducing substance use were those which combined strong behavioral life skills development content, emphasized team-building and interpersonal delivery methods, and emphasized introspective learning approaches focusing on self-reflection. The programs which were most effective were also based upon a clearly articulated and coherent program theory that provided intense contact with youth. Programs utilizing these positive program components produced consistent and lasting reductions in substance use 2000). Horn (2004), found a positive association between number of program hours and youth outcomes. In evaluating high school smokers, 2 central Appalachian states received the American Lung Association's 10-session Not on Tobacco program or a 15-minute brief self-help intervention. The study compared the efficacy of both programs by examining group differences in the smoking quite rates and found that participants in the 10-session program had high quit rates in comparison to peers who received the 15-minute brief intervention.

Variation of attendance or participation: Similar to degree of participation, level of contact has been associated with positive outcomes in mentoring relationships. A review of mentoring programs identified greater interaction as a characteristic of effective mentoring (Jekielek, Moore et al., 2002). For instance, adolescents whose mentors contacted them most frequently experienced more positive outcomes (e.g., better school attitudes, more positive reaction to drug use situations, less absence) compared with less frequently contacted adolescents. Mentees who spoke with their mentors rarely "did not experience benefits from program participation, and may even have experienced harm" (Jekielek, Moore et al., 2002). A study of almost 700 mentors found a significant correlation between spending more time with youth and feeling closeness and support in the relationship (Herrera et al., 2000). Study authors concluded that the "tipping point"

distinguishing the most close and supportive relationships from those less so was 10 or more hours per month of mentor-mentee interaction (Herrera et al., 2000).

There are several caveats to consider with regard to these findings. Attendance- and participation-based findings may be vulnerable to spurious relationships, and should therefore be understood with certain limitations. Given that attendance is necessarily voluntary, the participant who chooses to attend (or attend frequently) and the participant who chooses to skip sessions probably have different, largely unmeasurable individual characteristics with regard to motivation or responsibility, for example. These characteristics may make them more or less likely to experience positive outcomes, regardless of exposure to the program; unfortunately, these were not studied in the reviewed sources.

Another example of how individual characteristics of the participant may carry program effects is highlighted in the 2003 evaluation of the 21st Century Community Learning Centers program. In the 21st Century middle school study, very frequent participation (104 days over the school year) was significantly associated with certain measures of better in-school behavior and fewer school absences than median participation (44 days over the school year).[8] However, frequent participation was also significantly associated with indicators of inferior social skills. These findings, combined with the lack of an association between attendance and other positive outcomes, lead study authors to conclude that "the analysis does not suggest that higher levels of center attendance lead to improved outcomes" (Dynarski et al., 2003, p. 81). Particular to this study, the population of children who attended most frequently appeared to have several high-risk characteristics. The composition of this particular high-attendance group may have counteracted the usual pattern described above.

Summary of Dosage Findings

As is done for the other findings presented in this chapter, it is important to consider that dosage findings based on associations do not provide conclusive evidence of whether, and in what direction, the relationship is causal. For example, the association between greater amounts of interaction in mentoring relationships and positive relationships or outcomes may be misleading; it is unclear whether more time with the mentor benefits the youth, or whether a positive relationship between the youth and mentor leads to their spending more time together. Finally, it is important to note that a meta-analysis of early intervention programs shows that more intensive programs do not necessarily lead to better outcomes for children with disabilities (Innocenti & White, 1993), and a review of after-school child care programs finds that there may be cases (e.g., in regard to social competence) for which "moderate involvement" in activities is associated with more positive outcomes than extensive involvement (Vandell & Shumow, 1999).

[8] It should be noted that less than 10% of the study participants attended for more than 76 days over the school year; over half attended for less than 26 days over the school year.

Caveats regarding self-selection, in particular, confound the importance of attendance and hours of exposure; however, it has been associated with positive outcomes for many ages and contexts. Given this, we conclude that dosage appears to be a **positive-neutral** program component.

Duration

Duration—typically understood as the length of time during which a participant is enrolled in a program—has also been associated with positive program outcomes. This pattern is particularly strong in the context of mentoring relationships.

In a large study of 4-H programs, adolescents who participated for more than a year had more positive outcomes (e.g., "leadership, conflict resolution, communication, self-confidence, ability to make healthy choices, knowledge of nutrition and food safety, and record keeping") than youth who participated for less than a year (Rodriguez, Hirschl, Mead, & Goggin, 1999). A study of a substance abuse prevention program that controlled duration of exposure to program sessions found a similar pattern. Students in schools that were randomly assigned to a series of booster sessions subsequent to the previous year's full program experienced significantly lower cigarette smoking than students in schools that received the initial program but not the booster sessions (Botvin, Baker, Filazzola, & Botvin, 1990). Similar results were found in an earlier study of the same program that was implemented by regular classroom teachers (Botvin, Renick, & Baker, 1983). Finally, for young children in cognitive-enhancing programs, duration as well as continuity of participation in the program (i.e., not spending substantial time out of contact with the program) is associated with positive outcomes (Wasik & Karweit, 1994).

Duration appears to be particularly important for mentoring programs. A review of numerous mentoring studies concluded that the length of time a mentee and mentor had maintained their relationship was associated with significantly different outcomes; positive outcomes were found for those in relationships of 6 months or more, no outcomes were found for 3- to 6-month-long relationships, and negative outcomes (e.g., less academic confidence, lower self-worth) were found to be associated with relationships that lasted less than three months (Jekielek, Moore et al., 2002). There is also limited information suggesting that the positive effects associated with participation in mentoring relationships dissipate when the relationship ends (Jekielek, Moore et al., 2002). A different review of mentoring programs concluded that for trust to develop in the relationship, it generally required roughly six months of "regular meetings" (Sipe, 1996).

Summary of Duration Findings

Like dosage findings, those concerning duration are accompanied by several caveats. Length of continuation with a program is necessarily voluntary. The family or individual who chooses to stay in a program and the family or individual who drops out

probably have particular characteristics that make them more or less likely to experience positive outcomes, regardless of exposure to the program. For example, a family's inability to secure stable housing will most likely affect their ability to maintain participation in a program as well as affect various child outcomes. Also, as explained in other sections, another caveat to consider is that findings based on associations do not provide conclusive evidence on whether, and in what direction, the relationship is causal. For example, the association between length of mentoring relationship and positive outcomes may be misleading; it is unclear whether a longer relationship with the mentor benefits the youth, or whether factors that may make the youth more likely to experience positive outcomes (e.g., responsibility, respectfulness, etc.) facilitates the continuation of the relationship.

Overall, longer participation in a program appears to be a positive practice, though self-selection lends some doubt to its singular effect. We identify duration as a **positive-neutral** program component.

Intensity: Dosage in combination with duration

Four meta-analyses found that increased dosage and duration, together, are associated with positive outcomes. In the context of successful prevention programs for youth, the authors concluded that the most effective programs typically had a greater number of sessions and a longer duration of intervention (Eisen, Pallitto, Bradner, & Bolshun, 2000). In the context of juvenile delinquency programs, meta-analyses showed that as both duration and dosage increased, participants data showed larger effect sizes for reducing re-arrest and/or reconviction; the authors identified "substantial duration and intensity" as one of six common elements of program success (Lipsey, 1992). Similarly, a greater general level of participation was associated with more positive school-age through youth outcomes in a review of academic achievement programs (Redd, Cochran, Hair, & Moore, 2002).

The same caveats around self-selection and the unknown direction of causation, discussed separately in regard to dosage and duration, both apply to this section. Overall, the combination of dosage and duration appears to be a **positive-neutral** program component, as well.

Cultural Variation

More recently, studies have shown that cultural variances in program design have measurable effects on youth outcomes. Harvey (2004) examined the effects of an Africentric youth and family rites of passage program on at-risk African American youth and their parents. Data were collected from a multi-year evaluation using therapeutic interventions based on Africentric principles. The study demonstrated that the culturally competent programming such as the program's holistic, family-oriented, strengths-based approach and indigenous staff contributed to the gains exhibited gains in youth's self-esteem and accurate knowledge of the dangers of drug use. While, there were no

statistically significant differences, but parents showed improvements in parenting skills, racial identity, cultural awareness, and community involvement (Harvey, 2004).

Similarly, King et al. (2001) examined the effectiveness of an Afrocentric treatment program for male, juvenile, felon offenders in Cincinnati, Ohio. The evaluation used a two-group, quasi-experimental design to compare the 281 African American youths in the Afrocentric treatment program. The youths assigned to the Afrocentric treatment program performed slightly better than the probationers. Findings by Belgrave et al. (2004) support the effectiveness of a culturally enhanced drug prevention program for increasing drug refusal efficacy and resistance among African American girls. Belgrave and colleagues found that cultural variables such as ethnic identity and Afrocentric values are related to lower drug use and more negative attitudes toward drugs. Though evaluations of these programs have shown significant findings for cultural variables, but have been less successful in altering attitudes toward drugs and actual drug use (Belgrave, 2004).

Overall, cultural variances in program design with regard to affecting youth outcomes seem to have minimum effects on youth outcomes. Additionally, the studies examined review only Africentric programming. More evaluative research, both quantitative and qualitative, with varying racial and ethnic groups is needed to test the significance of culturally competent programs.

Clear and consistent message

Findings from the field of prevention/intervention programs appear to champion the practice of communicating clear and consistent messages about program goals and accepted behaviors. A review of 21 "successful" prevention programs finds that a common element of program success was "clearly delineated and articulated goals for behavior change" (Eisen et al., 2000). One commonality of effective pregnancy prevention programs from the last 20 years was that effective curricula "gave a clear message by continually reinforcing a clear stance on these behaviors" (Kirby, 1999). Fagan and Forst (1996) suggest that key components of successful programs include "clear rules and contingencies." Additionally, a large overview of crime prevention programs identified "clarifying and communicating norms about behavior through rules, reinforcement of positive behavior, and school-wide initiatives" as a school-based method that works to decrease crime, delinquency and substance use (Sherman et al., 1998).

Overall, communicating clear and consistent program messages appears to be a promising and positive program practice. Given that these associations have not been tested in relation to the effects of other program, participant, or caregiver factors, however, it is possible that the practice of providing a clear and consistent message is necessarily acting independently to improve outcomes. As such, this component can be considered a **positive to neutral** program practice.

Diversity of activities

Two studies suggest that the diversity of activities made available in programs affects child outcomes.

In a study of mostly Caucasian third- to fifth-graders in after-school programs in and around Madison, WI, "the presence of a greater number of different types of program activities was [significantly] associated with staff having more frequent positive or neutral interactions with children" and participant reports of more emotional support; also, it was marginally associated with participant reports of better overall climate (Rosenthal & Vandell, 1996). Conversely, a study of mostly Caucasian first graders in after-school programs in the same metropolitan area found that the presence of a large number of different activities had negative associations with positive behaviors in boys: "boys who attended programs offering a larger number of different activities had [significantly] more internalizing and externalizing problems, and poorer grades in reading and math," net of family structure (Pierce et al., 1999).

As there appears to be little information on the effects of diverse activities, and since the studies' findings are seemingly contradictory, it is unclear whether this component would be a positive, negative, or neutral program feature. It is possible that the different findings are due to children's different developmental needs at different ages. Also, it is possible that the effects of the "large number" of different activities in the Pierce et al. (1999) study were confounded with the effects of too many activities or an otherwise distracting or over-stimulating environment or selection into a program with many activities. Earlier in the paper, we concluded that the effects of structure and flexibility in conceptual approach differed by age; it is possible that too many activities at a young age compromise the degree of program structure from which younger children appear to benefit. As such, based on very limited research, **a greater diversity of activities appears to be a positive to neutral program feature for older elementary school-age participants,** while it appears to be **a negative to neutral feature for younger children**.

Interesting, engaging and enjoyable activities

It appears, unsurprisingly, that the provision of interesting and enjoyable activities may engage and retain program participants. In a study of almost 700 mentors, Herrera et al. (2000) found that "the extent to which youth and mentors engage in social activities [was] the strongest contributing factor for both community-based and school-based programs on all three measures of positive relationship quality: closeness, emotional and instrumental supportiveness;" academic and job activities, typically considered less entertaining than social activities, appeared to contribute less to the relationship. Frequency of participation among high-risk youth in a gang prevention program was significantly higher for those who reported that program activities were interesting (Arbreton & McClanahan, 2002). However, this pattern was not found for youth in the intervention program who were already involved in gangs. A study of Boys and Girls Club participation produced a similar finding, with Boston youths' attendance related

only to how interesting they found activities; however, there was no such association found at the study's New York site (Herrera & Arbreton, 2003). Additionally, recent research shows that service that exposes participants to compelling social causes (e.g., disadvantaged populations, issues of inequality or injustice), and empowers them to be agents in helping those in need, is associated with greater concern for social issues and intended civic activity than social work unrelated to these more compelling causes, such as tutoring (Metz, McLellan, & Youniss, 2003). These findings were significant net of various personality and background characteristics.

In sum, we find that providing engaging, enjoyable activities appears to be a **positive to neutral** program component, net of certain potentially influential characteristics, but not others (e.g., site location). This provides an interesting contrast for the importance of diverse activities, discussed above; it suggests that engaging activities may be more important than varied activities.

Family involvement

Family involvement has been viewed as particularly promising in achieving program goals. Communication and positive relationships with participants' families have been associated with positive outcomes in mentoring and after-school programs. For example, a 1999 review of after-school programs observed that "nearly everyone agrees that the most effective programs emphasize ongoing outreach to and communication with families," facilitating family engagement with the program, and understanding family needs and culture (Fashola, 1998). Similarly, a meta-evaluation of mentoring programs found that mentor's involvement with the mentee's family was associated with effective mentoring (i.e., "improves youth outcomes, or it results in successful mentor-mentee relationships"). In fact, "when students perceived that their mentor knew their parents well, these youth had better GPAs and higher levels of college attendance than non-participants" (Love, Schochet, & Meckstroth, 1996). According to a review of 12 pro-abstinence sex education programs, programs that involved parental participation had a small but significant positive association with program effect sizes for adolescents (Silva, 2002); programs that did not involve parent participation had a "substantially" smaller mean weighted effect size. Limited weight should be given to this finding, as only two of the twelve studies involved parental participation; however, both showed support for this component.

There are several caveats to consider with regard to findings on family participation. Parents that are able to and choose to participate in their child's program may be different from parents who do not participate in significant ways that affect their child's outcomes. As participation is necessarily voluntary, parents who choose to engage in programs may be more committed to involving themselves in their child's development. Alternately, parents who have an accommodating schedule and sufficient resources to spend time participating in their children's programs may have more time to spend with their families than parents who are unable to participate in programs. These and similar factors call into question the singular effect of family program participation

when compared to the effect of family characteristics that enable them to participate; unfortunately, these questions cannot be answered by the sources reviewed here.

Nevertheless, family participation appears to be a mostly **positive, though sometimes neutral**, program component. Parental participation appears to be particularly positive for older children and youth, especially in programs conducted outside of educational settings. Family participation did not appear to matter as much, and could be considered largely neutral, for younger children in child care settings.

4. Conceptual Approaches

Fidelity to program design ("strength of implementation")

Our discussion on conceptual approaches begins with addressing whether the program implements, with fidelity, the conceptual design on which it is based. Typically, thoughtful research and theory-based work goes into designing programs to fit particular populations, address certain needs, and produce certain outcomes. Indeed, program and curricular success have been associated with the program's basis in established theories (Eisen et al., 2000; Kirby, 1999). Adherence to the intended program design, therefore, is most likely in the best interests of the program. This thinking is intuitive and is corroborated by the research on a variety of program contexts.

In certain fields of evaluation, "issues of program implementation have recently emerged as some of the most important topics" to investigate (Catalano, Berglund, Ryann, Lonczak, & Hawkins, 1999). Unlike many other regulatable or manipulable features in this discussion, degree of implementation is virtually never manipulated in studies. Accordingly, the importance and role of implementation fidelity is typically discovered when unintended variations in this construct are found between program sites.[9] Fidelity to program design is often assessed by overall impressions of evaluation staff (Botvin et al., 1990) or by more concrete measures, such as the extent to which the program changes from baseline in specific "observational measures of program practices" (Battistich, Schaps, Watson, & Solomon, 1996).

There seems to be agreement—across program types and participant populations—that implementing a program with low fidelity is associated with unsuccessful and even negative outcomes, and that achieving a high "strength" of implementation is associated with positive outcomes. A study of correctional intervention programs for juvenile delinquents found that sites with the strongest fidelity to theory and program design had youth with the most positive outcomes—in this case, lower recidivism (Fagan & Forst, 1996). Furthermore, compared with their matched schools, only the high-implementation group experienced declines in carrying a weapon, vehicle theft, skipping school, or threatening someone with harm; some of these effects continued to the second year of data collection (Battistich et al., 1996). A study of

[9] Often, declines in the quality of program implementation (and, later, program quality) occur when model programs that had been rich in resources when evaluated and deemed successful are replicated "with inadequate budgets and less skilled personnel" (Shonkoff & Phillips, 2000, p. 19).

substance abuse prevention programs in 10 suburban New York junior high schools found that outcomes from teacher-led interventions were non-significant or negative until analyses were limited to teachers who "were judged to have implemented the intervention with fidelity and completeness;" in this high-fidelity group, negative effects disappeared and were replaced by positive effects for female participants (Botvin et al., 1990).

Adherence to program design appears to be important for smaller-scale services as well. In multisystemic therapy for violent and chronic juvenile offenders and their families, higher therapist adherence to program practices and principals was significantly associated with lower juvenile-reported delinquency and juvenile- and parent-reported re-arrest (Henggeler, Melton, Brondino, Scherer, & Hanley, 1997). Conclusions from meta-analyses concur with these study findings. Data from 196 school-based violence prevention programs suggest that "implementation quality made the largest contribution of any variable to effect size" (Mihalic, Fagan, Irwin, Ballard, & Elliott, 2002). Similar observations were made in regard to drug prevention programs (Mihalic et al., 2002).

There are certain caveats to consider in regard to these findings, particularly because these studies did not (and, in many cases, could not) statistically control other potentially influential factors. In particular, fidelity to program design and positive program outcomes may have a spurious relationship. Strong implementation may be less likely to occur at sites without sufficient resources or staff training (Hallfors & Godette, 2002), for example, which are characteristics that would most likely compromise program outcomes, independent of the degree of implementation. Furthermore, it is speculated that high implementation fidelity and the presence of program researchers are often conflated (Lipsey, 1992); therefore, it is possible that strength of implementation is a marker of the presence of evaluation or research staff, which could be as much or more responsible for positive outcomes. Nonetheless, there is extensive non-experimental information that links strength of implementation to positive outcomes; we conclude it appears to be a **positive to neutral** program component.

Participant-centered approaches

Programs are necessarily oriented toward their participants: eliciting changes in their behavior or well-being is the crux of virtually every program's design. However, accommodating participants is not the same as making program practices and processes "participant-centered." As discussed below, processes and interactions that distinguish themselves by addressing participant needs initially and as they evolve and—perhaps most importantly—not prioritizing staff or mentor needs over those of participants, are linked with positive program outcomes.

Studies of mentoring programs have consistently found that positive outcomes are linked to youth-centered mentor-mentee relationships. For example, a study of 669 mentors found that these adults felt their relationships were closer and more supportive when the pair made decisions together (Herrera et al., 2000). This association of positive relationships and mentor-mentee teamwork was corroborated by a review of mentoring programs (Sipe, 1996).

Conversely, relationships in which the mentor largely determines the development and "agenda" of the relationship have been linked with less positive relationships and outcomes. When adults enter the mentoring relationship with goals to reform the youth or with expectations of achieving particular goals, the matches are described as "prescriptive;" this is in contrast with entering the relationship with only the goal of developing a relationship with the youth. An overview of Big Brother/Big Sister evaluations found that while "the majority of prescriptive matches faltered or closed (22 of 28), most developmental matches (50 of 54) persisted and continued to develop" (McGill et al., 2001). Herrera et al.'s (2000) study found that "the least positive relationships resulted when decisions about activities were made primarily by the mentor or established in advance by the program."

One caveat to consider is that adults who go into the mentoring relationship with the primary goal of developing a friendship may differ in significant ways from adults who enter the relationship with prescriptive agendas. Nevertheless, the youth-centered approach appears to be a **positive though possibly neutral**, practice for mentoring programs.

Youth Development Approach

These approaches speak to positive youth development theory, which is based on the perspective that "problem-free is not fully prepared" (Quinn, 1999). In other words, working only to prevent problems in participants may not be as effective an approach to achieving this goal as would also addressing antecedents of, and the lifestyles associated with, problem behaviors. Addressing the whole person requires a "more complete and balanced approach to… development," taking into account the various factors—such as peer pressure, academic difficulty, or family turbulence—that contribute to or detract from the individual's well-being (Moore & Zaff, 2002, p. 3). Similarly, programs with a multi-component approach are designed to affect change in the individual by addressing—and possibly changing certain elements of—the various domains in which he or she exists (e.g., school, neighborhood, family). These two approaches are designed to address the varied needs of participants and the various contexts in which the participants live. Programs that speak to the experiences and environments of the whole child, adolescent, or youth appear to be associated with significantly more positive outcomes than do less comprehensive programs. In his review of sexual health programs, Kirby (2002) touches on the logic of such approaches: "It seems likely that programs that address both sexual and nonsexual antecedents effectively will be more effective than those programs that address only one group of antecedents."

A meta-analysis of 207 school-based drug prevention programs by Tobler (2000) suggests that program effects increase with the comprehensiveness of approach. The most effective program types—all described as interactive and psychosocial in nature and designed to increase participants' interpersonal skills—were system-wide change programs (which incorporate the participants' school, community, family, and media environments), comprehensive life skills (e.g., refusal skills, goal setting, communication

skills) programs, and social influence programs. Similarly, one of six common elements of program success found in a review of higher-quality prevention program studies was the "multiple-component interventions," which largely involved the adolescent participants' community, parents, and/or "same-age or older-age peer leaders" (Eisen et al., 2000).

In promoting positive youth development and risky behavior prevention, the "whole person" approach is necessarily linked to the youth's social contexts and his or her ability to navigate them; social skill-building is nearly always a key feature of such programs. Authors of a review of successful prevention programs identified the following as one of six common elements found among the most successful programs: "interactive student-to-student and student-to-instructor skill-building methods---including role-playing and rehearsal, guided practice, and immediate feedback--- to address the target problem behavior" (Eisen et al., 2000). Communication issues, behavior modeling, social influences, assertiveness skills, problem-solving, and decision-making were also cited as components of the successful approaches. Kirby's (1999) review of pregnancy prevention programs from the last 20 years concludes that effective curricula "included activities that address social pressures on sexual behaviors," and "provided modeling and practice of communication, negotiation, and refusal skills."

Institutions that are able to execute multi-component programs may have certain characteristics (e.g., high-quality staff, greater financial resources, etc.) that affect positive program and participant outcomes, regardless of approach; unfortunately, many of these factors were unable to be controlled in these meta-analyses and reviews. Also, not every child or youth has deep or pervasive needs requiring a comprehensive intervention (Moore, Vandivere, & Redd, 2003). Nonetheless, multi-component, "whole person" approaches and social skill-enhancing practices appear to be positive **to neutral** program practices.

Shock approaches

Shock-based programs are based on the logic that exposure to "the horrors and difficulties of life in prison" through visits to or short-term imprisonment in maximum security prisons will motivate offenders to abandon criminal activity and other negative behaviors (MacKenzie, 1998, sect. 3.2). The "Scared Straight" program is based on this rationale, and targets juvenile offenders in particular. We have found no studies that compare shock approaches with non-shock approaches in any methodologically sound way. However, a meta-analysis and a review of juvenile delinquency and crime prevention programs speak to typical outcomes associated with the shock approach.

The meta-analysis of 443 juvenile delinquency programs suggested that "scared straight" and "shock" treatments can actually lead to increases in participant re-arrest and conviction (Lipsey, 1992). The studies included in the meta-analysis were experimental or quasi-experimental and conducted between 1950 and 1987. A large overview of crime prevention programs, commissioned by the Department of Justice, reveals the same

pattern. The review shows that the prison visit-oriented Scared Straight program does not deter and may, in fact, increase criminal activity (MacKenzie, 1998). Similarly, shock parole and probation, in which adult offenders are temporarily incarcerated, are linked with null or negative outcomes as well.

From these reviews, "shock" programs appear to have a negative effect, if any; we conclude that this is a **negative to neutral** program approach.

D. Chapter Summary

Non-experimental research is plentiful, but not as useful as experimentally-driven research because findings are necessarily accompanied by numerous caveats. Given that non-experimental findings are generated from mostly correlational analyses, causal relationships cannot be established. Furthermore, reviews and meta-analyses provide a rich overview of extant research, but may be victim to the "file drawer" effect. With these caveats, however, non-experimental studies can contribute to what we know about program implementation.

1. The ideal program, according to non-experimental research

Based on non-experimental research studies, one could describe a rather complete ideal program. Staff for this program would have higher levels of formal education, as well as training for the particular program. A lower ratio of participants to staff in child care programs, higher staff wages, lower group sizes, and positive provider-participant relationships would be part of an ideal program as well. It would not be necessary to match staff and participants by demographics, especially if they had similar interests to link them. Any adult volunteers in the program, particularly mentors, would undergo comprehensive screening, training, and going supervision.

A variety of different approaches would merit use in the ideal program. The approach could be participant-centered, regularly changing in order to best address participant needs, or it could be multi-component and "whole person," addressing the various contexts in which the participants live, and promoting social and life skills or both. However, lessons and skills-trainings would be interactive in nature, not didactic or information-only. Also, activities would be designed to be engaging and interesting to the participant group. The program would take steps to communicate a clear and consistent message, and would implement a theory-based program design with fidelity to the program design. "Shock" or "Scared Straight" approaches would not be employed.

For younger children, the ideal program would feature a relatively high degree of structure in its day-to-day activities; for older children and youth, however, higher levels of program flexibility would be beneficial. Parental involvement in the program would be an important component for school-age children. Overall, the program would be of high quality—particularly for child care centers, in which quality is typically measured in established environmental scales—and participants would experience it at higher levels of dosage and duration.

2. Future research

The field has research on many other regulatable and manipulable features that may be promising; however, these "lone findings" were not included in this discussion because there were no other studies with which to affirm or reject them. These components, on which we encourage more research to be done, include single-sex facilities (McGill et al., 2001), style of caregiver speech (McCartney, 1984), the effect of distractions near the program site (McCartney, 1984), the effects of new participants on existing participants (Arbreton & McClanahan, 2002), and the importance of support services (St. Pierre & Layzer, 1996) and of home-based visiting (St. Pierre, Gamse, Alamprese, Rimdzius, & Tao, 1998; St. Pierre et al., 1995).

Other features, such as site safety (Arbreton & McClanahan, 2002) and program strength and stability (Halpern, Spielberger, & Robb, 2001), are the subject of little empirical research and most likely affect participant and program outcomes; however, they are inappropriate to be manipulated in experimental studies. In these cases where research cannot inform the field, wisdom from program providers is particularly illuminating.

IV. WISDOM FROM THE FIELD

By Krystal McKinney and Christina Theokas

A. Introduction

This chapter presents information from "the field," including individuals who work directly with children and youth in out-of-school time programs, intermediaries that fund programs and train youth workers, and researchers who collaborate with practitioners to understand their view on what constitutes program quality and effective program implementation. It differs from the prior two chapters that summarize independent research studies that have been conducted within actual programs for children and youth and instead focuses on the knowledge of individuals and groups that is gained from experience. As such, this information complements and extends what is learned from rigorous experimental and quasi-experiment studies in that it is gained from hands-on experience and can provide information about program components that cannot be manipulated in traditional research studies (e.g., providing a safe environment).

It should be noted that the information in this chapter is not solely gained from front line program delivery staff. Often the voices of practitioners are linked with youth development researchers as the field has grown as a collaborative effort. The ideas in this chapter therefore incorporate elements of practice and scholarship, in fact, at times it is hard to separate the two. This chapter presents an opportunity to document the ideas that are arising from and influencing the direction of the field and hopefully will lead to more rigorous research and evaluation studies when appropriate and feasible.

B. Methodological Approach

1. Information Collection

To obtain information for this chapter, the websites of key national youth development organizations, national, state and local after school partnerships and intermediaries, and national youth serving organizations were reviewed to find information about program implementation, essential program components and guidelines for practitioners to follow when delivering programs (see Appendix E for list of organizations and websites). These sites were chosen as they serve the out-of-school time field, guide their work, and/or advocate on their behalf. They represent clearinghouses of information that are publicly documented and available. These organizations speak for groups of individuals, across different program types, so they tend to provide thematic summaries of information, as opposed to the more find grained analyses in other chapters. At times, information is gleaned from individuals, who are experts in the field, and these ideas reflect unique values, experiences and the work view of that person.

A second source of information for this chapter is research conducted with youth workers. Literature searches were conducted to locate research that utilized youth

workers and out-of-school time practitioners as the primary respondents reporting, ranking or sharing their ideas about program implementation. In addition, journals that cater to and include articles by practitioners were individually examined and searched for reports about ideal program implementation practices. Also, the opinions of youth workers based on their personal experience is also incorporated, to provide insight derived from this on-the-ground experience. Furthermore, the thoughts and ideas of youth who participated in the Youth Roundtable held by Child Trends in 2006 are incorporated.

2. Identifying "positive" and "negative" program components

This chapter reviews program components from the perspective of what works to enhance participants' development and well-being, and what is important to programs and service delivery; all of the components identified in the discussion that follows are considered "positive" for this reason. When endorsed by multiple sources and dissent is not identified, this represents consensus in the field on the importance of the program element. Components that represent the opposite of any given positive component would be understood as "negative." For example, if staff retention is highlighted, the opposite, high staff turnover, is considered negative. A component is identified as neutral if it is generally considered to make no significant impact—positive or negative—in the program or participants.

C. Summary of Findings

The following sections represent the general consensus of the field on the essential elements, "best practices," or key ingredients for success for out-of-school time programs. Ideas are grouped into major headings. These groupings represent all facets of running a program including infrastructure needs, staffing concerns, activities, and partnerships. The sources reviewed for this chapter tend to generate lists of standards for program quality as opposed to focusing on program implementation specifics. This likely reflects the reality that they are speaking across program types with varying financial resources and community dynamics that must be considered. For example, 4-H is a national youth serving organization that runs programs in all counties in the country. Instead of providing detailed implementation guidelines, 4-H specifies essential elements of all programs across diverse content, activities, and practices, including creating a sense of belonging, mastery, independence and a spirit of generosity (Kress, 2004). This approach does not address fine details (e.g., must meet 3 times per week for an hour for a minimum of 12 weeks) but attempts to focus on being developmentally appropriate and context and culturally relevant. A youth development orientation underlies many of these standards (as opposed to a prevention science approach, which more commonly underlies the specific experimental research discussed in the first chapter). Nonetheless, across these many sources, there is reasonable agreement on key ingredients for success; for example, interesting activities and supportive relationships (Granger & Kane, 2004).

1. Vision and Program Goals

Vision is not commonly discussed in relation to quality out-of-school-time programs. It is more typically associated with business models of success; however, in our review, the term "vision," "mission-driven," or "focus" did arise a number of times in relation to quality standards, making it worthy of discussion. In addition, although vision may not have been mentioned directly, associated ideas were frequently cited with regard to program operations including having a clear focus and goals or having a philosophy (e.g., youth development) guide program activities.

A vision statement articulates the inspiration for the program. It describes where a program wants to go, but not how it is going to get there. The underlying assumption regarding why vision is important to program quality is that this inspiration, or passion, guides strategic planning and keeps the organization focused on improving itself in the space it operates in. Having a clear program focus is vital to a knowledge-centered organization, according to McLaughlin (2000). Based on an examination of effective organizations in California programs that focused on something in particular and had purpose attracted youth to get involved and stay committed. During site observations, research staff noted that, although programs on the surface seemed casual and relaxed with informal relationships between staff and participants, they were deliberate in both the content of activities and the type of environment created. The successful organizations were confident and consistent about offering something of value to participants and responded to participants' strengths, interests and needs.

Similarly, the Finance Project (2005) interviewed 18 program providers and intermediaries to determine what separates successful, sustainable youth development programs from those that struggle to survive. A vision and mission-driven focus were deemed highly important and included in the list of seven essential characteristics. The respondents felt that having a vision helps programs to have a clear understanding of what they do, the niche they have, and what should be sustained. Without this focus, mission slip is possible (SRA SIG, March 2006). In this era of restricted funds and competition, programs can easily be convinced to alter their mission to attract funds. However, this slippery slope away from core elements and values of the program can lead to confusion, turnover, and inconsistency in programming that turns participants away. After interviewing staff at successful after-school programs in Massachusetts, Hughes (2006) also articulates a vision and mission for quality as a prerequisite for a successful program. This sets the stage for hiring competent staff and implementing curriculum in line with program goals. Other summaries of research have also established a clear mission and goal setting as integral to program quality (Bodilly & Beckett, 2005; National Research Council & Institute of Medicine, 2000; U.S. Department of Education and U.S. Department of Justice, 2000).

Related to vision, is the idea of a strong philosophy to guide program actions. After conducting an extensive literature review and having discussions with experts and innovative program leaders, The American Youth Policy Forum (2006) arrived at the conclusion that the foundation of any out-of-school time program must be youth

development principles including being strength based, youth-focused, and developmentally appropriate. The DC Children and Youth Investment Trust Corporation, which is an intermediary funding organization in the District of Columbia, takes this program quality standard seriously. They offer a 32-hour training to their grantees in youth development principles. This is not simply a service that is available to grantees; at least one staff member from each of their grantees is required to participate yearly. Being youth-centered is a paradigm shift for many out-of-school time programs. Historically, out-of-school time programs were established to keep kids safe and occupied during the hours after school (Halpern, 2002). Children and youth were viewed as problems waiting to happen, so programming often focused on fixing and teaching. The new generation of programs that have been found to be successful (i.e., that achieve positive outcomes) put youth at the center of programming and understand youth needs and respond to their diverse talents, skills, and interests (McLaughlin, 2000). Activities, environments, and content can vary, but at the heart of the organization is a philosophy that values youth, listens to youth, and aims to build their strengths.

Building on the idea of a strong program philosophy, programs must provide specific and detailed program goals and outcomes they wish to achieve. Providers agree that it is crucial to set clear goals that have measurable outcomes and understand what it will take to reach these outcomes. "The use of outcome indicators that reflect common sense and common understanding may be the most powerful force to focus attention on agency mission rather than rules, and to permit the necessary flexibility and autonomy at the front-end" (Schorr, 1992). Having clear goals and outcomes helps to improve the program effectiveness and accountability as well as identify where changes need to be made.

When identifying outcomes, general opinions acknowledge the need to identify outcomes not only for the program and the agency, but for program participants as well. Informing participants of the desired program outcomes will provide then with a sense of the program's purpose. Many providers agree that sharing outcomes in a manner that reflect what you don't want participants to do may lead to resistance from parents and well as from participants. For example, a program that explicitly states to its parents and participants that its program goal is to prevent substance use among teens may prohibit some from becoming involved. Parents may feel their child does not need to be involved because they believe their child would never use drugs. If the outcome is formed to state what the program wants the child to do, such as make healthy decisions, people tend to be more responsive and even recognize how the outcome applies to their personal development. When developing outcomes for program participants, "we need to be intentional about expecting and measuring what we want them to do, not just want we do not want them to do" (Pittman, Irby, Tolman, Yohalem, & Ferber, 2001).

Vision, focus, and goals are not enough though to sustain a successful program. Instead, this orientation forms the foundation from which to build. A necessary ingredient to be added is staff.

2. Staff

Staff are considered *the* key resource in out-of-school time programs by most organizations, intermediaries and even by youth themselves (e.g., United States Department of Education and United States Department of Justice, 2000). For example, when leaders in the field were asked to identify the single most important ingredient for program quality, the overwhelming response was program staff, indicating that they need to be well trained, well compensated and able to foster youth leadership (Harvard Family Research Project, Spring 2004). The ability of staff to develop positive relationships, create the desired experiences, maintain safety, communicate effectively, and engage youth in the process is integral to the success of any program. Children and youth "vote with their feet," so if they do not feel welcome, or enjoy the staff, they will go elsewhere.

Staff are described in varied ways in relation to best practices. With regard to the individual, characteristics include personal qualities (e.g., optimism, creativity), education (e.g., child development background, attainment level), and essential competencies (e.g., cultural competence, artistic, etc.). For the organization, staff development and training, staff-to-participant ratios, compensation and retention are considered in relation to program management and operations.

Personal and professional qualities

Working with children and adolescents in structured out-of-school time programs requires an excitement and enthusiasm for children and what they can accomplish. School is mandatory and has known elements, for example, one teacher for a group of children and specific content to be covered and mastered. However, out-of-school time programs are voluntary, and the content and process of engagement can vary across programs. Some adults shy away from young people as they transition to adolescence because of their changing attitudes and behavior. Adults may feel that they are not needed as much, as young people turn to their friends to share confidences and ask for advice. However, program providers see beyond these surface behaviors and recognize that adult support and guidance are necessary. Bob Granger, President of the William T. Grant Foundation, suggests that line staff should be hired based on their personal characteristics and social skills that allow them to connect with youth (Harvard Family Research Project, 2004). Having a positive attitude, charisma and belief that all children can succeed inspires trust and liking by youth that is essential to programs. Children and adolescents recognize when a staff member is motivated and committed to them and the program, which increases their enjoyment and participation.

The majority of sources reviewed for this chapter identified youth worker competencies. The National Collaboration for Youth (www.nydic.com) specifies characteristics including the following personal attributes:

- Respects and honors cultural and human diversity
- Demonstrates the attributes and qualities of a positive role model
- Interacts and relates to youth in ways that support asset building.

Other core competencies include:

- Understands and applies basic child and adolescent development principles
- Facilitates and evaluates age appropriate activities
- Involves and empowers youth
- Communicates and develops positive relationships.

Staff who are able to form supportive relationships so children and teens can experience emotional guidance and practical support create an environment that participants want to be a part of and come back year after year. More so than any other skill, the ability to have positive interactions between staff, participants, parents and communities is most frequent characteristic mentioned on lists of program quality standards (e.g., the After School Corporation, New YorkDepartment of Youth Community Development, National After School Association).

One core component of positive relationships is being able to relate to youth in developmentally appropriate ways. Familiarity with children's changing developmental needs allows program staff to know what practices, program design, subject matter, teaching styles, and variations on many other components are appropriate for the children, teens or youth in the program. Children of different ages need different things out of programs. School age children often benefit from consistency and structure, whereas older teens benefit from options and autonomy to choose among them (Forum for Youth Investment, 2003).

Staff Education and Experience

There are few key mandates from the practice community on minimum education and qualifications. As mentioned above, there are ideal competencies, such as having a child or adolescent development background. However, this is not typically operationalized as a minimum degree requirement, major in college, or number of years of work experience, as is typical, for example, with teachers or social workers who also work with children and teens. Programs struggle to attract and retain staff to programs due to the typical hours of operation, often part-time work, and traditionally low compensation. This often leads to programs drawing staff from the local community, finding mentors at local colleges or universities, and finding individuals who have other part-time jobs, but can work after-school or on weekends. This makes it difficult to find staff that are able and willing to make long term commitments to programs that research has shown to be particularly important (Hall, Israel & Shortt, 2004), especially for mentors (Tierney, Grossman & Resch, 1995).

Staff Training and Professional Development

Closely aligned with the issues of ideal competencies, education, and personal qualities is the staff training and professional development offered by the organization to prepare youth workers for the unique demands of the program. As opposed to focusing on the individual, this best practice focuses on the organizations and their management and administration. Most organizations recognize that they need to provide some training

to cultivate and retain competent and committed staff. For example, The After-School Corporation (TASC; www.tascorp.org) identifies staffing and professional development as one of the ten essential categories of quality after-school programming. They suggest that professional development should focus on training staff to value participants, understand their developmental needs and to work closely with families and school partners. Programs may not be able to hire college graduates, but they can provide training in program philosophy, interactions, and activities that will ultimately benefit the program and participants. Moreover, if programs do not have the resources to conduct trainings themselves, a number of intermediary organizations have been developed for the express purpose of providing training and technical assistance to program providers. For example, the Partnership for After School Education (www.pasesetter.com) offers workshops, institutes, best practice panels and annual conferences, free of charge, to New York City Out-of-School time grantees as part of a large initiative to ensure quality programming fro children and youth and building program capacity.

The National After School Association (NAA) has developed the most comprehensive set of quality guidelines (NAA, 1998). They are intended to describe best practices in out-of-school time programming for children and youth 5-14 years old and can act as a self-study guide to program improvements and accreditation, if desired. There are six keys of quality:

- Human relationships
- Indoor environment
- Outdoor environment
- Activities
- Safety, health and nutrition
- Administration.

The administration section describes program organization, procedures and policies. Within this category, a number of individual keys focus on staff training and development. First, the NAA standards specify that orientation to the job before beginning work should be provided for new staff and training is needed that is relevant to the responsibilities of each job. A specific number of annual hours of training per job type is also recommended; Assistant Group Leaders should receive at least 15 hours of training, Group Leaders at least 18 hours, Senior Group Leaders at least 21 hours, Site Directors at least 24 hours, and Program Administrators at least 30 hours. Training is supposed to be ongoing and cumulative. In addition, they suggest that training should focus on 4 areas:

1. Staff receive training in how to work with families and how to relate to children in ways that promote their development.
2. Program directors and administrators receive training in program management and staff supervision.
3. Staff receive training in how to set up program space and design activities to support program goals.
4. Staff receive training in how to promote the safety, health, and nutrition of children

Many local intermediary organizations also include staff training as an essential quality element. Some have adopted and adapted the NAA Standards for Quality School-Age-Care (e.g., DC Children and Youth Investment Trust Corporation; Baltimore Safe and Sound Campaign, Philadelphia; and Youthnet in Kansas). Those that have not adopted the NAA guidelines also generally include qualified staff and training on their lists of elements of quality after school programming. For example, the Stark Education Partnership in Ohio surveyed parents and providers. School personnel conducted focus groups to learn about key program quality elements. There was agreement across all the groups that staffing quality was essential. Boston's After School for All Partnership include well-trained staff in their position statement on quality out-of-school time learning, as does the Providence After School Alliance.

Recent summaries of the literature and work conducted with out-of-school time programs and interviews with program leaders have also pinpointed staff training as essential (Beckett, Hawken, and Jacknowitz, 2001; Pechman and Fiester, 2002; Raley, Grossman, and Walker, 2005). For example, Raley and colleagues (2005) provide the following insights about how to develop a strong, stable staffing structure from lessons learned from over 100 out-of-school time programs: hire the right staff, align staff skills with tasks, make training substantive and accessible, provide day-to-day staff development, and monitor activity quality. Their summary highlights that training should be daily and ongoing and that one way to assess the success of training is to monitor activities. This recommendation was also made by Pechman and Fiester (2002). They suggest continuous staff development and using peer modeling and team work to develop staff capabilities, as opposed to just investing in a one-shot training curriculum.

Compensation and Retention

Compensation, similar to educational standards for out-of-school time staff, is not frequently included on lists of quality standards or best practices in program implementation, although the common wisdom is that poor compensation and turnover are related and are barriers to the growth and stability of the workforce and field (Tolman et al., 2002). Advocates for the out-of-school time field recognize that youth work tends to be characterized as low status, which correlates with the low compensation. Professionalizing the field, establishing credentials, and creating opportunities for public recognition are suggested as one way to change perceptions and are being discussed in numerous forums, including offering undergraduate and graduate degrees in youth and community development, as well as certificates from training programs (Walker, 2003). Similarly, providing attractive compensation could improve recruitment and retention (Beckett, Hawken, and Jacknowitz, 2001; CS Mott Foundation, 2005). However, most programs operate on very tight budgets and are always seeking new funding to maintain and grow services, which limits their ability to guarantee salaries. Often, programs rely on volunteers to provide program services. Interestingly, many of the sources reviewed for this chapter suggest the use of volunteers to help address the lack of resources and increase the financial stability of programs (Beckett, Hawken, and Jacknowitz, 2001).

Once a program has managed to obtain qualified staff, it needs to identify and incorporate effective ways to keep staff on board. Most providers acknowledge that they have felt the effects and disappointment of losing qualified staff after a short time of employment. This can definitely cause some disturbance to a program. Staff turnover negatively affects program participants and their involvement. The more that children and youth lose people whom they trust and respect, the harder it is for them to develop attachments to the next person, due to fear of another loss. Low staff turnover allows stable relationships to develop between staff and program participants. "Training is critical to retention of quality staff members and volunteers" (U.S. Department of Education & U.S. Department of Justice, 2000, p. 39).

What role can programs play in maintaining staff? Key advice from program directors is to recognize that working with young people is not an easy task, regardless of how qualified a person is and how much this person loves the population. It can be very emotionally draining. Providers agree that allowing time for staff to discuss their own concerns and receive feedback is critical to their own development as well as retention. It is critical that directors provide emotional support and guidance, as well (The Forum for Youth Investment, 2003); this allows staff to feel that they are understood, valued and respected for their work. Just as important is being in tune to staff morale and having strategies for boosting morale.

Staff-Participant Ratios

A question frequently raised by programs is: how many adult leaders are needed to maintain safety and adequately implement the curriculum, program, or intervention? With child care for younger children, guidelines have been set. For children and youth, the guidelines are less clear. The After School Corporation suggests a 1:10 ratio. The US Department of Education and the United States Department of Justice (2000) suggest that, for school age children and older adolescents, a ratio between 1:10 and 1:15 is adequate. Other sources just say that maintaining a low child-to-staff ratio is a model after-school practice as it increases the likelihood that each child will have one-on-one time with an adult and be able to develop a meaningful personal relationship (Beckett, Hawken, and Jacknowitz, 20010; Caplan and Calfee, 1998). Minimum standards are necessary for safety and quality; however programs differ in their goals and activities, which will affect the desired ratio. For example, a teen drop in center may have fewer staff, because the environment is supposed to be a safe place for youth to congregate and talk with one another, whereas a youth advocacy training program may have one or two adults for a small group of six or seven youth as the focus is on transmitting skills and collaboration.

3. Program Practices

Activities

Similar to staff, activities are at the core of any program. They can range from sports, to arts, to civic engagement, to recreation and leisure or all of the above. A range

of high-interest activities is often thought of as ideal, as it can appeal to more potential participants (American Youth Policy Forum, 2006; Raley, Frossman, and Walker, 2005). In addition, a large number of programs are being, or have been developed, to support academic learning and enrichment, and are seen as extensions to the school day including tutoring, instruction, and homework help (e.g., 21st Century Community Learning Centers). According to The After-School Corporation, if programs have an academic focus, it is important to align activities with school standards and curriculum (see also, Beckett, Hawken, and Jacknowitz, 2001; National Partnership for Quality After School Learning www. Sedl.org/afterschool). To accomplish this integration, it is suggested that after-school staff work with school staff to develop programming that builds upon what is covered during the school day (e.g., CS Mott Foundation, 2005). However, after-school programs can be different and many suggest that learning activities should be project based, youth-led, with interactive teaching methods to hold the attention of participants (American Youth Policy Forum, 2006). Programs for older teens and youth, as opposed to those that focus on academic instruction, are instead often focused on exposure to career possibilities, marketable skills, apprenticeships, and job training (e.g., After School Matters in Chicago). For these programs, rather than collaborating directly with school staff, practitioners suggest building alliances with community and business leaders (Hall, Israel, and Shortt, 2004).

Regardless of the content of the activity, practitioners clearly articulate that activities should be engaging and build on youth needs and interests (Bodily, and Beckett, 2005; Carnegie Corporation, 1992; National Research Council, 2002). One suggestion from the National Partnership for Quality Afterschool Learning (www.sedl.org/afterschool) is to survey participants about what they would like to do in an out-of-school time program and to build programming/curricular around those interests. In addition, the majority of lists of program quality standards include the idea that activities should emphasize skill building and provide children and teens with challenging and interesting learning experiences, whether academic or not (Baca, Walters, Orloff, and Salerno, 2004; Boston's After School for All Partnership; Providence After School Alliance). Also, many of the national youth serving organizations emphasize in their mission statements or essential program elements that curricula must be engaging and meaningful. For example, core beliefs for 4-H are leadership, citizenship, and life skills through hands-on experiential learning. Similarly, Girl Scouts emphasizes non-formal, experiential, and cooperative learning as a core program element.

Several ways that activities can be delivered are through small group discussions, arts and crafts, educational videos, sports, and guest speakers. For younger children, in particular, "quality programs give children the opportunity to follow their own interests and curiosity, explore other cultures, develop hobbies, and learn in different ways, such as through sight, sound, or movement" (U.S. Department of Education & U.S. Department of Justice, 2000). For older children, teens, and youth, utilizing teaching methods that offer opportunities to model and practice skills, such as role-plays, is a practice that is likely to help participants remember the information being taught. "Young people, just like adults, learn best through active participation and learning

occurs in all types of settings and situations" (Newman, Smith, & Murphy, 2001). Effective role-plays incorporate ample time for youth to thoughtfully process what was learned. For example, a discussion on the effects of various decisions reinforces the potential consequences, helps children learn to generate multiple possible scenarios, and clarifies the importance of prudent decision-making. As children get older, many program providers agree on the importance of incorporating activities that encourage the participants to personalize the information being taught whereby they can truly relate to the issues.

The After School Alliance offers a specific list of the types of activities that after school programs should offer to youth by age group (5-10, 10-14, and 14-18). For example, they suggest that for 5-10 year olds a wide variety of activities and choices should be offered, but there should be a clear routine and structure; 10-14 year olds need connections to real world experiences; and 14-18 year olds enjoy presentations and projects that involve current trends (see www.afterschoolalliance.com for a full list). The National After School Association (1998) offers a slightly different list and focuses on how activities are implemented rather than the types of activities offered. Programs are supposed to assess:

- If the daily schedule is flexible, and offers enough security, independence, and stimulation to meet the need of all children and youth;
- If there are a wide variety of activities to choose from for participants;
- If the activities reflect the mission of the program and promote the development of all children and youth in the program; and
- If there are sufficient materials to support program activities.

Engaging Participants

Many providers have found that an effective strategy for engaging young people is to develop programs that are based on youth interests and input and allow opportunities for program participants to belong. One way to accomplish this is to involve young people in planning program activities and to encourage them to offer feedback on the program structure and activities. When young people are involved in this process and their views are heard, experience suggests that they are more likely to take ownership and feel a sense of belonging. However, as avenues open for young people to voice their opinions and criticisms, all responses will not be positive. It is important that program providers respect all comments and not take them personally.

Young people will learn and enhance their own skills in areas such as communicating effectively, respecting others' opinions and building self-confidence when given opportunities to share. Program providers will ideally use all opportunities, including those accompanied by criticism, to enhance these skills. An environment where young people feel cared for and respected enhances engagement and retention as well as youth's ability to learn the skills and information the program offers.

For example, After School Matters in Chicago (www.afterschoolmatters.org) works with adolescents and has drawn from the research of Hall, Istrael, and Shortt (2004) on out-of-

school time needs for urban teens, which emphasizes incorporating youth voices into decision making and offering employable skills and connections to job training and employment. Core program elements of the After School Matters program are thus focused on providing youth with real-world, relevant, challenging, and meaningful apprenticeship experiences to help prepare them for the job market as adults. This program seeks out adult mentors who are both caring as well as experts in their field. Instructors are coached on listening to teens, treating them with respect and also being able to provide personalized instruction in relevant job skills and having high expectations for performance.

Interactions among Participants, Staff, and Families

Staff and activities are the main ingredients of daily program operations. However, the process by which they come together with participants is through the actions and interactions with staff. Human relationships are the first major category of the National After School Association's list of quality standards (1998) that has been adopted by many organizations. The five best practices associated with human relationships are:

1. Staff relate to all children and youth in positive ways.
2. Staff respond appropriately to the individual needs of children and youth.
3. Staff encourage children and youth to make choices and to become more responsible.
4. Staff interact with children and youth to help them learn.
5. Staff use positive techniques to guide the behavior of children and youth.

Each best practice is associated with a set of guiding questions and specific objectives to help programs evaluate and rate themselves. For example, relating in positive ways to participants is a core practice. Some of the associated questions for this best practice are: How do we greet children? Are they happy to be with us? Some of the objectives include: Staff treat children with respect and listen to what they say. Staff make children feel welcome and comfortable. Staff respond to children with acceptance and appreciation. Staff are engaged with children.

Positive, promotion-oriented interactions seem to form the core of most best practice standards lists (International Youth Foundation, 1999; National AfterSchool Association, 1998; The After School Corporation). The National Research Council and Institute of Medicine (2002) report has provided a seminal list of eight features of positive developmental settings. The first seven all describe the nature of relationships and types of experiences that should be fostered through interactions with program participants. For example, safety, both physical and psychological, needs to be created, along with appropriate structure and positive social norms, supportive relationships, opportunities to belong and support for efficacy and mattering.

These lists developed by intermediary organizations highlight the goals of out-of-school time programming and emphasize the unique types of relationships and interactions that can be formed in this context of development. They build on what is known about families, parental relationships and the developmental needs of children and

youth. The standards cover basic needs including safety and support, but also emphasize the idea of belonging and worth and usefulness for young people. Also, of note, behavior management and discipline are recognized as a crucial element of effective programs, but in a less traditional way. What is emphasized is the need for structure, norms, and role models. As previously mentioned, the dynamic in out-of-school time programs is different than the structure of schools where interactions are more prescribed. Based on the roles staff and participants, out-of-school time staff may be more of a friend or confidante than a teacher would be. However, clear boundaries and consistency need to be established so all members of the program know what is acceptable or not.

Youth engagement and youth leadership are additional components of the nature of relationships established with participants that are often discussed, but less consistently so that many standards have been developed for younger elementary children. For those programs working with older youth, leadership and voice are considered more vital. Incorporating a place for youth at the table has been found by many programs to improve programming to better meet the needs and interests of young people. Additionally, including youth in the development and content of standards helps young people to understand and value the role of standards.

Cultural competence is another facet of interactions that is less commonly included on lists, but it is becoming increasingly recognized (Baca, Walter, Orlof, and Salerno, 2005; Carnegie Council on Adolescent Development, 1992; International Youth Foundation, 1999). Acknowledging and respecting differences in cultures is an important program provider characteristic to help support the healthy development of youth and to understand the unique issues that may be faced by individuals from minority cultures. Understanding the culture of participants can also help reduce conflicts that arise from competing values that may arise as individuals assimilate to the dominant culture. No research or quality standards that we have found mandates that program providers need to be of the same race/ethnicity or socioeconomic background, but young people need to feel that adults in the program appreciate and can relate to their cultural circumstances. Programs develop in certain communities for particular reasons; to survive they must be relevant, sensitive, and knowledgeable about community and family goals for children and youth to be successful.

4. Conceptual Approaches

Participant-Centered Approach

At its core, practitioners believe that a program should be participant-centered. This has to do with flexibility on the part of the program staff in order to accommodate what they learn to be the particular needs of their participants as a group and individually. This approach is very similar to individualized service delivery, which "matches well-defined goals to the specific needs and resources of the children and families who are to be served" (Shonkoff & Phillips, 2000, p. 18).

How do providers move towards meeting the needs of young people? A variety of strategies are used from reading and facilitating open discussions with participants to talking with parents and having participants complete surveys. As program providers take the time to get to know children and youth on a personal level and learn about their personal interests, strengths, and goals, they can focus on identifying other individual needs. Once needs are identified, providers can plan activities and provide resources that focus on the issues that are relevant to the population. For example, if participants are involved in risky behaviors, leaders can offer activities that provide opportunities for youth to acknowledge the effects of these behaviors. Additionally, alternatives to participating in the risky behaviors can be provided. "Programs that work recognize that, the more young people's needs are met, the more likely they are to gain competency" (National Youth Development Information Center, 2000).

Keeping the program participant-centered requires a great deal of flexibility on the part of the program staff. When programs exhibit some degree of flexibility, specifically around the activities offered, they are better able to accommodate not only the learning styles and varying interests, but also the daily moods, behaviors and attitudes, of the participants. Similarly, offering diverse activities is a good practice because it gives participants various opportunities in which to pursue their interests—an important part of positive youth development (National Research Council & Institute of Medicine, 2002). This flexibility also extends to addressing the needs and context of the participants' larger community (e.g., rural families and many working parents; A. Segal, personal communication, July 19, 2003). This would be especially important for younger children whose participation is tied to parents' schedules and access to transportation, for example. Programs could most likely maximize participation and relationships with families by taking such specific community factors into account when planning hours of operations, services offered, et cetera.

The participant-centered approach also has to do with assuring that staff agendas and ways of providing services are in a way secondary to the needs of participants—that staff respond to, and not dictate, participant involvement. Indeed, many providers agree that when staff listen and ask questions that guide a young person's own thinking, they are much more effective than simply telling that young person what to do.

Social Skills Development Approach

Despite pressure for academic and cognitive gains in out-of-school time programs, achieving social and emotional competence is an extremely important part of child development (Halpern, 2005). Effective programs recognize their role in facilitating this process. "Children can develop important interpersonal skills during out-of-school hours as they work on learning activities or join in recreation together" (U.S. Department of Education & U.S. Department of Justice, 2000). Effective programs provide opportunities for participants to develop social, emotional, intellectual, and physical skills (see Appendix C, Table C1). When young people learn skills such as teamwork, problem solving, and communication skills, they become better prepared to enter the workforce and more able to function in society. "Business people from one end

of the country to the other are saying that America's economic growth depends on all youngsters coming into the work force with skills and motivation to participate in a high-tech world" (Schorr, 1992, p. 34). "To achieve positive outcomes, programs must also find ways to expose young people to the world beyond their immediate experience; to raise their expectations of themselves and their ability to make their lives better; as well as improve the communities they live in" (Miller, 2003). Developing a young person's social skills also usually benefits their self-esteem. As young people feel more confident about and see a future for themselves, they are more likely to want to set goals and take action to reach those goals.

Comprehensive, "Whole Person" Approach

Providers generally believe that programs that are comprehensive and don't seek a "quick fix" are more effective: "Successful programs are comprehensive, intensive, and flexible" (Schorr, 1992). Within this comprehensive approach, the goal is to focus on the participant as a whole person, providing a program to meet a variety of participant needs, not just a single issue. For younger children, this often means that programs and staff attend to all the developmental domains of their young participants (A. Segal, personal communication, July 19, 2003). For older youth, it may mean addressing and linking the other areas of their lives that affect them as well (e.g., school and family) (National Research Council & Institute of Medicine, 2002). For example, a successful pregnancy prevention program recognizes that preventing pregnancy is more than just learning about sex and sexuality. It should focus on developing the adolescent's goals, relationships with positive role models, school engagement, self-confidence and decision-making skills, as well as provide information on the body and sexuality (Kirby, 2001).

Many program providers view life skills development as a key component to implementing effective programs that will benefit participants not only today but in the future. This focuses on developing skills such as communication, decision-making and anger management, as well as future goal-setting. In circumstances where the program cannot meet specific needs, providing participants with referrals and/or resources that can help is critical.

Many providers work hard to ensure that young people learn new skills as a result of programs. An effective strategy provides opportunities for learning skills as well as gaining new knowledge. As discussed above, this incorporates providing interactive, stimulating activities while engaging participants in learning about their world and their role in shaping it. These activities provide the opportunity to learn, explore and contribute, thereby teaching skills that lead to learning. It is widely believed by service providers that knowledge is gained when accurate, age appropriate, and culturally relevant information is provided. However, a variety of activities and teaching methods are needed that meet the various needs and learning levels of children and youth. Just as children and youth develop at different levels, the ways they learn also varies.

5. Program Design

Intensity, Dosage, and Duration

It is important that programs last a long enough period of time to complete key activities and, hopefully, meet expected goals. It is widely believed by service providers that a program that lasts just a few sessions is less effective than one that lasts 20 or 30 sessions. Also, shorter programs with fewer sessions allow for less relationship building between participants and staff, less flexibility in dealing with unexpected participant issues, and possible diversion from the proposed plan of a given session.

Gottfredson and Gottfredson (2002), researchers in the field of program implementation, have proposed various program component standards based on practices from successful programs as well as from "well-reasoned" and "commonsense" standards (p. 21). The authors suggest that 16 program sessions is a reasonable number of meetings for school-based prevention programs (as based on the Life Skills Training program). For mentoring programs, they propose a frequency of weekly contact for at least one year, based on the Big Brothers/Big Sisters model. Finally, the authors suggest that a commonsense standard for program duration is that it last longer than one month.

It is important to note that providers of programs that are primarily information and fact-based have witnessed the benefits of offering sessions at least two times a week.

Environment

The environment or physical space an organization operates in is integral to its functioning. Ensuring high quality or even adequate facilities is quite expensive for individual programs and communities to create an infrastructure of out-of-school time opportunities. For example, in 2000, the evaluation of Making the Most of Out-of-School Time (MOST) initiative in Chicago indicated that facilities and utilities account for up to 15-20 percent of programs' yearly costs and can be as high as 30 percent of all expenditures (Halpern, Spielberger, and Robb, 2001). Facilities must be safe, free from hazards, and healthy (Forum for Youth Investment, The After School Corporation; National After School Association, 1998). In addition, they must be inviting and have adequate resources for the activities of the program (e.g., sports equipment or arts and crafts materials). The New York Department of Youth and Community Development established a series of goals for out-of-school time providers. Providing a safe environment was the first goal. In addition to necessary resources, program requirements for safety included, complying with regulations (e.g., Americans with Disabilities Act), establishing a security plan, including emergency evacuation procedures, screening staff and volunteers for criminal conviction history and managing the safe transportation of children to programs.

Another key feature of the environment is its accessibility. Parents often regulate attendance and transportation for younger children; however as children get older it often becomes more difficult due to school and work schedules for both parents and children.

Programs often strive to operate programs where the children and youth are, such as in schools or in neighborhood locations. However, transportation is quite expensive and difficult to manage for many small programs. Moreover, even if a program is easy to walk to, changes in weather and daylight can influence attendance.

Organizational Capacity

Organizational capacity refers to the ability of programs to deliver quality services effectively. These include strength of management, soundness of finances, and a commitment to ongoing improvement, among others as a measure of performance and capacity. For programs to run smoothly, they must have leadership and clear operating practices (The Finance Project, 2005). Leadership can be at the program level with a program director, or it can extend to executive and boards of trustees, which are responsible for cultivating financial support that will sustain programs over years. Engaging children and youth for a single year and then going out of business is not useful in meeting community needs. Building strong managerial and administrative capacity to engage communities, government bodies, donors, and stakeholders is as important as having competent and committed staff who work with the youth in the program (International Youth Foundation, 1999). The size of the organization, the numbers of young people served, the types of services and the number of different revenue streams all affect the size and scope of the management team needed. All organizations do need someone who can manage and utilize staff to the fullest potential, develop policies and procedures, and management account for financial resource.

Partnerships and External Collaborations

Out-of-school time programs are distinct child development institutions with their own values, goals, and rules of interaction; however they are also neighborhood resources. In this later capacity, they cannot operate independently. In particular, best practice standards suggest that collaborating with families, school, and other community institutions is important to maximize their effectiveness (International Youth Foundation, 1999; National Research Center & Institute of Medicine, 2002). Communication facilitates acquiring the trust and relationships described as "social capital" and increases the likelihood of adequate structure. When settings, such as families, schools, and community based organizations are inconsistent, developmental opportunities will be missed and young people can be confused about adult expectations and deviant behaviors and values can take root.

Communication with families is one key avenue for programs to build meaningful connections with other key developmental contexts. Family involvement in programs increases consistency and synergy amount the different settings of children's lives. Moreover, parents' program expectations can be incorporated into operations that can help improve programs to meet the needs of the community. Although not easy to accomplish, in particular due to scheduling issues, parents can be included in many ways. For example, some programs have holiday parties in which families are included, while other programs offer events, lectures, and training just for parents. Working with

families emphasizes a team orientation and opens lines of communication. In addition to communication, families can be a source of volunteers and can assist staff with fundraising (Beckett and Jacknowitz, 2001).

Partnerships with schools and principals is another important avenue to consider, especially if the program is hoping to affect educational outcomes including school engagement, grades, achievement, persistence, and motivation. In addition, schools are often great spaces to operate after-school programs. If programs are offered in schools, transportation and safety issues are minimized.

Developing meaningful linkages with the community also supports program sustainability and ensures relevance. For example, employers may really want to hire teens, however, the youth who apply may not have adequate skills. Youth programs can incorporate relevant skill-building activities into programming and have career nights where young people can learn about different careers and local opportunities for employment. Moreover, programs can facilitate the process of bringing youth potential employers through apprenticeship programs and community service activities.

Furthermore, instead of competing with other programs also trying to survive, developing agreements with other agencies for comprehensive services or sharing expertise is one way to maintain stability and as well as support participants to get the services they need (University of CA Cooperative Extension, 2003). For example, a tutoring and mentoring program could collaborate with an after-school recreation program to provide homework help to students in need.

Research and Evaluation

Measuring outcomes and evaluating program effects and impacts are becoming increasingly intertwined with effective youth serving organizations and being included in lists of standards (e.g., American Youth Policy Forum, 2006; Carnegie Council on Adolescent Development, 1992; C.S Mott Foundation Committee on After-School Research and Practice, 2005; Bodilly and Beckett, 2005; Caplan and Calfe, 1998; McLaughlin, 2000; Pachman and Fiester, 2002; Stebins, Deich, and Hayes, 2005; U.S. Department of Education and U.S. Department of Justice, 2000). Evaluation is often the final standard and considered ideal by program staff, although it may not be fully implemented. The field is changing, and out-of-school time programs are raising the bar for themselves and taking accountability seriously. This practice has its roots in the accreditation process, but is also for continuous evaluation of program progress and effectiveness, as well as for making program improvements and mid-course corrections to ensure program quality. For example, The American Youth Policy Forum (2006), after two years of work gathering research, visiting communities, and inviting national experts and youth leaders to share their ideas about effective programs, concludes that evaluation should be considered a tool of self-improvement. They warn the programs should not necessarily hold themselves accountable for academic outcomes, but they must consider what services they provide and what other outcomes are meaningful and valuable for their program. In addition, they caution against solely relying on impact

evaluations and suggest that formative evaluation should be build into program design so that it is not seen as something separate or a burden, but as a useful means to gather information to build a strong program.

C. Chapter Summary

The ideas in this chapter suggest what matters, though not necessarily the matter, for running a high-quality out-of- school time program. Programs can be considered from many viewpoints including the human resources (staffing, leadership), guiding philosophy, organizational structure, funding, facilities/infrastructure, processes and interactions between participants, staff, and families. Most lists of quality standards make sense and when all components are combined together, a program will be well on its way to affecting positive change in its participants and community.

Developing and implementing successful programs for children and youth can be very challenging. In an era where young people are surrounded by many choices, attracting them to youth serving organizations is a challenge. The good news is that there are practices that enhance program implementation such that young people want to become involved.

1. The ideal program, according to wisdom from the field

Based on wisdom from the field, the ideal program would begin with a defined vision to guide program actions and set a strong program foundation. This sets the stage for strategic design and implementation strategies that will help the program achieve its vision. Moreover, having clear and consistent program goals are important. However, it has been found occasionally detrimental to program success when behavioral goals, particularly those related to negative or stigmatized behaviors, are articulated.

Depending on the vision, the type of program, and the children and youth the program serves, a variety of program approaches would be employed to optimize participation and positive outcomes. The participant-centered approach would allow flexibility with regards to the needs and interests of the children and youth they are serving. The social skills development approach would focus on the social, emotional, and intellectual skill development in children and youth participants. The other ideal approach is the comprehensive "whole person" approach, which focuses on the participant as a whole person by providing a program to address the many different needs of children and youth. Although these approaches differ in their focus, one or a combination of these approaches would be incorporated into the ideal program.

Wisdom from the field stresses the importance of having trained, culturally competent, and passionate staff. Staff who are able to form supportive relationships with the children and youth in their program are vital to the ideal program. The ideal program would pay staff well and provide job training and professional development opportunities. Furthermore, the activities that a program offers are vital to its success, and wisdom from the field suggests that a range of high-interest activities that engage and

build on children and youth needs and interests is most effective. The ideal program would engage youth by allowing opportunities for participants, particularly older children and youth, to take on leadership roles giving them ownership of their program. According to the voices of youth themselves, "friends, food, and fun," are crucial to the ideal out-of-school time program.

Furthermore, a successful program must have a safe, accessible environment and the organizational capacity to achieve their vision and serve children and youth, including strong management, sound finances, and a commitment from staff. Having a research component is important and would be included in the program design. Lastly, the ideal program would strive to collaborate and create relationships with families, schools, and other community organizations to provide the most positive out-of-school time program experience for children and youth.

V. CONCLUSIONS

What is known currently about program implementation? We have examined experimental studies (Chapter II), quasi-experimental and non-experimental research (Chapter III), and provider wisdom (Chapter IV) to inform this critical question. We have identified some important high-quality experimental studies. Unfortunately, they are so few in number and spread across such disparate domains that it is difficult to cobble together definitive conclusions about what program components "work" based on these studies. On the other hand, there is a host of non-experimental research on manipulable or regulatable features; unfortunately, this type of information is almost always accompanied by significant caveats. In particular, because the study designs are not experimental, it is not possible to know whether the factors associated with positive outcomes actually caused these outcomes. Finally, wisdom from the field is easily accessible and grounded in experience, but it cannot inform the field with the certainty inspired by controlled studies. Moreover, it tends to reflect the values, experiences and world view of the person providing the opinion. Taken together, however, we feel that findings from these three sources can give program providers, researchers, and funders a good idea of which regulatable features seem to matter, which are promising, which appear not to matter, and which, if any, appear to hinder program goals.

Conclusions regarding all of the specific manipulable or regulatable features that were drawn from Chapters II, III, and IV are summarized in Tables D1 and D2 (see Appendix D). Table D1 locates specific program features along the scale of negative to positive with regard to their apparent effect on child, youth, or program outcomes. The letters within the table indicate the source of information from which the conclusions originate (experimental, non-experimental, or provider wisdom). Reviewed together, it is evident that there are some differences between the findings from experimental sources and non-experimental sources. One explanation of this difference is provided by Glazerman, Levy, and Myers (2002), who found that studies of quasi-experimental design often do not produce the same results as the same investigation done with experimental design, which is generally considered the "gold standard" method. We advise the reader to review this table in conjunction with the write-ups on which its findings are based, and with Table D2—discussed below in section C—which indicates to which age groups and institutional settings the assessments specifically apply.

Below, in prose, we present overarching conclusions and descriptions of "ideal" programs based on information from experimental, non-experimental, and provider wisdom sources. The first "ideal" program is in regard across ages 6-17 and across all program types (see section B); the subsequent two descriptions are of "ideal" programs according to age groupings—middle childhood and teens an youth (see section C).

A. Overarching Conclusions

The conclusions culled from experimental, non-experimental, and provider wisdom sources are presented in detail in Tables D1 and D2 (described in sections B and C, below, and found in Appendix D). In all, almost 50 different components were addressed among the three chapters, several of which are addressed only in one chapter. There are 15 components, however, on which two or all three types of sources inform our conclusions. Often, information from these three disparate sources produced complementary findings, which is valuable to highlight. Nearly as often, however, experimental, non-experimental, and provider wisdom sources provide slightly (or sometimes wholly) contradictory conclusions; we find these particularly important to discuss. We remind the reader, again, that definitions of the terms "positive," "negative," "neutral," and combinations thereof differ slightly between the experimental, non-experimental, and provider wisdom sources; definitions are described in each chapter's methodology section. Our 14 overarching conclusions follow.

1. Staff training

Information from experimental, non-experimental, and provider wisdom sources all conclude that staff training (i.e., courses of study, workshops, or supplemental classes relevant to the program's specific participant population or program design) is linked with positive program and participant outcomes.

2. Staff experience

Practitioner wisdom suggests that experience, in combination with high-quality interpersonal skills and program-specific training, is a positive and important program component in center-based or after-school programs for older children, teens and youth. We suspect that experience is not a linear variable. Specifically, some experience seems valuable; however, too much experience may signal a service provider who is "burned out." Also, it is possible that years in the field affect providers differently, according to their institutional setting, or that working with different populations "matures" providers in different ways. For now, the effects of staff experience remain unclear. We look forward to experimental studies on this construct, across different settings and age groups.

3. Participant-staff ratio

Non-experimental information suggests lower participant-staff ratios may be beneficial for older children in center-based programs. However, provider wisdom suggests that for groups of older children, teens, or youth, there should be a limit to the number of staff present: too many can be detrimental to the comfort level of the participants and the youth-centered dynamic. The correct ratio from a provider's perspective should be left to the discretion of well-informed staff, who will decide according to their specific setting, program type, and participant group.

4. Staff wages

Both non-experimental and provider wisdom sources find that higher, more competitive, and/or more satisfying staff wages and benefits are beneficial to program outcomes. Non-experimental findings are based exclusively on studies of child care centers, whereas provider wisdom sources indicate that this component is important for staff who serve any age of participant. Provider wisdom suggests that this component is important because staff who are satisfied with program compensation are more likely to stay with their programs and may also develop higher levels of engagement with the program.

5. Interactive approach vs. information-only approach

All three types of sources—experimental, non-experimental, and provider wisdom—indicate that interactive program approaches, especially as compared to didactic, information-only approaches, are extremely promising in producing positive outcomes in older children, teens, and youth. Programs that engage participants through activities such as role-playing and group discussions in center- or classroom-based sessions, for example, have been linked with positive behavioral outcomes (e.g., decreased substance use, deferring sexual initiation). Interactive programs are unique in the opportunity they provide for communication and the exchange of ideas; it is possible that this kind of program approach particularly appeals to the way older children and teens internalize program messages. This thought is consistent with meta-analytic findings that suggest the gap in effectiveness between interactive and non-interactive programs declines as the programs reach group sizes of "a few thousand" (Tobler et al., 2000); a meaningful exchange appears much less likely in such a setting. In regard to non-interactive, information-only approaches, all three sources suggest that they are not a promising approach for generating notable positive outcomes with the exception of targeting a specific gender. Information-only approaches have been linked to changes in the level of participant knowledge about a targeted program topic, but not to changes in behaviors.

Many program approaches can be described as interactive or non-interactive. As evident in the two discussions that immediately follow, slight variations on the interactive program approach have been found to be promising as well.

6. Flexible, participant-oriented approach

Non-experimental and provider wisdom sources find dynamic, participant-oriented approaches to be particularly positive. Specifically, this supports programs that change and adapt to accommodate participants' developmental changes. Positive provider behaviors for this approach—for regular staff as well as for adult volunteers in mentoring programs—include being flexible in their interaction or teaching styles, and not imposing a preexisting, rigid "agenda" they may have for participants. This is also closely tied to encouraging teen and youth participants to have a voice in program development and activity planning, which is particularly important according to provider

wisdom; this practice addresses youths' needs to feel useful and to find reason to engage in the program. With regard to school-age and older participants, flexibility for participants to have some autonomy in selecting activities and their degree of participation in them is identified as a promising approach, in contrast to enforcing rigid provider-determined schedules and activities.

7. Multi-component and "whole person" approach

Both non-experimental research sources and provider wisdom identify multi-component and "whole person" approaches as particularly promising. These approaches are designed to address the varied needs of participants and the various contexts in which they exist (e.g., academic, family, social, residential/neighborhood-based), as well as the cross-cutting pressures endemic to playing multiple roles and having multiple needs (e.g., school-work conflict, sexual decision-making, pressure to join local gang). The goal of the whole person approach is often the promotion of multiple positive skills and changes in order to avoid or minimize negative behaviors. The multi-component program typically affects change through addressing participant needs in various domains of his or her life, or through working with and involving people (e.g., parents, teachers) and institutions (e.g., schools, communities) from these other domains. Often, programs with these approaches teach a battery of social, problem-solving, and general life skills to enhance competency in these various domains. These approaches have been discussed mostly in regard to teens and youth, and are complementary to the ideals set forth in youth development theory, but they are in line with child development researchers as well (Shonkoff & Phillips, 2000).

8. Clear and consistent program goals

Both non-experimental research and provider wisdom sources provide information on the role and importance of clear and consistent program goals. However, the chapters' conclusions address different contexts in which to communicate these goals, and are—conceptually—slightly contradictory. Non-experimental sources suggest that in prevention programs, the practice of staff articulating to participants the program goals (e.g., to prevent drug use) or the programs' stance on certain behaviors (e.g., zero tolerance for drug use) facilitates achieving program goals for the participants (e.g., no drug use). However, provider wisdom concludes that it is occasionally detrimental to program success when behavioral goals, particularly in regard to negative or stigmatized behaviors, are articulated. For example, participants and/or their families may be resistant to involvement in programs that address behaviors they feel do not apply to them (e.g., sexual health programs, if parents do not believe their teen is/would be/or should be sexually active). However, provider wisdom suggests that it is appropriate and beneficial to articulate positive program goals (e.g., teaching problem solving, refusal skills), which often attract participation; furthermore, achieving these positive goals typically affects other (potentially stigmatized) behaviors.

9. Group size

For school-age children, an experimental study found smaller group size to produce more positive academic outcomes. Non-experimental sources found, overall, smaller group sizes to be a positive, though sometimes neutral, feature for all ages. However, for youth and adolescents, provider wisdom suggests that the ideal group size may not necessarily be small. The ideal size would vary depending on the kind of program, participant needs, staff ability, and program resources. Hence, smaller group sizes appear to be generally, but perhaps not inevitably, a positive program feature.

10. Diversity of activities

Both non-experimental and provider wisdom sections suggest that it is beneficial to provide a variety of activities—that are interesting, engaging, and enjoyable—to help meet the needs of the various learning styles and interests of school-age, teen, and youth participants. There is non-experimental evidence, however, that for very young school-age children—first-grade boys, specifically—attending a program with "a larger number of different activities" was associated with negative behavioral, emotional, and academic outcomes (Pierce et al., 1999). It is possible that a great number of different activities may undermine younger children's needs for significant structure in their programs.

11. Incentives for participation

Experimental research and provider wisdom sources find that incentives, in general, are a positive program component. An experimental source found that cash incentives helped students to be more engaged in the academic programs in which they were enrolled. Provider wisdom suggests that incentives and rewards, such as trips and snacks, encourage youth to participate more frequently and, possibly, increase enthusiasm for the program.

12. Parental involvement

Experimental research, non-experimental research, and provider wisdom find that parental involvement can be a positive component in programs; some experimental and non-experimental sources, however, also show that it sometimes it appears not to matter. Overall, effects of parental involvement appear to vary by level of participation, kind of participation and reasons for participation. Experimental sources found that parental participation in literacy program and two obesity-intervention programs—one of which involved parents in weight-loss treatment, the other which focused on affecting children through improving parental knowledge, parenting skills, and parents' ability to role-model proper behaviors—produced significantly more positive impacts than involving the school-age children only. Parental involvement did not appear to matter in a social skills training program, however. Non-experimental reviews found similarly positive findings for parental interaction with school-age and teen participants' mentoring

relationships and after-school programs, but little apparent effect in services for young children, perhaps because parental involvement is ubiquitous at these ages. Provider wisdom sources, however, suggest that engaging parents in programs for younger children clearly benefits participants.

13. Dosage and duration

All sources suggest that, in certain contexts, increased intensity of involvement—through dosage and length of time in which a participant is involved—leads to more positive participant outcomes than less intensity. This pattern was not significant for all ages or contexts, however. Experimental studies found several examples of higher program dosage producing more positive and/or longer-lasting outcomes among teens and youth in mental health and family reunification programs and among younger children in education and low birth-weight intervention programs,[10] for example, but not for certain risky behavior prevention or juvenile justice probation programs. Importantly, it became evident that slight differences in frequency often did not produce significantly different outcomes. Non-experimental research has found both higher dosage and longer duration to be positive; duration was identified as particularly important in the context of mentoring relationships. Provider wisdom identifies sufficient dosage and duration as one of the most important aspects of programming. The program design must incorporate enough time and meetings for program providers to be able to address participant needs, and for participants to achieve program goals despite any unanticipated delays that may occur. Between experimental, non-experimental, and provider wisdom sources, all age ranges and program settings recorded on Table D2 are found to benefit from some combination of dosage and duration.

14. Overall program quality

It seems intuitive and self-evident that "overall quality"—often described in the context of a program meeting or surpassing a battery of established standards, or following various best practices—is a program characteristic that would be linked to positive participant and program outcomes. However, it is important to note that non-experimental research and provider wisdom sources both corroborate this association. Non-experimental sources find that quality—most often, environmental, as measured by several different established instruments—is related to positive child outcomes in the child care context even after controlling for various different other potentially influential participant, center, and location-based factors. The provider wisdom discussion finds overall quality to be important, as well, identifying it as the end product of different practices and processes that a given program employs and the quality with which each of these individual components is implemented.

[10] The positive outcomes in the early childhood education program are attributable to higher rates in both dosage and duration.

B. What Would A Generic "Ideal" Program Look Like?

According to all of the information we have amassed from experimental, non-experimental, and provider wisdom sources, what would be an ideal program across different age groups and institutional settings? Answering this question is a bit like assembling an ancient vase from fragments found in an archeological dig. The pieces are not all there, and many of the pieces are chipped. In other words, the evidence is incomplete and variable in how definitive it is. Nevertheless, there are instances were the evidence from all three methodologies converges. Moreover, when evidence is available from multiple methodologies, it tends to point in the same direction in a number of instances.

An ideal program, according to all three methodologies, would benefit from having staff with training that is specific to the program and participant age group, and having more rather than less dosage and duration (i.e., greater quantity, intensity and length). In addition, an ideal program would use an interactive approach to teach and communicate messages to children or youth. In most cases, didactic, information-only approaches would not be employed. Many publications reviewed in these three sources also suggest that an ideal program would incorporate parental involvement to some degree. Two studies (one experimental, one non-experimental) found that parental involvement appeared not to impact children's outcomes, but some type of parental involvement for young children seems both very likely and essential. In most cases, the ideal program would enroll smaller rather than larger groups of participants, and have lower rather than higher participant-to-staff ratios.

The evidence base is thinner for several additional program features, but these program components are still supported by two methodologies. Non-experimental research and provider wisdom jointly support a number of program features. These include the flexible, participant-centered approach, which dictates that the program adjusts and refocuses as necessary to accommodate changing participant needs, and the multi-component and "whole person" approaches, which address the multiple domains of development and the multiple pressures that participants experience in the various contexts they inhabit (e.g., school, family, neighborhood). Communicating clear and consistent messages about positive program goals would also typically be a positive practice according to these two sources; provider wisdom suggests, however, that programs would do well to downplay addressing stigmatized behaviors, and highlight the development of positive behaviors. Other positive practices include paying staff higher and more competitive salaries, establishing positive and warm caregiver-participant relationships, and offering diverse, engaging, and interesting program activities. Also, provider wisdom and non-experimental research suggest that the ideal program would have a high level of overall program quality—often described as the end result of a program meeting or surpassing a battery of established standards, or following various best practices. Finally, experimental research and provider wisdom support the practice of providing incentives for participants to achieve identified goals or attend the program.

Finally, a number of ideas come up in just one source, often from providers or non-experimental research. Some of the provider wisdom reflects common sense, while some of it may not be obvious or even salient for many adults. These program elements include program accessibility, developmentally appropriate behavior management and discipline, a safe and stable atmosphere, targeting the participation of hard-to-reach populations, staff retention (minimal staff turnover), program evaluation, and a positive, strength-based program philosophy. The provider perspective also notes that staff need a combination of commitment, positive personality traits, training, cultural competence, and experience with participant populations. These staff then need competent and consistent supervision, management, and support from directors, as well as opportunities and support for professional development. Finally, provider wisdom identifies collaboration with institutions and professionals in the community as a promising program practice. Non-experimental research identifies implementation fidelity to theory-based program design, and the screening, training, and supervising of adult program volunteers (e.g., mentors), as beneficial. The evidence with regard to the degree of flexibility or structure in programs seems to vary by age of child. Ideally, greater program flexibility would be featured in programs for older children and teens, while less flexibility and more structure would be featured in programs for younger children.

Several program elements consistently do not seem to matter. For example, in this ideal program, work with teens would not necessarily have to be delivered by adults; experimental findings show that trained peer leaders can work as well as (and sometimes better than) trained adults. Also, non-experimental research suggests that program leaders or mentors would not have to be of the same demographic characteristics as participants; it may be more important that the mentoring pair have the same interests, or that program providers are respectful of and sensitive to cultural differences.

Finally, this generic, ideal program would do well to avoid certain program practices. Non-experimental analyses have linked greater program flexibility (for younger boys only) and the use of "shock" approaches with negative participant outcomes.

C. What Would an Age-Specific "Ideal" Program Look Like?

The characteristics of a given program obviously need to vary depending upon the age of the participant. Our primary goal is to identify broad program elements that work across age groups. However, most studies are conducted on narrow age groups, and—unless research has been conducted across a variety of ages—it is difficult and even somewhat risky to extrapolate the results to all age groups. Accordingly, in this section, we summarize the evidence based on findings for two specific age groups: middle childhood (ages 6 to 11 or 12) and teens and youth (ages 12-13 and older). Additionally, we discuss in the last subsection implementation findings specific to program staff.

Age-specific findings, along with findings specific to the program's institutional setting, are organized in Table D2 (see Appendix D). This table is meant to represent the

various ages and contexts to which studies that informed our assessments applied. We advise the reader to review this table in conjunction with the write-ups on which it is based, particularly because most of its assessments—divided across different age groups and institutional settings—are based on a smaller number of studies than are our overall component assessments (Table D1). The letters within the table and within the summaries, below, indicate the source of information—experimental (Ex), non-experimental (N), or provider wisdom (PW)—from which the findings and conclusions originate.

Middle childhood

For participants in middle childhood—roughly ages 6 to 11 or 12—the ideal program would be staffed by caregivers who had received training specific to the program context and/or age group (Ex, N, PW; particularly in regard to their unique developmental needs), who receive regular and appropriate professional management, support, and development, and who had a combination of experience in the field and a positive outlook toward the program and participants (PW). Also, adult volunteers such as mentors would undergo comprehensive screening, significant program training, and staff supervision (N). In general, the program would have a smaller, versus larger, participant-to-staff ratio (N, PW); however, provider wisdom asserts that if the goals or format of the program are considered by staff or program designers to benefit from fewer staff per children, the ratio would be determined according to these needs (PW). Staff would receive competitive salaries and benefits packages (PW), turnover would be minimal (PW), and staff would foster positive, warm relationships with the participants (N, PW). It would be important that staff be culturally competent, and able to address and accommodate cultural differences (PW). It would not necessarily be important to match caregivers to participants based on race, ethnicity or gender (N).

The conceptual design of the program would be flexible and participant-centered, meaning the programs would adjust to and change with the participant's needs (N, PW). It would communicate clear and consistent messages about desired participant goals and program philosophy (N), it would be interactive in approach rather than information-only (Ex, N, PW), and it would address the "whole person," teaching the socio-emotional skills needed to flourish in every arena of the participants' lives (social, academic, neighborhood, family, etc.; N, PW). The program would be implemented with fidelity to its research-based design (N).

Many different program practices could be used in this ideal program. In general, the program would have a smaller, versus larger, group size (N, PW); however, if the goals or format of the program are considered by staff or program designers to benefit from a large number of participants, the group size would be determined according to these needs (PW). The program would offer diverse activities to all but the youngest participants, who tend to need structured activities (N, PW); all activities would be designed to be interesting, engaging, and enjoyable to children in this age range (N, PW). Staff would employ developmentally appropriate behavior management and discipline (PW). The program might provide incentives for participation or for achieving certain

goals (Ex). The program would be geographically accessible to its target participants, and its environment would be safe and stable (PW). If possible, the program would include parental involvement, though one of four experimental studies reviewed found that parental involvement not be an essential component for older children (Ex, N). In programs structured around caseworkers, participants and/or their families would receive intense case management or multisystemic therapy, rather than traditionally less intense services (Ex). In cases of child custody disputes, the ideal program would resolve the dispute with mediation instead of litigation (Ex). At predetermined intervals, the program would undergo assessment and evaluation, and implement changes accordingly (PW). In the community, the program staff would have positive relationships with schools and other institutions, in order to maximize various resources available to them (PW). The program design would involve higher, rather than lower, levels of dosage and duration (Ex, N, PW). Finally, the ideal program, as a whole, would be of high quality (PW).

Teens and youth

Surprisingly little experimental research informs program implementation for teens and youth. For teens and youth, the ideal program would be staffed by caregivers who had received training specific to the program context and/or age group, particularly in regard to their unique developmental needs (Ex, N, PW), who receive regular and appropriate professional management, support, and development, and who had a combination of experience in the field and a positive outlook toward the program and participants (PW). Also, adult volunteers such as mentors would undergo comprehensive screening, significant program training, and staff supervision (N). In most cases, the optimal participant-to-staff ratio would be determined by well-informed staff or program designers according to the particular needs of the participants and program staff, and the goals and format of the program (PW). Staff would receive competitive salaries and benefits packages (PW), turnover would be minimal (PW), and staff would foster positive, warm relationships with the participants (N, PW). Staff would be culturally competent, able to address and accommodate cultural differences (PW). It would not necessarily be important to match caregivers to participants based on race, ethnicity or gender (N). For prevention programs, it may be more promising to employ peer leaders than adult leaders (Ex).

The conceptual design of the program would be flexible and participant-centered, meaning the programs would adjust to and change with the participant's needs (N, PW). It many cases, it would be a promising approach to communicate clear and consistent messages about desired participant goals and program philosophy (N, PW). However, for programs with goals addressing stigmatized behaviors (e.g., sexual or criminal activity), it may be more productive to emphasize positive secondary program goals (e.g., decision-making or job skills); this approach helps engage youth and families who otherwise may feel alienated or offended by the focus on negative behaviors (PW). Similarly, the ideal program would embrace a "strength-based" philosophy, which emphasizes positive development and considers youth to be "at promise" rather than "at risk" (PW). The program would be interactive in approach rather than information-only

(N, PW), and it would address the "whole person," teaching the socio-emotional skills needed to flourish in every arena of the participants' lives (social, academic, neighborhood, family, etc.; N, PW). The program would be implemented with fidelity to its research-based design (N). Finally, "shock" approaches—in which youth are made to visit prisons and observe the "horrors and difficulties of life" there—would not be employed (N).

Many different program practices could be employed in this ideal program. In general, the program would have a smaller, versus larger, group size (N, PW); however, if the goals or format of the program are considered by staff or program designers to benefit from a large number of participants, the group size would be determined according to these needs (PW). The program would offer diverse activities (PW), all of them designed to be interesting, engaging, and enjoyable to children in this age range (N, PW). Staff would employ developmentally appropriate behavior management and discipline (PW). Furthermore, leadership opportunities would be available for participants (PW). The program might provide incentives for participation or for achieving certain goals (PW), though experimental research shows that case management may be even more successful at promoting positive outcomes than incentives (Ex). The program would be geographically accessible to its target participants, and its environment would be safe and stable (PW). Special efforts would be made to attract participants in this age group, as program participation must compete with myriad other potential distractions and activities that the youth can choose to do instead (PW). If possible, the program would incorporate collaboration with the community (e.g., financial supporters, professional contacts for participants, informal participant recruiters; PW) and parental involvement (N). If possible, it would employ case management instead of financial incentives in order to keep participants engaged (Ex). If employing videotape instruction, the ideal program would accompany it with interaction with and skills training from a program leader (Ex). At predetermined intervals, the program would undergo assessment and evaluation, and implement changes accordingly (PW). The program design would involve higher, rather than lower, levels of dosage and duration (Ex, N, PW). Finally, the overall quality of the program—the combination of its various practices and processes—would be of high quality (PW).

D. Population Effects

It is important to note, in light of our findings, that every one of our promising components may not work with every population. Research from across the field of program evaluation has shown that program participation may have significantly different effects among different populations, and that population characteristics may, in fact, may moderate program effects. Several studies have found that high-risk (i.e., some combination of disadvantaged (Currie, 2000; Jekielek, Moore et al., 2002; Peisner-Feinberg & Burchinal, 1997), and/or minority (McGill et al., 2001; Sullivan-Temple & Ravid, 1991) children from low income (Michelson, Zaff, & Hair, 2002; Simpkins, 2003), high-crime (Vandell & Shumow, 1999) neighborhoods, and/or mothers with less education (Vandell & Wolfe, 2000)) participants appear to experience positive program effects at a greater magnitude than lower-risk peers (Dishion & Andrews, 1995; Jekielek,

Cochran, & Hair, 2002; Peisner-Feinberg & Burchinal, 1997; Peisner-Feinberg et al., 2001; Webster-Stratton, Reid, & Hammond, 2001). Age may also moderate program effects. Mentoring relationships appear to be closer between adults and elementary school mentees than adults and high-school mentees (Herrera et al., 2000), and employment programs may benefit younger participants more than older participants (Jekielek, Cochran et al., 2002). Participation in out-of-school time programs is more strongly associated to positive outcomes among adolescents than elementary school-age children (Simpkins, 2003).

E. Next Steps

In the future, we hope that the funders, policy-makers, and service providers increasingly sees the value of more, and higher-quality, experimental studies. Components discussed as "lone findings" in Chapter III and as "gaps" in the field cited throughout Chapter II would benefit from more research, as would studies of historically understudied groups, such as Asian American, Latino American, immigrant, and sexual minority populations. In addition, we need cost-benefit data to inform discussions about program features to implement. Work on these "next steps" will benefit from the numerous insights and the many sources reviewed in this report, as well as many other rich sources that may not have been included here. A partial list of other sources valuable to the discussion on program implementation is presented in Appendix E.

We hope that this report provides a starting point from which to assemble stronger programs. It is an appropriate and attainable goal that practitioners, researchers and funders have access to solid and clear guidance on "what works" in program implementation.

Appendix A
Source Information from Chapter II (Experimental Research)

F. Table A1

III. Findings on Staff and Staffing from Experimental Research

	Study Design		Program and Sample Descriptions			Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Program Information**	**Population**	**Difference in Effectiveness**
Jemmott III, J.B., Jemmott, L.S., Fong, G.T., (1998). Abstinence and safer sex HIV risk-reduction interventions for African American adolescents. *Journal of American Medical Association, 279*(19), 1529-1536.	Type of activity leader	Random assignment to: Experimental peer-led program or adult-led program	Health Risk/Abstinence and safer-sex HIV intervention program	Behavior theory based HIV prevention program 2-week program by trained facilitators. Students were placed in safer sex program, abstinence program, or health promotion program. Location. Inner city of Philadelphia.	Low-income African American children. Mean age was 11.8 years. A little over half were female.	NEUTRAL: The study did not find any differences in the outcomes by type of faciliator. Trained peer facilitators did not differ in effectiveness from trained adult facilitators.
Murray, D.M., Richards, P.S., Luepker, R.V., & Johnson, C.A. (1987). The prevention of cigarette smoking in children: Two- and three-year follow-up comparisons of four prevention strategies. Journal of Behavioral Medicine, 10(6), 595-611.	Type of activity leader	Random assignment to: Experimental peer-led or adult-led program (by aproach: health consequences, usual care comparison, and social influences curricula)	Health Risk/Substance abuse preventiion (cigarette smoking)	Students from three Minnesota districts participated in smoking prevention program with a social influences curriculum, health consequences curriculum, or a usual-care curriculum. One of the peer-led programs used a video supplement.	Seventh grade students in 8 mostly suburban junior high schools. Mean age at the initiation of the study was 12 years of age. The sample was mostly White and contained about an equal proportion of males and females.	POSITIVE TO NEUTRAL: Peer led, social influences condition was more effective than adult-led in delaying the onset of smoking. Peer and adult leaders otherwise equivalent.

	Study Design		Program and Sample Descriptions			Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Program Information**	**Population**	**Difference in Effectiveness**
Nye, B., Hedges, L.V., & Konstantopoulos, S. (1999). The long-term effects of small classes: A five-year follow-up of the Tennessee Class Size Experiment. *Educational Evaluation and Policy Analysis,* 21(2), 127-142.	Group size and child-teacher ratio	Random assignment to: Classes with 13-17 students; or classrooms with 22 to 26 students per classroom; or classrooms with 22-26 students with a teacher along with a paid full-time classroom aide.	Education/Early elementary grades	Elementary school classes in Tennessee.	Children in kindergarten through third grade. The follow-up data were based on 4,944 of the original 6,572 students who were randomly assigned to 331 classes in 76 schools.	POSITIVE: Assignment to smaller classes resulted in better academic outcomes through the five-year follow-up.
Olds, D. L., Robinson, J., O'Brien, R., Luckey, D. W., Pettit, L. M., Henderson, C. R., et al. (2002). Home visiting by paraprofessionals and by nurses: a randomized, controlled trial. *Pediatrics*, 110(3), 486-496. Olds, D. L., Robinson, J., Pettit, L. M., Luckey, D. W., Holmberg J., et al. (2004). Effects of Home Visits by Paraprofessionals and by Nurses: Age 4 Follow-up Results of a Randomized Trial. Pediatrics, 114(6), 1560-1568.	Types of health home visitor	Random assignment to: paraprofessionals or nurse home visitors	Health/Maternal and child health home visiting	Home visiting services in Denver, CO metropolitan area.	735 mostly Hispanic (about 45%) or White (about 35%) pregnant low-income women experiencing their first live births and their children from pregnancy to two years after birth.	POSITIVE TO NEUTRAL: Nurses perform better than controls on numerous measures, while paraprofessionals perform better than control on only a few measures. Size of effects of nurses about twice the size of those for paraprofessionals. Outcomes for children in nurse-visited homes and paraprofessional-visited homes did not differ statistically on most outcomes, with the exception of language development (for which children in nurse-visited homes were superior).
Perry, C.L., Telch, M.J., Killen, J., Burke, A., & Maccoby, N. (1983). High school smoking prevention: The relative efficacy of varied treatments and instructors. Adolescence, 18(71), 561-566.	Type of instructor	Random assignment to: health classes with experimental trained college students or trained school teachers	Health Risk/Smoking prevention	The High School Smoking Prevention Program for high school students at four schools in northern California.	10th grade students in 20 classes.	NEUTRAL: No difference by instructor-type in level of success in decreasing smoking among 10th graders. Some empirical analyses suggested that younger instructors more effective with social pressure curriculum than the health effects curriculum, but not experimentally based.

	Study Design		Program and Sample Descriptions			Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Program Information**	**Population**	**Difference in Effectiveness**
Ross, J.G., Nelson, G.D., & Kolbe, L.J. (Eds.). (1991). Teenage health teaching modules evaluation [Special issue]. Journal of School Health, 61(1).	Type of instructor	Random assignment to: Training.	Health Risk/Health education for reducing substance use.	Healthy behavior promotion and health risk behavior prevention program at around 85 junior and senior high schools in Maryland, Vermont, and San Diego County, CA. The Teenage Health Teaching Modules (THTM) curriculum, developed using CDC funding, was used.	4,806 students.	POSITIVE: Training had positive impacts on the implementation of the curriculum and a variety of health risk outcomes.
Sussman, S., Sun, P., McCuller, W. J., & Dent, C. W. (2003). Project towards no drug abuse: Two-year outcomes of a trial that compares health educator delivery to self-instruction. Preventive Medicine, 37, 155-162.	Mode of instruction, service delivery	Random assignment to: Standard care; Self-instruction; or Health Educator-led classroom instruction.	Substance use prevention	12 session drug abuse prevention program	High school students from 18 schools.	Self-instruction was not effective compared to control; Instruction by health educator produced positive impacts.
Valente, T. W., Hoffman, B. R., Ritt-Olson, A., Lichtman, K., & Johnson, C. A. (2003). Effects of a social-network method for group assignment strategies on peer-led tobacco prevention programs in schools. *Americal Journal of Public Health, 93*(11), 1837-1843.	Instructor; Leader selection (method of selection)	Random assignment to: Random condition; Social Network condition; or Teacher condition	Substance use prevention	Leaders were selected based on condition criteria and were linked with students	6th graders in 84 classrooms in 16 schools (1961 students)	Network method was the most effective way to structure the program and was related to improved attitudes compared to random condition

G.

H. Table A2

IV. Findings on Program Practices from Experimental Research

	Study Design		Program and Sample Descriptions			Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Program Description**	**Population**	**Difference in Effectiveness**
Burrow, F. B. (2001). The effect of parental involvement on social skills training for children with and without attention deficit hyperactivity disorder. *Dissertation Abstracts International: Section B: The Sciences & Engineering (62),* 4-B, 2048.	Parental involvement	Random assignment to: experimental group with social skill training plus parental training or control group with social skill training	Social Functioning/Social Skill Training	Social skill training.	42 elementary school children, including those with attention deficit hyperactivity disorder (ADHD).	NEUTRAL: Both programs improved social behaviors but no statitistical differences between the programs with and without parental involvement
Emery, R.E., Laumann-Billings, L., Waldron, M.C., Sbarra, D.A., & Dillon, P. (2001). Child custody mediation and litigation: Custody, contact, and coparenting 12 years after initial dispute resolution. *Journal of Consulting and Clinical Psychology, 69*(2), 323-332.	Mediation	Random assignment to: litigation versus mediation approach	Social Functioning/Child Custody Disputes	Litigation or mediation services for recently divorced couples with a child custody dispute in central Virginia.	52 families were included in this 12-year follow-up. The sample was mostly low SES and predominantly White (79%) and Black (21%).	POSITIVE: Mediation better than litigation in improving various family functioning outcomes
Epstein, L. H., Valoski, A., Wing, R. R., & McCurley, J. (1990). Ten-year follow-up of behavioral family-based treatment for obese children. *Journal of American Medial Association, 264*(19), 2513-2523.	Parental involvement	Random assignment to: child and parent-focused; child-focused; or no focus.	Obesity	Child and parent focused program treated parents' obesity aslong with their children's.	67 children ages 6 to 12.	POSITIVE: Child-parent focused program was the only program that maintained a long-term loss of percent overweight.
Flay, B. R., Graumlich, S., Segawa, E., Burns, J. L., & Holliday, M. Y. (2004). Effects of 2 prevention programs on high-risk behaviors among African American youth - A randomized trial. *Archives of Pediatrics & Adolescent Medicine, 158*(4), 377-384.	Mode of instruction	Random assignment to: Social Development Curriculum (SDC); or School/Community Intervention (SCI); or an attention-placebo Health Enhancement Curriculum (HEC)	Risky behavior prevention (sexual, substance use, and delinquency).	Varied by curriculum.	Participants in grades 5-8 and their parents and teachers	Positive for both SEC and SCI compared to control. SCI was more effective than SDC on combined behavioral measure.

	Study Design		Program and Sample Descriptions			Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Program Description**	**Population**	**Difference in Effectiveness**
Golan, M., Fainaru, M. and Weizman, A. (1998). Role of behaviour modification in the treatment of childhood obesity with the parents as the exclusive agents of change. *International Journal of Obesity (22),* 1217-1224.	Parental involvement	Random assignment to: parents-only program or children-only program.	Obesity	Either children or parents were solely responsible for losing children's weights.	6 to 11 year olds (n = 60)	POSITIVE: Parents-focused had much larger impacts on children's behavior and weight loss than children-focused.
Lederman, R. P., & Mian, T. S. (2003). The Parent-adolescent Relationship Education (PARE) Program: A curriculum for prevention of STDs and pregnancy in middle school youth. *Behavioral Medicine, 29*, 33-41.	Mode of instruction	Random assignment to: Social learning education group; or Traditional didactic teaching	Pregnancy and STD prevention	Family communication program with sessions held covering issues on risks of sexual activity and booster sessions held at 6 month intervals	Middle school students and their parents	Social learning experimental group expressed stronger intentions to postpone sex; No impact (difference) on other outcomes.
Leve, L. D., Chamberlain, P., & Reid, J. B. (2005). Intervention outcomes for girls referred from juvenile justice: Effects on delinquency. Journal of Consulting and Clinical Psychology, 73(6), 1181-1183.	Mode of Instruction	1.) Multidimensional treatment foster care; or 2.) Group care	Delinquency prevention program (to prevent recidivism)	Delinquency intervention to reduce number of days spent in locked settings and fewer reports of criminal activity.	81 Adolescent girls with chronic delinquency	MTFC had greater reduction in locked settings and in delinquency (as reported by caregivers).
Lipkus, I. M., McBride, C. M., Pollak, K. I., Schwartz-Bloom, R. D., Tilson, E., & Bloom, P. N. (2004). A randomized trial comparing the effecs of self-help materials and proactive telephone counseling on teen smoking cessation. *Health Psychology, 23*(4), 397-406.	Mode of service delivery	Random assignment to: Written self help materials plus a video; or Written self help materials plus video and telephone counseling.	Substance use prevention	Teen smoking cessation program	402 adolescent smokers from southeaster U.S.	No significant difference were found between the smoking cessation rates for the two groups, although the group with the telephone counseling had slightly (ns) higher cessation rates at all three follow-up points.
Lochman, J. E., & Wells, K. C. (2004). 72. *Journal of Consulting and Clinical Psychology, 72*(4), 571-578.	Parent involvement	Random assignment to: child only component; full program with child and parent component; control condition	Delinquency prevention	Counseling plus	Aggressive pre-adolescent boys and their parents.	POSITIVE: Larger effects for full program with parent and child component.

	Study Design		Program and Sample Descriptions			Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Program Description**	**Population**	**Difference in Effectiveness**
Morrow, L. M. and Young, J. (1997). A family literacy program connecting school and home: effects on attitude, motivation, and literacy achievement. *Journal of Educational Psychology(89),* 4, 736-742.	Parental involvement	Random assignment to: literacy program with family components or literacy program without family components	Education/Literacy	School-based literacy program. The program with family components taught parents how to conduct literacy activities with children at home.	56 children in grades 1 through 3.	POSITIVE: The program with family literacy components produced larger impacts on children's literacy skills than school-based only program.
Reid, W. J., Bailey-Dempsey, C., Cain, E., Cook, T. V., & Burchard, J. D. (1994). Cash incentives versus case management: Can money replace services in preventing school failure? Social Work Research, 18(4), 227-236; and Reid, W.J. & Bailey-Dempsey, C. (1995). The effects of monetary incentives on school performance. Families in Society: The Journal of Contemporary Human Services, 331-340.	Incentives	Random assignment to: Cash incentives offered versus case management.	Education/School failure prevention	Case management program or cash incentive program designed to prevent school failure and to improve school attendance in two middle schools and one high school in a medium-sized New England city. Michael Klein Project for Students at Risk.	56 low-income White girls at risk of school failure, ages 11 to 17 (mean age=14)	POSITIVE TO NEUTRAL: Both cash incentives and case management were more effective than no intervention. Case management more effective than cash. But not getting cash reward-not meeting strict improvement requirements to gain reward seemed to drive cash incentive program findings that girls had lower self-esteem. Authors hypothesized incentive requirements may have been too stringent given that many girls did improve, and that it may have been seen more punitively rather than as an incentive.
Rowland, M. D., Halliday-Boykins, C. A., Henggeler, S. W., Cunningham, P. B., Lee, T. G., Kruesi, M. J., & Shapiro, S. B. (2005). A randomized trial of multisystemic therapy with Hawaii's Felix class youths. *Journal of Emotional and Behavioral Disorders, 13*(1), 13-23.	Mode of service delivery	Random assignment to: Multisystemic therapy or Traditional Coninuum of Care services	Mental health problem prevention	Multisystemic Intervention designed to improve mental health and reduce minor criminal activity	31 young people diagnosed with serious emotional disturbance (SED)	MST participants had lower internalizing, externalizing symptoms and less involvement in minor criminal activity and fewer days in out of home placement.

	Study Design		Program and Sample Descriptions			Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Program Description**	**Population**	**Difference in Effectiveness**
Schaeffer, C. M., & Borduin, C. M. (2005). Long-term follow-up to a randomized clinical trial of multisystemic therapy with serious and violent juvenile offenders. *Journal of Consulting and Clinical Psychology, 73*(3), 445-453.	Mode of service delivery	Random assignment to: Multisystemic therapy; or Individual therapy.	Delinquency prevention	Multisystemic Intervention designed to lower recidivism rates	176 10-15 year olds with a history of detention for violent offenses.	MST participants had lower recidivism rates than IT participants
Schinke, S., & Schwinn, T. (2005). Gender-specific computer-based intervention for preventing drug abuse among girls. The American Journal of Drug and Alcohol Abuse, 31, 609-616.	Mode of instruction	1) Gender specific computer intervention or 2) Conventional intervention (not on computer)	Substance use prevention program	Technology based intervention designed for girls to prevent drug and alcohol use	Adolescent girls (7th graders in NYC middle schools)	GSI girls had better attitudinal outcomes
Telch, M.J., Miller, L.M., Killen, J.D., Cooke, S., & Maccoby, N. (1990). Social influences approach to smoking prevention: The effects of videotape delivery with and without same-age peer leader participation. Addictive Behaviors, 15, 21-28.	Mode of instruction	Random assignment to: Classrooms with one of three conditions: videotape social resistance training with peer leader; videotape lesson alone; or control group receiving no treatment.	Health Risk/Substance use prevention (smoking)	Smoking prevention program in Southern California using Resisting Pressures to Smoke videotape. Based on the peer-led Project C.L.A.S.P. program.	540 7th grade students in 15 classrooms.	POSITIVE: Group with peer leaders and tape had significantly lower smoking adoption rate for in comparison to group viewing just the tape.

	Study Design		Program and Sample Descriptions			Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Program Description**	**Population**	**Difference in Effectiveness**
Winett, R.A., Anderson, E.S., Moore, J.F., Taylor, C.D., Hook, R.J., Webster, D.A., Neubauer, T.E., Harden, M.C., & Mundy, L.L. (1993). Efficacy of a home-based Human Immunodeficiency Virus prevention video program for teens and parents. Health Education Quarterly, 20(4), 555-567.	Mode of instruction	Random assignment to: experimental program using a home-based video with informational and skills training or control group program using a home-based video with information-only	Health Risk/HIV prevention and to increase family communication.	Family/Media AIDS Prevention Project in Roanoke, VA.	70 families with a 12 to 14 year old boy or girl were recruited through their primary care physicians.	POSITIVE TO NEUTRAL: Both informational and skills-focused video treatments increased knowledge for parents and teens, but skills-training plus video demonstrated increased family problem-solving and assertiveness behaviors.

I. Table A3

V. Findings on Dosage and Duration from Experimental Research

	Study Design		Program and Sample Descriptions			Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Program Information**	**Population**	**Difference in Effectiveness**
Barton, W. H., & Butts, J. A. (1990). Viable options: intensive supervision programs for juvenile delinquents. *Crime and Delinquency, 36*(2), 238-256.	Dosage and duration	Random assignment to: experimental program with higher dosages or traditional aftercare program with lower dosages	Social functioning/ Aftercare probation for juvenile offenders	Social casework and surveillance	500 males at the mean age of 15 mostly with prior charges with serious crimes	NEUTRAL: No impact on the average number of criminal charges at a two-year follow up after the street time was controlled.

	Study Design		Program and Sample Descriptions			Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Program Information**	**Population**	**Difference in Effectiveness**
Burns, B.J., Farmer, E.M.Z., Angold, A., & Costello, E.J. (1996). A randomized trial of case management for youths with serious emotional disturbance. *Journal of Clinical Child Psychology*, *25*(4), 476-486.	Dosage and duration	Random assignment to: experimental case management or primary mental health clinician.	Mental health/Serious emotional disturbance	Case management by multiagency treatment team.	167 children with serious emotional disturbance.	POSITIVE: Intensive case management related to longer participation; wider use of community based services; and fewer inpatient days.
Fraser, M.W., Walton, E., Lewis, R.E., Pecora, P.J. and Walton, W.K. (1996). An experiment in family reunification: Correlates of outcomes at one-year follow-up. *Children and Youth Services Review, 18*(4/5), 335-361.	Dosage and duration	Random assignment to: two groups with different dosages.	Social functioning/ Family reunification services	Social casework.	110 families	POSITIVE: Produced positive impacts on family reunification.
Greenwood, P. W., Deschenes, E. P., & Adams, J. (1993). *Chronic juvenile offenders: Final results from the Skillman After-care Experiment*. Santa Monica, CA: The RAND Corporation.	Dosage and duration	Random assignment to: experimental program with higher dosages or traditional aftercare program with lower dosages	Social functioning/ Aftercare probation for juvenile offenders	Social casework and surveillance	More than 50 youth at the average age of 17 in each of two sites. Predominantly black males and habitual delinquents.	NEUTRAL: No impact on recidivism and drug use.
Kamb, M.L. et al (1998). Efficacy of risk-reduction counseling to prevent Human Immunodeficiency Virus and Sexually Transmitted Diseases. *Journal of American Medial Association, 280* (13), 1161-1167.	Dosage and duration	Random assignment to: two experimental programs with different dosage levels of interactive counseling sessions or one control group that received two brief messages.	Health Risk/ HIV and STD prevention	Counseling regarding HIV and STDs at STD clinics.	More than 4300 HIV-negative adults and adolescents who visited inner-city STD clinics	NEUTRAL: Both experimental groups produced positive impacts on sexual behaviors and contraction rates but no difference was found between the two groups.

	Study Design		Program and Sample Descriptions			Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Program Information**	**Population**	**Difference in Effectiveness**
Land, K.C., McCall, P.L., & Williams, J.R. (1990). Something that works in juvenile justice: An evaluation of the North Carolina Court Counselors' Intensive Protective Supervision randomized experimental project. Evaluation Review, 14(6), 574-606.	Dosage and duration	Random assignment to: experimental program with higher dosages or traditional aftercare program with lower dosages	Social functioning/Juvenile crime prevention	Social casework and surveillance	93 females ages 14 and 15	POSITIVE: Reduced the percentage of rearrests but only among first-timers.
National Council on Crime and Delinquency. (1987). *The impact of juvenile court intervention*. San Francisco, CA: Author.	Dosage and duration	Random assignment to: experimental program with higher dosages or traditional aftercare program with lower dosages	Social functioning/Juvenile crime prevention	Social casework and surveillance	Serious delinquents in Utah.	NEUTRAL: No impact on recidivism.
Project MATCH Research Group. (1998). Matching patients with alcohol disorders to treatments: Clinical implications from Project MATCH. *Journal of Mental Health (UK), 7*(6), 589-602.	Dosage and duration	Random assignment to: two experimental programs with different dosage levels of counseling sessions or one control group with no session.	Health risk/ alcoholism treatment	Counseling regarding alcoholism and risk behaviors.	Not limited to children and youth.	NEUTRAL: Both experimental groups produced positive impacts on risk behaviors but no difference between the two groups was found.
Rhodes, F., & Malotte, C. K. (1996). HIV risk interventions for active drug users. In S. Oskamp & S. C. Thompson (Eds.), *Understanding and Preventing HIV Risk Behavior: Safer Sex and Drug Use* (pp. 211-214). Thousand Oaks, CA: Sage Publications Inc.	Dosage and duration	Random assignment to: standard two sessions or standard two sessions plus individual counseling and behavioral skills training	Health risk/ HIV, STD and drug use prevention	Counseling and behavioral skills training.	Not limited to children and youth.	POSITIVE: Higher dosages produced positive impacts on risk behaviors.

	Study Design		Program and Sample Descriptions			Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Program Information**	**Population**	**Difference in Effectiveness**
Sontheimer, H., and Goodstein, L. (1993). An evaluation of juvenile intensive aftercare probation: aftercare versus system response effects. *Justice Quarterly, 20*(2), 197-227.	Dosage and duration	Random assignment to: experimental program with higher dosages or traditional aftercare program with lower dosages	Social functioning/ Aftercare probation for juvenile offenders	Social casework and surveillance	90 juveniles who had at least one preceding adjudication for serious crimes	POSITIVE: Reduced the frequencies of rearrests but had no effects on the percentage of rearrests.
Walton, E., Fraser, M.W., Lewis, R.E., Pecora, P.J. and Walton, W.K. (1993). In-home family-focused reunification: an experimental study. *Child Welfare, 72*(5), 473-488.	Dosage and duration	Random assignment to: two groups with different dosages.	Social functioning/ Family reunification services	Home-based services	57 families.	POSITIVE: Had positive impacts on family reunification.
Wiebush, R. G. (1991). *Evaluation of the Lucas County Intensive Supervision Unit: Diversionary impact and youth outcomes, final report*. Toledo, OH: Lucas County Juvenile Court.	Dosage and duration	Random assignment to: experimental program with higher dosages or traditional aftercare program with lower dosages	Social functioning/ Aftercare probation for juvenile offenders	Social casework and surveillance	Incarceration-bound juveniles in Ohio.	NEUTRAL: No impact on recidivism.

Appendix B
Source Information from Chapter III (Non-Experimental Research)

J. Table B1

VI. Findings on Staff and Staffing from Non-Experimental Research

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
STAFF FORMAL EDUCATION					
Rosenthal, R., & Vandell, D.L. (1996). Quality of care at school-aged child-care programs: Regulatable features, observed experiences, child perspectives, and parent perspectives. Child Development, 67, 2434-2445.	Staff education	Of 70 potential centers that were listed in a "resource and referral directory," served the target population, and were willing to participate in the study, 30 programs were selected for the evaluation. In this group, "The number of programs from any one affiliation was limited to four... [and] programs were selected to reflect variations in director-reported curriculum focus, number of children attending, and program activities. " The selected group programs were larger, enrolled more children in the target age group, and had more teachers (though not a significantly different child-staff ratio), than those programs not selected for the study.. The 180 participants were selected from a group of 265 whose parents approved participation at one of the 30 day care sites (response rate of 57%). For each site, one boy-girl pair from the three target grade levels was randomly selected for the study (i.e., 6 children per center). "When it was not possible to obtain this distribution, other children were selected in an attempt to obtain similar distributions by grade and gender."	After-school programs	180 third- through fifth-grade children enrolled in after-school programs in or near Madison, WI. Grade distribution: 94 in 3rd, 55 in 4th, 21 in 5th. Mean age: 9.2 years. 57% male; 90% Caucasian.	Staff with less formal education were associated with more frequent negative staff-child interactions (.31+) (Unit of analysis was the program center (N = 30).

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Vandell, D.L., & Shumow, L. (1999). After-school child care programs. The Future of Children: When school is out, 9(2), 64-80.	Teacher education	Review of several large-scale studies on after-school programs.	Afterschool child care programs	Variety of programs and populations: Ecological Study of Afterschool Care, Study of After-School Care and Children's Development, Child Development Project, and Boston After-School Time Study. Among all the programs, 1038 children were involved.	"In programs with... less-educated teachers, the staff were more negative with children." "When staff members had higher levels of education, they were more likely to use positive behavior management strategies and were less likely to be harsh with their children."
STAFF TRAINING					
Eisen, M., Pallitto, C., Bradner, C., & Bolshun, N. (2000). Teen risk-taking: Promising programs and approaches. Washington, DC: Urban Institute.	Training/technical	Review of 51 "successful" prevention programs. Minimum criteria were: "treatment group is matched with a comparison… group; there is at least one follow-up review at three months after program completion; initial sample size consists of 100 adolescents or more; the study retains at least half of the participants at the final follow-up review; there is a statistically significant improvement for at least one target behavior in at least one target group." Secondary review of 21 high-quality evaluations. For inclusion in this analysis, additional criteria were: "Each program collected data about participants' problem behavior prior to the program beginning; the follow-up review period is longer (at least 12 months or the full school year); no fewer than 150 individuals are included in the treatment and comparison; the program retains more participants--- 67% in each group by the final follow-up date."	Prevention programs	Varied. Adolescents.	Common element of program success (one of six), culled from the review of higher-quality studies: "written curriculum and trainer feedback are provided" for leaders/presenters.

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Gottfredson, D.C., & Gottfredson, G.D. (2002). Quality of school-based prevention programs: Results from a national survey. Journal of Research in Crime and Delinquency, 39(1), 3-35.	Staff training	Data is based on the National Study of Delinquency Prevention in Schools. Principals in sampled schools answered a survey identifying activities that their school used to prevent or reduce delinquency, drug use, or other problem behavior or to promote a safe and orderly school environment. 17,110 prevention activities were identified by the principals in the screening stage. For each of the sample schools one activity for each of the 14 discretionary categories examined were randomly chosen. In all, 3,691 completed questionnaires were obtained from 554 schools which had an average return rate of 83% of their surveys. The survey measured such items as: program intensity, organizational capacity, organizational support, program structure, and integration into normal school operations.	School-based prevention programs	1, 287 schools were selected (143 schools for each cell in the sample design) from the 50 states and the District of Columbia. Public, private, and Catholic schools were selected and stratified by location and grade level.	Quality of program provider training was significantly correlated ($p < .01$) with proportion of best practices used for delivery method and with proportion of best practices used for program content (pp. 13, 25). The content-based best practices scale comprised four variables: employed prevention curriculum/instruction/training, employed behavioral programming/modification, made improvements to instructional practices, and employed classroom organization/management practices. The methods-based best practices scale comprised the same four variables as the content-based outcome, as well as two additional items: employed a counseling/social work/psychological or therapeutic activity; and employed mentoring, tutoring, coaching, or job apprenticeship/tutoring (p. 22).
Sipe, C. L. (1996). Mentoring: A synthesis of P/PV's research: 1988-1995. Philadelphia: Public/Private Ventures.	Training	Training and, especially experience in similar relationships, has been found to be positively related to positive mentoring relationships.	Mentoring programs	Mentoring pairs of adults and disadvantaged teenagers.	Observations based on findings from 10 reports from 1988 to 1995 by P/PV

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Wang, S. Keats, D.(2005) Developing an innovative cross-cultural strategy to promote HIV/AIDS prevention in different ethnic cultural groups of China. AIDS Care; Oct2005, Vol. 17 Issue 7, p874-891, 18p	Peer-Led Health Message	This study carried out in Sichuan Province is the first to explore how to use cultural resources for developing an effective strategy for promoting HIV prevention in different cultural groups in China. After training, these participants spread safe sex messages to other contacts who became an indirect peer diffusion group. A third group of 150 male volunteers from the same three cultural groups but from another relatively comparable community acted as controls. Each participant was interviewed before and after the intervention to assess knowledge, attitudes and behavioural intentions regarding HIV/AIDS prevention.	HIV Prevention	One hundred and fifty male volunteers drawn from the Yi (50), Tibetan (50) and majority Han (50) cultural groups were assigned to a direct training programme.	The study examined the cultural appropriateness and effectiveness of peer-led health message diffusion in promoting condom use through a traditional oral communication approach from the direct training groups to the indirect intervention groups and broad peer networks within the Yi, Tibetan and Han cultural communities. Key findings showed that the peer-based oral communication strategy was effective for encouraging condom use with casual sexual partners in both the direct training group and the indirect peer diffusion group in all three cultural groups. There was no significant change in any of the comparison groups. Although change in the majority Han cultural group was generally greater than in the ethnic minority groups, the results clearly suggest that the methods can be successfully adopted to promote safe sexual behaviour in different cultural groups of China.
STAFF-PARTICIPANT RATIO					
Rosenthal, R., & Vandell, D.L. (1996). Quality of care at school-aged child-care programs: Regulatable features, observed experiences, child perspectives, and parent perspectives. Child Development, 67, 2434-2445.	Child-staff ratio.	Of 70 potential centers that were listed in a "resource and referral directory," served the target population, and were willing to participate in the study, 30 programs were selected for the evaluation. In this group, "The number of programs from any one affiliation was limited to four... [and] programs were selected to reflect variations in director-reported curriculum focus, number of children attending, and program activities. " The selected group programs were larger, enrolled more children in the target age group, and had more teachers (though not a significantly different child-staff ratio), than those programs not selected for the study.. The 180 participants were selected from a group of 265 whose parents approved participation at one of the 30 day care sites (response rate of 57%). For each site, one boy-girl pair from the three target grade levels was randomly selected for the study (i.e., 6 children per center). "When it was not possible to obtain this distribution, other children were selected in an attempt to obtain similar distributions by grade and gender."	After-school programs	180 third- through fifth-grade children enrolled in after-school programs in or near Madison, WI. Grade distribution: 94 in 3rd, 55 in 4th, 21 in 5th. Mean age: 9.2 years. 57% male; 90% Caucasian.	Larger child-staff ratios were associated with more frequent negative staff-child interactions (.53**). (Unit of analysis was the program center (N = 30).

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Vandell, D.L., & Shumow, L. (1999). After-school child care programs. The Future of Children: When school is out, 9(2), 64-80.	Staff-child ratio	In programs with lower staff-child ratios, "staff were more negative with children." Programs with higher staff-child ratios were "more flexible, and staff appeared warmer, more sensitive, and more supportive of the children. Children in these programs spent less time waiting and in transition, less time watching television, and more time interacting positively with staff."	Afterschool child care programs	Variety of programs and populations: Ecological Study of Afterschool Care, Study of After-School Care and Children's Development, Child Development Project, and Boston After-School Time Study. Among all the programs, 1038 children were involved.	Review of several large-scale studies on after-school programs.
STAFF "MATCHING" PROGRAM POPULATION					
Blueprints for Violence Prevention. (2001B). Big Brothers/Big Sisters program. Boulder: Center for the Study and Prevention of Violence, University of Colorado.	Matching (race, gender)	"No significant differences were found in the rates of interaction (i.e., longevity of the BBBS match and rate of interaction between volunteers and youth) occurring in the same-race and cross-race matches."	Mentoring programs	Varies between program sites. Typically, youth, 6-18 years old, from single parent homes.	Overview of various program evaluations.

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Botvin, G. J., Baker, E., Filazzola, A. D., & Botvin, E. M. (1990). A cognitive-behavioral approach to substance abuse prevention: One-year follow-up. Addictive Behaviors, 15, 47-63.	Matching by age: Peer-led vs. teacher-led approach	Schools were randomly assigned to either "(a) prevention program implemented by older peer leaders, (b) prevention program implemented by regular classroom teachers, (c) prevention program with booster sessions implemented by older peer leaders, (d) prevention program with booster sessions implemented by regular classroom teachers," and control. All program leaders were trained using the same procedures.	Substance abuse prevention	Whole sample: "998 eighth-graders from 10 suburban New York junior high schools" were followed over two years. Sample was "predominately [80%] white and…. from middle-class families." 49% of the sample was male; and participants came from largely intact families (83%).	"The peer-led conditions continued to be more effective than the teacher led conditions, both in terms of substance use behavior and in terms of impacting on the cognitive, attitudinal , and personality mediating variables." Monthly, weekly, and daily cigarette smoking, monthly and weekly marijuana smoking, and indices of both behaviors were significant behavioral changes. However, qualitative information suggests that "teachers failed to implement the prevention program according to the protocol and that some teachers simply failed to implement substantial portions of the intervention altogether." Analyses with a restricted, high-fidelity sample of teachers revealed more positive effects.
Herrera, C., Sipe, C. L., McClanahan, W. S., Arbreton, A. J. A., & Pepper, S. K. (2000). Mentoring school-age children: Relationship development in community-based and school-based programs. Philadelphia: Public/Private Ventures.	Matching (race, gender)	Telephone interviews were conducted with mentors; "interviews and focus groups with youth, school and agency staff from eight exemplary programs." Mentor-reported measures.	Mentoring programs	669 mentors in community- and school-based programs serving school-aged children (this, a subset of the programs used in authors' 1997 surveys). Between the two types of programs, mentors were 33-40% male; 16-28% non-Caucasian; and of various ages.	"From the mentor's perspective, cross-ethnic matches are as close and supportive as same-ethnic matches. In addition, same-gender matches do not differ from cross-gender matches in closeness and supportiveness." Benchmark for most close and supportive relationships: similar interests
Jekielek, S., Moore, K. A., & Hair, E. C. (2002). Mentoring programs and youth development: A synthesis. Washington, DC: Child Trends.	Matching (race, gender)	Review of numerous experimental, quasi-experimental, and non-experimental analyses.	Mentoring programs	Varied. Most often, disadvantaged youth.	"Cross-race matches are just as successful as same-race matches for improving eleventh-grade GPA, college attendance, and college retention." (from Johnson, A. (1999). Sponsor-A-Scholar: Long-term impacts of a youth mentoring program on student performance. Princeton: Mathematica Policy Research, Inc.

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Jemmott, J. B., III, Jemmott, L. S., & Fong, G. T. (1992). Reductions in HIV risk-associated sexual behaviors among black male adolescents: Effects of an AIDS prevention intervention. American Journal of Public Health, 82(3), 372-377.	Matching (race, gender)	Random assignment to treatment (AIDS risk reduction) or control (career opportunities) groups. Facilitators were 11 Black women and 16 black men, all with a 4-year college degree, who were all trained for the program at the same time.	AIDS/HIV prevention	157 Black male adolescents (mean age: 14.64) "who were recruited from among outpatients at a medical clinic in West Philadelphia (44%); students attending the 10th, 11th, and 12th grade assemblies at a local high school (32%); and adolescents at a local YMCA (24%)."	"Results provide scant support for the view that matching the gender of facilitator and intervention recipient enhances the effectiveness of AIDS interventions with Black male adolescents." Comparisons: with male facilitators, participants experienced "greater increases in post-intervention AIDS knowledge... but this advantage of male facilitators was not evident on post-intervention attitudes or intentions, and it vanished at follow-up;" with female facilitators, participants experienced "less self-reported risky behavior and less positive attitudes toward such behavior" at a three-month follow-up.
Jemmott, J. B., III, Jemmott, L. S., Fong, G. T., & McCaffree, K. (1999). Reducing HIV risk-associated sexual behavior among African American adolescents: Testing the generality of intervention effects. American Journal of Community Psychology, 27(2), 161-187.	Matching (race, gender)	Participants were recruited from public schools via announcements… during 7th- and 8th-grade assemblies or lunch periods." The teens were paid $20 for per each of the three program steps: initial intervention, and follow-ups 3 and 6 months later. All facilitators were trained and repeatedly instructed to teach the program the same way in order to minimize "the effects of idiosyncrasies in the personalities of individual facilitators."	HIV prevention	"496 inner-city African American adolescents (mean age = 13 years)" from enrolled in school in Trenton, NJ. The sample was almost 54% female. Over 55% of the sample had had sex pre-intervention.	Null finding. "The effects of the HIV risk-reduction intervention did not vary as a function of the facilitator's race or gender, participant's gender, or the gender composition of the intervention group." The matching hypothesis was not supported, and "there was even a moderate negative correlation between the number of matching factors and participants' reactions."

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Jucovy, L. (2002). Same-race and cross-race matching (Technical Assistance Packet #7). Philadelphia: Public/Private Ventures.	Adult-youth matching	Reviewed a handful of studies that examined the different outcomes and quality of relationships between same- and cross-race matched mentoring pairs. Sample sizes were small, and the populations going into same- vs. cross-race matches may be substantially different.	Mentoring programs	Mentoring pairs in Big Brother/Big Sister programs. (Adults and disadvantaged teenagers.)	"Taken together, the research findings suggest that race, in itself, does not play a significant role in determining whether or not a mentor and mentee form a strong relationship and the extent to which that relationship leads to positive changes for the youth." Slight differences in specific measures were found between the groups in some programs, but these were thought to be potentially related to the ways in which youth in same- vs. cross-race matches differed. (e.g., "Parents or guardians who were most worried about their child might have agreed to a cross-race match because they wanted their child to have a mentor as quickly as possible...")
Lovato, C. Y., & Shoveller, J. A. (1999). Youth smoking cessation in school settings. In L. W. Green, C. J. Frankish, P. McGowan, P. Ratner, J. Bottorff, C. Y. Lovato, et al. (Eds.), Smoking cessation: A synthesis of the literature on program effectiveness. Vancouver: University of British Columbia Institute of Health Promotion Research.	Matching by age: Peer-led vs. adult-led leadership	Review of 9 studies, generated between 1983-1997 in the US, Canada, England or Australia, on school-based smoking-cessation programs. Some were published in peer-reviewed journals, others were unpublished; majority were quasi-experimental.	Smoking cessation programs	Varied. Programs designed for 12-18 year-olds enrolled in school.	"There does not appear to be a clear advantage to peer versus adult leadership; however, more research is needed in this area before any final conclusions are made." As evidenced in Prince, 1995, a program using TNT education materials saw no significant differences between outcomes of classes taught by trained classroom teachers or by trained college-age leaders.

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
McCurdy, K., Gannon, R. A., & Daro, D. (2003). Participation patterns in home-based family support programs: Ethnic variations. Family Relations, 52(1), 3-11.	Provider-participant matching	Data come from evaluations of the Healthy Families America (HFA) home-visiting programs. Programs studied in this paper (n = 17) had voluntary service enrollment, employed three concurrent home visitors, were able to serve at least 50 families, had been operating for at least three years, and had or would acquire within 8 months HFA credentials. The study was designed to collect data on families who "had the opportunity to receive at least 2 years of home visitation services" (p. 5).	Family support services (home-based)	670 African American, Latino, and European American mothers of newborns "who have been identified as at-risk for poor parenting through a structured interview with the Family Stress Checklist" (p. 5). 153 family support workers.	Ethnicity-based matching with a family support worker was associated with higher participation in a home-based family support program for African American mothers, but not for Latino or Caucasian mothers. Similarly, closeness in age between the mother and the provider was important for participation of Latino mothers only. Provider-mother similarity (in regard to race, age, education, marital status, residential location, etc.) was not associated with any differences in participation by Caucasian mothers. See study for additional findings.
Sipe, C. L. (1996). Mentoring: A synthesis of P/PV's research: 1988-1995. Philadelphia: Public/Private Ventures.	Matching (race, gender)	Observations based on findings from 10 reports from 1988 to 1995 by P/PV	Mentoring programs	Mentoring pairs of adults and disadvantaged teenagers.	"None of the 'objective' factors (e.g., age, race and gender)… correlates very strongly with either frequency of meeting, length of match or effectiveness." "Analyses of BB/BS data uncovered no differences in outcomes for youth involved in same-race versus cross-race relationships."
CAREGIVER-PARTICIPANT INTERACTION					
Arbreton, A.J.A., & McClanahan, W.S. (2002). Targeted outreach: Boys & Girls Clubs of America's approach to gang prevention and intervention. Philadelphia: Public/Private Ventures.	Staff-youth relationship.	Evaluation began in October 1997 with a focus on evaluating Club achievement of outreach, recruitment, and youth outcome goals. The youth questionnaire at baseline and 12 month follow-up focused on relationship behaviors, positive use of leisure time, school behaviors, delinquent behaviors, gang behaviors. The comparison group was selected to have similar demographic characteristics to the experimental group. The analysis took into account statistically the risk factors that might affect the positive or negative outcomes and examined the average effect on participants regardless of there amount of participation.	Program for prevention of gang involvement.	236 youth (avg. of 44 per site) at risk of gang membership. 48% aged 13 or older, 64% male, 51% African American, 29% Hispanic. Virtually all 'low income' (25% in public housing, 78% free/reduced-price lunch)	Frequency of participation was higher for youth in the prevention program who "report that the adults at the Club are supportive (.16*)... and that adults at the Club knew them well (.12+)." These factors did not predict the participation of youth in the intervention program (described elsewhere).

	Study Design		Program and Sample Description		Findings
Source	Manipulable Feature	Methodological Approach	Domain	Population	Differences in Effectiveness
Pierce, K.M., Hamm, J.V., & Vandell, D.L. (1999). Experiences in after-school programs and children's adjustment in first-grade classrooms. Child Development, 70(3), 756-767.	Staff positivity/negativity	Of 92 potential centers in the Madison, WI area, 38 programs were selected for the evaluation; this was based on some non-Caucasian enrollment, a minimum enrollment of three first-graders per program site, the selection of a roughly even distribution of for- and non-profit sites, and participation in/cooperation with the participant recruiting process. There was no difference in participant gender or minority status between participating and nonparticipating programs. Correlational study. Participants were selected "using a conditional random sampling strategy so that approximately one half were boys. All minority-race children and all children living in single-parent homes were selected. Other children who did not have these characteristics were selected randomly. "Families who chose to participate... had better educated fathers... and families of boys were more likely to refuse participation than families of girls."	After-school programs	150 first-grade children who attended after-school programs in or near Madison, WI, at least three days per week. Mean age: 6.5 years. 51% male; 87% Caucasian; 75% from two-parent families.	More positive staff regard was associated with fewer internalizing and externalizing behaviors in boys. "Negative staff regard was associated with negative academic impacts, resulting in significantly lower reading ($p<.01$) and mathematics ($p<.05$) grades. "Staff positivity... was associated with boys displaying fewer internalizing and externalizing problems, whereas staff negativity was related to boys obtaining poorer grades in reading and math."

K. Table B2

VII. Findings on Conceptual Approaches from Non-Experimental Research

	Study Design		Program and Sample Description		Findings
Source	Manipulable Feature	Methodological Approach	Domain	Population	Differences in Effectiveness
STRENGTH OF IMPLEMENTATION/FIDELITY TO PROGRAM DESIGN					

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Battistich, V., Schaps, E., Watson, M., & Solomon, D. (1996). Prevention effects of the Child Development Project: Early findings from an ongoing multisite demonstration trial. Journal of Adolescent Research, 11(1), 12-35.	Program implementation	All regular classroom teachers at each of the demonstration and comparison schools were assessed four times annually for 90 minutes per observation. Additionally an annual teacher questionnaire were collected. Students were assessed only in grades 4-6 and only when written parental consent was obtained. However, assessments of drug use and delinquent behaviors were limited to students in the fifth and sixth grade.	Child Development Project (classroom, school, and family involvement to make schools more hospitable to child development needs)	"Students and teachers [in all regular classrooms] at 24 elementary schools (12 demonstration and 12 comparison) from six school districts around the United States." Great variety in school's demographic make-up. Sample size for this study's analyses comprised only students from the top grade (5th or 6th) in each school (n = 1429-1745 over the course of the three academic-year evaluation). Male: 48%. Avg. age: 11.69.	"Use of marijuana also declined consistently from baseline among students at the high-implementation schools, bun increased at their matched comparison schools, $F(2,1368)= 8.56$, $p<.001$, resulting in a 9% greater prevalence rate among comparison than demonstration students in the second year of program implementation, $t(1454) =4.57$, $p<.001$, ES = .38. There were no reliable differences in trend for marijuana use in the analyses for the moderate- and low-implementation groups." Subgroup analyses showed there were several effects that only the high-implementation group experienced: "Differences in trend over time between program students in the high-implementation schools and their matched comparison students were found for carrying a weapon, $F(2, 1375)=7.28$, $p<.002$, vehicle theft, $F(2,1368)=7.11$, $p<.002$, skipping school, $F(2,1376)=2.41$, $p<.09$, and threatening someone with harm, $F(2,1366)=2.31$, $p<.10$." "By the second year of program implementation, students at the high-implementation schools showed reliably lower rates of skipping school, carrying weapons, and vehicle theft than did comparison students, ts $(1461) > 2.56$, $ps<.01$, ESs=.23-.31."

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Botvin, G. J., Baker, E., Filazzola, A. D., & Botvin, E. M. (1990). A cognitive-behavioral approach to substance abuse prevention: One-year follow-up. Addictive Behaviors, 15, 47-63.	Implementation fidelity	Schools were randomly assigned to either "(a) prevention program implemented by older peer leaders, (b) prevention program implemented by regular classroom teachers, (c) prevention program with booster sessions implemented by older peer leaders, (d) prevention program with booster sessions implemented by regular classroom teachers," and control. All program leaders were trained using the same procedures. Unit of analysis was the individual. For restricted analyses on classes in which teachers implemented the program with fidelity to protocol: 145 students.	Substance abuse prevention	Whole sample: "998 eighth-graders from 10 suburban New York junior high schools" were followed over two years. Sample was "predominately [80%] white and…. from middle-class families." 49% of the sample was male; and participants came from largely intact families *83%). Characteristics of this sample were very similar to the larger sample.	Overall, outcomes from teacher-led programs were non-significant and even negative. However, when limiting analyses to teachers who "were judged to have implemented the intervention with fidelity and completeness," negative effects disappeared and were replaced by positive effects for female participants.
Cave, G., & Quint, J. (1990). Career Beginnings impact evaluation: Findings from a program for disadvantaged high school students. New York: Manpower Demonstration Research Corporation.	Program implementation	Studied over two years, 1,574 students were initially randomly split into intervention and control groups; after two years and two follow-up interviews, 1,233 participants were still in the study. Sites were not randomly selected; chosen to be roughly representative of all 24 original site locations.	Career development program	1,233 (621 treatment) high school juniors in seven sites. 50% economically disadvantaged, 80% from families where neither parent had a college degree, 45% male, in the middle of their classes academically.	"Sites that were judged to have implemented the program most effectively produced the largest impacts, while sites judged the least successful at implementation had the smallest impacts." Activities, locations, and frequencies varied considerably between sites.

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Eisen, M., Pallitto, C., Bradner, C., & Bolshun, N. (2000). Teen risk-taking: Promising programs and approaches. Washington, DC: Urban Institute.	Approach: Theory-based	Review of 51 "successful" prevention programs. Minimum criteria were: "treatment group is matched with a comparison… group; there is at least one follow-up review at three months after program completion; initial sample size consists of 100 adolescents or more; the study retains at least half of the participants at the final follow-up review; there is a statistically significant improvement for at least one target behavior in at least one target group." Secondary review of 21 high-quality evaluations. For inclusion in this analysis, additional criteria were: "Each program collected data about participants' problem behavior prior to the program beginning; the follow-up review period is longer (at least 12 months or the full school year); no fewer than 150 individuals are included in the treatment and comparison; the program retains more participants--- 67% in each group by the final follow-up date."	Prevention programs	Varied. Adolescents.	Common element of program success (one of six), culled from the review of higher-quality studies: "All are theory-based."
Fagan, J., & Forst, M. (1996). Risks, fixers, and zeal. The Prison Journal, 76(16), 5-21.	Quality of implementation. (i.e., degree to which four different sites implemented the theory and program design.)	Experimental and control group youth and their contexts were studied through "participant observation, staff questionnaires, youth interviews [at several points throughout their involvement in the program], archival data," etc. Participants were followed over time. (Great diversity in the way programs achieved threshold of strong implementation…)	Intervention programs for chronically violent delinquents.	Chronically violent delinquents in Detroit, Memphis, Newark, and Boston juvenile justice systems.	"Recidivism rates were lower in the Detroit and Boston programs, those with the strongest overall implementation." In the Memphis program, recidivism rates were lower where "implementation of the experimental intervention was strongest," despite the site's control group have in stronger implementation of the theoretical and structural elements. The two sites that had the strongest overall implementation were the sites that had the lowest overall recidivism rates.
Harnett, P. & Dadds, M.R. (2004). Training school personnel to implement a universal school based prevention of depression program under real-world conditions. *Journal of School Psychology*. 42, 343-357.	Program Implementation	Conducted in two independent girls school in Brisbane, Australia. RAP Training Program Questionnaire, a 27-item self report questionnaire was used to evaluate the program Training was delivered to participants over a 1 day period that was essentially suppose to be a two day training session	Universal Prevention of Depression Program	A total of 212 female students from the two school. Participants were eight facilitators from School A: school psychologist and seven teachers (two male and six female)	Assumption is that a 1 day training program successfully facilitate the acquisition of relevant knowledge is important. Adherence to the program varied. Deviation does not distinguish between flexibility in curriculum delivery or deviations due to incorrect administration. Relationship between program implementation and student outcomes: students exposed to a higher percentage of key concepts responded better to the intervention as compared to students who were exposed to fewer. Maximizing features: Training and Supervision pg. 354

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Henggeler, S. W., Melton, G. B., Brondino, M. J., Scherer, D. G., & Hanley, J. H. (1997). Multisystemic therapy with violent and chronic juvenile offenders and their families: The role of treatment fidelity in successful dissemination. Journal of Consulting and Clinical Psychology, 65(5), 821-833.	Fidelity to program design.	Youths and primary caregivers were randomly assigned to treatment or control groups. Data collection at T1 and 1.7 years later. Various established measures were employed to gauge effects.	Multisystemic therapy for violent/chronic juvenile offenders	155 violent/chronic juvenile offenders (11-17 years old) and their families	"Outcomes were substantially better in cases where treatment adherence ratings were high." "High adolescent reports of [self-reported delinquency] index offenses were significantly associated with low therapist adherence to MST principles, $F(1, 65) = 5.40$, $p < .03$… On the basis of parental reports of MST fidelity, high rates of re-arrest after T1 were significantly associated with low therapist adherence to MST principles, $F(1, 51) = 9.55$, $p < .004$."
Kallestad, J. & Olweusa, D. (2003). Predicting teachers' and schools' implementation of the olweus bullying prevention program: a multilevel study. *Prevention and Treatment*. 6(21).	Program Implementation	Analyzed 89 teacher surveys from 37 Schools Provided data at two different time points, separated by 6 months. All the teachers used varying degrees, the Olweus Bullying Prevention Program The sampling procedure was based on classes/schools in a quasi-experimental design with four different cohorts of students who were in Grades 6-9 in the presented analyses Only teachers with valid data were included in the multilevel analyses.	Bullying/Victim problems in schools	Teachers over 6-9 grade students	83% of the teachers used at least one specific measure proposed in the Olweus Bullying Prevention Program --49% use of literature --11% role-playing --6% directly involved parents Almost 90% of teachers had at least one positive score on Classroom Intervention Measures, and as much as 83% had at least one positive score on Classroom Intervention Measures
Kelly, Cherly M., Baker, Elizabeth A., Williams, Deidre, Nanney, M.S., Haire-Joshu, Debra. (2004). Organizational Capacity's Effects on the Delivery and Outcomes of Health Education Programs. Journal of Public Health Management Practice 10(2), 164-170	Organizational Capacity of program	6 sites were identified to measure the factors influencing implementation of this dietary change program. High needs families were selected to participate in the dietary change program if they met two of the following criteria: single parental low income; minority ethnicity; or living with stressors such as illnesses or recent divorce or death in the family. Additionally, the number of modules delivered to families by PEs were tracked through out the intervention. Lastly, several organizational variables, including both structural and process dimensions of organizational capacity, were measured, both quantitatively and qualitatively, through the course of the intervention period.	Dietary change program	Parent and Teacher program directors High Needs Families	Organizational structural and Process Capacities affect program implementation which in turn affect health outcomes. Reasonable caseload experience, training, and education of program and teacher program directors substantially impact the eating behavior or high needs families.

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Kirby, D. (1999). Reflections on two decades of research on teen sexual behavior and pregnancy. Journal of School Health, 69(3), 89-94.	Theoretical basis	Observations based on research from the last 20 years	Pregnancy prevention programs	Varied older children, adolescents and/or youth.	Effective curricula "were based upon theoretical approaches that have been demonstrated to be effective in influencing other health-related risky behaviors…"
Lipsey, M. W. (1992). Juvenile delinquency treatment: A meta-analytic inquiry into the variability of effects. In T. Cook, H. Cooper, D.S. Cordray, H. Hartmann, L.V. Hedges, R.J. Light, et al, Eds., *Meta-analysis for explanation: A casebook* (p. 83-127). New York: Russell Sage Foundation.	Significant researcher involvement/influence in program delivery. [Tied to design fidelity]	Meta-analyses of experimental and quasi-experimental studies (N = 443) conducted between 1950 and 1987. One of the criteria for the inclusion of studies was that the study had to be experimental based on random assignments or if it was quasi-experimental, the study had to provide: 1) pre- and post-outcome measures, and 2) "some evidence of matching between the two groups prior to treatment" (or measures that show the similarity of the two groups prior to treatment). The study used "a weighted (stepped) multiple regression in which the contribution of each case (study) to the analysis is weighted by the inverse variance of the effect size" in order to adjust for different sample sizes. All the measures related to methodologies of the studies were "stepped into" the regressions prior to treatment measures.	Juvenile delinquency treatment.	Varied. The sample size ranged from 25 or fewer to 801 and more participants.	Larger effect size for reducing re-conviction/reconviction. **SEE Experimental write-up on this**
Mihalic, S., Fagan, A., Irwin, K., Ballard, D., & Elliott, D. (2002). Blueprints for violence prevention replications: Factors for implementation success. Boulder: Center for the Study and Prevention of Violence, University of Colorado.	Fidelity to program design.	Review of successful anti-violence programs.	Anti-violence programs	Varied.	Lots of information on positive outcomes w/program design fidelity (p2.12) Also info on positive outcomes with "best practices" (p2.5).

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Supovitz, J., and Henry M. (2004). A study of the links between implementation and effectiveness of America's choice comprehensive school reform design.. *Journal of Education for Students Placed at Risk* 9(4), 389-419.	Implementation and Student Learning	Data was collected as part of the Consortium for Policy Research in Education's (CPRE) national evaluation of America's Choice. Teachers in all of the America's Choice schools in the district were asked to complete a survey Students were given an achievement test Students whose teachers did not complete a survey were dropped from the analyses, as were teachers without valid student information.	Comprehensive school-based instructional program	Teachers 114 General Elementary 1572 Students from 1st thru 6th grade in the 10 America's Choice Schools	Explore what teacher characteristics were associated with teacher implementation of America's Choice. Identifying where teachers with certain characteristics were more or less likely to implement America's Choice. Characteristics of Teachers • Gender • Ethnicity • Subject-certified teachers Classroom Characteristics and Teacher's Overall Implementation • Classroom teachers • Class size Implementation and Teacher Attitudes • All students can learn • Same Standards should be applied to all students
PARTICIPANT-CENTERED APPROACH					
Herrera, C., Sipe, C. L., McClanahan, W. S., Arbreton, A. J. A., & Pepper, S. K. (2000). Mentoring school-age children: Relationship development in community-based and school-based programs. Philadelphia: Public/Private Ventures.	Youth-driven approach	Telephone interviews were conducted with mentors; "interviews and focus groups with youth, school and agency staff from eight exemplary programs." Mentor-reported measures.	Mentoring programs	669 mentors in community- and school-based programs serving school-aged children (this, a subset of the programs used in authors' 1997 surveys). Between the two types of programs, mentors were 33-40% male; 16-28% non-Caucasian; and of various ages.	"mentors reported closer relationships and more supportive relationships both emotionally and instrumentally when decisions were made together. The least positive relationships resulted when decisions about activities were made primarily by the mentor or established in advance by the program." Benchmark for most close and supportive relationships: "getting ideas from youth and then deciding together."

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
McGill, D. E., Mihalic, S., & Grotpeter, J. K. (2001). Blueprints for violence prevention: Big Brothers/Big Sisters program. Boulder: Center for the Study and Prevention of Violence, University of Colorado.	Youth-driven vs. adult-driven approach. (Prescriptive vs. Developmental approach)	Overview of various program evaluations.	Mentoring programs	Varies between program sites. Typically, youth, 6-18 years old, from single parent homes.	"The majority of prescriptive matches faltered or closed (22 of 28), most developmental matches (50 of 54) persisted and continued to develop.
Sipe, C. L. (1996). Mentoring: A synthesis of P/PV's research: 1988-1995. Philadelphia: Public/Private Ventures.	Youth-driven approach	Observations based on findings from 10 reports from 1988 to 1995 by P/PV	Mentoring programs	Mentoring pairs of adults and disadvantaged teenagers.	This approach/characteristic (e.g., in which the pair decides together how they spend their time) was associated with more positive relationships between the mentor and youth. Adults who have "unrealistic expectations or reforming the youth will inevitably become frustrated and disappointed when these expectations are not met" (e.g., "reforming" the youth, achieving behavior changes, etc.).
DEGREE OF STRUCTURE vs. FLEXIBILITY					
Dungan-Seaver, D. (1999). Afterschool programs: An analysis of research about characteristics of effectiveness. Produced for the McKnight Foundation. Retrieved March 16, 2003, from the World Wide Web: http://www.mcknight.org/display_file.asp?FileID=28	Flexibility and structure	Review of several after-school studies.	Afterschool programs	Variety of programs and populations.	Great debate of this, but "studies seem to conclude… that structure should mean identifiable activities, progress and boundaries, but should not mean regimented participation or outcomes." Also, "younger children seem to benefit from more, older children and adolescents from less.

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Fashola, O. S. (1998). Review of extended-day and after-school programs and their effectiveness (Report No. 24). Washington, DC: Center for Research on the Education of Students Placed at Risk.	Flexibility of program	Review of 34 programs "with an educational focus that have been shown to have evidence of effectiveness for all children," and "that have been used as after-school programs by schools and/or communities, including extended day programs and some supplemental school programs that have potential after-school usage." Correlational research studies also discussed. Effect sizes were calculated across programs.	Extended school-day and after-school programs	Varied. School-age.	Social skills were higher for those involved in more flexible programs ($p<.05$); however, the written language grades were poorer in this group ($p<.05$).
Pierce, K.M., Hamm, J.V., & Vandell, D.L. (1999). Experiences in after-school programs and children's adjustment in first-grade classrooms. Child Development, 70(3), 756-767.	Program flexibility	Of 92 potential centers in the Madison, WI area, 38 programs were selected for the evaluation; this was based on some non-Caucasian enrollment, a minimum enrollment of three first-graders per program site, the selection of a roughly even distribution of for- and non-profit sites, and participation in/cooperation with the participant recruiting process. There was no difference in participant gender or minority status between participating and nonparticipating programs. Correlational study.	After-school programs	150 first-grade children who attended after-school programs in or near Madison, WI, at least three days per week were selected "using a conditional random sampling strategy so that approximately one half were boys. All minority children and all children living in single-parent homes were selected." Mean age: 6.5 years. 51% male; 87% Caucasian; 75% from two-parent families.	"Program flexibility was associated with boys having better social skills" (.29*). (Regressions control for family structure.)
CLEAR & CONSISTENT MESSAGE					

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Eisen, M., Pallitto, C., Bradner, C., & Bolshun, N. (2000). Teen risk-taking: Promising programs and approaches. Washington, DC: Urban Institute.	Approach: Specific behavior goals are targeted	Review of 51 "successful" prevention programs. Minimum criteria were: "treatment group is matched with a comparison… group; there is at least one follow-up review at three months after program completion; initial sample size consists of 100 adolescents or more; the study retains at least half of the participants at the final follow-up review; there is a statistically significant improvement for at least one target behavior in at least one target group." Secondary review of 21 high-quality evaluations. For inclusion in this analysis, additional criteria were: "Each program collected data about participants' problem behavior prior to the program beginning; the follow-up review period is longer (at least 12 months or the full school year); no fewer than 150 individuals are included in the treatment and comparison; the program retains more participants--- 67% in each group by the final follow-up date."	Prevention programs	Varied. Adolescents.	Common element of program success (one of six), culled from the review of higher-quality studies: "clearly delineated and articulated goals for behavior change."
Fagan, J., & Forst, M. (1996). Risks, fixers, and zeal. The Prison Journal, 76(16), 5-21.	Clear and consistent message	Experimental and control group youth and their contexts were studied through "participant observation, staff questionnaires, youth interviews [at several points throughout their involvement in the program], archival data," etc. Participants were followed over time.	Intervention programs for chronically violent delinquents.	Chronically violent delinquents in Detroit, Memphis, Newark, and Boston juvenile justice systems.	"These programs were notable for clear rules and contingencies, strong modeling of prosocial behaviors…"
Kirby, D. (1999). Reflections on two decades of research on teen sexual behavior and pregnancy. Journal of School Health, 69(3), 89-94.	Clear message	Observations based on a review of research from the last 20 years	Pregnancy prevention programs	Varied older children, adolescents and/or youth.	Effective curricula "gave a clear message by continually reinforcing a clear stance on these behaviors."
MULTI-COMPONENT, WHOLE-PERSON, SOCIAL SKILLS-BUILDING APPROACH					

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Tobler, M. S., Roona, M. R., Ochshorn, P., Marshall, D. G., Streke, A. V., & Stackpole, K. M. (2000). School-based adolescent drug prevention programs: 1998 meta-analysis. Journal of Primary Prevention, 20(4), 275-337.	Multi-component, whole youth approach	Meta-analysis of 207 school-based drug prevention programs whose data were reported between 1978 and 1998. Programs were categorized into Interactive and Non-Interactive groups based on content and delivery method. Weighted categorical and regression methods were used to determine the attributes that "most effectively reduce, delay, or prevent drug use, including program size, type of control group and leader, attrition, target drug, intensity, grade, special population and level of drug use".	School based drug prevention programs	Varied. Adolescents.	Interactive programs, that enhance the development of interpersonal skills, have greater impacts (WES=.15) than non-interactive lecture-oriented programs, which have only minimal impacts (weighted effect size or WES=.05). The largest gains were made in comprehensive life skills programs (WES=.17) and by efforts to implement systematic change (WES=.27). Life skill programs included training in refusal skills, goal setting, assertiveness, communication, and coping. Systematic changes involved school-based programs with community supports, media and family programming or school-wide restructuring efforts which emphasize bonding between students and the school, cooperative learning in small groups, and school-family communication.
Eisen, M., Pallitto, C., Bradner, C., & Bolshun, N. (2000). Teen risk-taking: Promising programs and approaches. Washington, DC: Urban Institute.	Approach: Multi-component.	Review of 51 "successful" prevention programs. Minimum criteria were: "treatment group is matched with a comparison… group; there is at least one follow-up review at three months after program completion; initial sample size consists of 100 adolescents or more; the study retains at least half of the participants at the final follow-up review; there is a statistically significant improvement for at least one target behavior in at least one target group." Secondary review of 21 high-quality evaluations. For inclusion in this analysis, additional criteria were: "Each program collected data about participants' problem behavior prior to the program beginning; the follow-up review period is longer (at least 12 months or the full school year); no fewer than 150 individuals are included in the treatment and comparison; the program retains more participants--- 67% in each group by the final follow-up date."	Prevention programs	Varied. Adolescents.	Common element of program success (one of six), culled from the review of higher-quality studies: "multiple-component interventions are especially promising." Many programs use a variety of techniques and delivery mechanisms… Eight programs involve the community in some capacity, and seven programs involve parents. Several include a strong peer education or support component. About eight recruit either same-age or older-age peer leaders."

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Eisen, M., Pallitto, C., Bradner, C., & Bolshun, N. (2000). Teen risk-taking: Promising programs and approaches. Washington, DC: Urban Institute.	Social skills-building	Review of 51 "successful" prevention programs. Minimum criteria were: "treatment group is matched with a comparison… group; there is at least one follow-up review at three months after program completion; initial sample size consists of 100 adolescents or more; the study retains at least half of the participants at the final follow-up review; there is a statistically significant improvement for at least one target behavior in at least one target group." Secondary review of 21 high-quality evaluations. For inclusion in this analysis, additional criteria were: "Each program collected data about participants' problem behavior prior to the program beginning; the follow-up review period is longer (at least 12 months or the full school year); no fewer than 150 individuals are included in the treatment and comparison; the program retains more participants--- 67% in each group by the final follow-up date."	Prevention programs	Varied. Adolescents.	Common element of program success (one of six), culled from the review of higher-quality studies: "interactive student-to-student and student-to-instructor skill-building methods--- including role-playing and rehearsal, guided practice, and immediate feedback--- to address the target problem behavior." Communication issues, behavior modeling, social influences, assertiveness skills, problem-solving, decision-making also cited.
Kirby, D. (1999). Reflections on two decades of research on teen sexual behavior and pregnancy. Journal of School Health, 69(3), 89-94.	Social skills building	Observations based on a review of research from the last 20 years	Pregnancy prevention programs	Varied older children, adolescents and/or youth.	Effective curricula "included activities that address social pressures on sexual behaviors," and "provided modeling and practice of communication, negotiation, and refusal skills."
Kirby, D. (2002). Effective approaches to reducing adolescent unprotected sex, pregnancy, and childbearing. Journal of Sex Research, 39(1), 51-57.	Comprehensiveness of approach	"Review of 73 studies and their respective programs"	Sexual health programs	Varied. Adolescents.	"It seems likely that programs that address both sexual and nonsexual antecedents effectively will be more effective than those programs that address only one group of antecedents."

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Tobler, M. S. (1992). Drug prevention programs can work: Research findings. Journal of Addictive Diseases, 11(3), 1-28.	Conceptual program approach	Meta-analyses of 91 drug prevention programs that measured, specifically, change in drug use (self-reported and/or observed). Only one effect size was computed per program. Selection criteria: "(1) Quantitative outcome measures, (2) a control or comparison group, (3) grades six through twelve, (4) prevention as a goal."	Drug prevention programs	Pre-teens and teenagers in grades 6-12.	Peer programs increased in their effect size using this new computing method, compared to their already high effect size (greatest among the 5 approaches) from the 1986 meta-analysis. Average effect sizes by type of approach: "knowledge-only [e.g., scare tactics] (ES = 0.07), affective-only [focused on psychological risk factors, values, no specific reference to drugs] (ES = 0.05), peer programs (ES = 0.42), knowledge-plus-affective (ES = 0.07), and alternatives (ES = 0.20)." (The latter was designed to have a greater number of program hours than the other programs and was usually administered to high-risk groups.) "The success of the peer program is not dependent on its leader" (see below). Peer programs involved a knowledge component, and "a group situation that promotes peer support for not using drugs... Peer interaction [between students] is the key component of these groups."
Information-only vs. other programs					
Kirby, D. (1999). Reflections on two decades of research on teen sexual behavior and pregnancy. Journal of School Health, 69(3), 89-94.	Information-only approaches	Observations based on a review of research from the last 20 years	Pregnancy prevention programs	Varied older children, adolescents and/or youth.	"Knowledge level is only weakly related to behavior, and that programs that focus on knowledge acquisition do increase student knowledge, but they do not significantly change sexual or contraceptive behavior." Effective curricula "employed a variety of teaching methods designed to involve the participants and have them personalize the information."

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Tobler, M. S. (1992). Drug prevention programs can work: Research findings. Journal of Addictive Diseases, 11(3), 1-28.	Conceptual program approach	Meta-analyses of 91 drug prevention programs that measured, specifically, change in drug use (self-reported and/or observed). Only one effect size was computed per program. Selection criteria: "(1) Quantitative outcome measures, (2) a control or comparison group, (3) grades six through twelve, (4) prevention as a goal."	Drug prevention programs	Pre-teens and teenagers in grades 6-12.	Peer programs increased in their effect size using this new computing method, compared to their already high effect size (greatest among the 5 approaches) from the 1986 meta-analysis. Average effect sizes by type of approach: "knowledge-only [e.g., scare tactics] (ES = 0.07), affective-only [focused on psychological risk factors, values, no specific reference to drugs] (ES = 0.05), peer programs (ES = 0.42), knowledge-plus-affective (ES = 0.07), and alternatives (ES = 0.20)." (The latter was designed to have a greater number of program hours than the other programs and was usually administered to high-risk groups.) "The success of the peer program is not dependent on its leader" (see below). Peer programs involved a knowledge component, and "a group situation that promotes peer support for not using drugs... Peer interaction [between students] is the key component of these groups."
Tobler, M. S., Roona, M. R., Ochshorn, P., Marshall, D. G., Streke, A. V., & Stackpole, K. M. (2000). School-based adolescent drug prevention programs: 1998 meta-analysis. Journal of Primary Prevention, 20(4), 275-337.	Information-only approaches	Meta-analysis of 207 school-based drug prevention programs whose data were reported between 1978 and 1998. All studies met the following criteria: "(a) evaluated a school-based drug prevention program available to all members of the student body, which may have included but did not specially target 'high risk youth;' (b) employed quantitative drug use measures; (c) used a control or comparison group...; (d) involved school grades six through twelve (plus a small number of K-6 programs...); (e) had goals of primary prevention and/or secondary prevention and/or early intervention, not targeting identified abusive or addicted drug users in treatment; (f) included universal participation of all ethnic groups that comprise the school's population; and (g) been reported or published from 1978-1998" (pp. 276-277). "All program outcome statistics have been converted to standard effect sizes (ES)" (p. 278). A program is the unit of analysis. Interactive approaches were social influences, comprehensive life skills, and system-wide change.	School based drug prevention programs	Varied. Mostly children and teens in grades 6-12.	Non-interactive lecture-oriented programs to have "minimal impact," whereas the "interactive programs that enhance the development of interpersonal skills have greater impact" When examined by specific prevention focus, programs with interactive approaches were significantly more effective than those with non-interactive approaches for tobacco and general substance use. There was no significant difference between the two approaches for programs that targeted alcohol use, which elicited smaller overall effects than the tobacco and general substance use programs. Regardless of the number of participants in the program, or whether it targeted "special populations" (see report) programs with interactive programs were associated with significantly greater effectiveness than those with non-interactive approaches (p. 289). The same pattern was found regardless of degree of program attrition or the quality of research design, as well.

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Tobler, N. (1986).Meta-analysis of 143 adolescent drug prevention programs: Quantitative outcome results of program participants compared to a control or comparison group. Journal of Drug Issues, 16, 537-567, as discussed in: Tobler, M. S. (1992). Drug prevention programs can work: Research findings. Journal of Addictive Diseases, 11(3), 1-28.	Conceptual program approach	Meta-analyses of 146 drug prevention programs, from 1986 article. Selection criteria: "(1) Quantitative outcome measures, (2) a control or comparison group, (3) grades six through twelve, (4) prevention as a goal." Effects were computed using, potentially, more than one effect size for a program.	Drug prevention programs	Pre-teens and teenagers in grades 6-12.	"Knowledge-only programs showed success on knowledge outcomes (ES = 0.59), but negligible change for attitudes (ES = 0.05) and drug use (ES = 0.03). The affective-only programs were ineffective across all outcome measures." The range of for affective affect was ES = 0.07 - 0.12. "The results for the third strategy, knowledge-plus-affective, were knowledge (ES = 0.47), attitudes (ES = 0.18), drug use (ES = 0.15), skills (ES = 0.07), and behavior (ES = 0.25)."
"SHOCK" APPROACHES					
Lipsey, M. W. (1992). Juvenile delinquency treatment: A meta-analytic inquiry into the variability of effects. In T. Cook, H. Cooper, D.S. Cordray, H. Hartmann, L.V. Hedges, R.J. Light, et al, Eds., *Meta-analysis for explanation: A casebook* (p. 83-127). New York: Russell Sage Foundation.	"Deterrence treatments" (scared straight, shock incarceration, etc.)	Meta-analyses of experimental and quasi-experimental studies (N = 443) conducted between 1950 and 1987. One of the criteria for the inclusion of studies was that the study had to be experimental based on random assignments or if it was quasi-experimental, the study had to provide: 1) pre- and post-outcome measures, and 2) "some evidence of matching between the two groups prior to treatment" (or measures that show the similarity of the two groups prior to treatment). The study used "a weighted (stepped) multiple regression in which the contribution of each case (study) to the analysis is weighted by the inverse variance of the effect size" in order to adjust for different sample sizes. All the measures related to methodologies of the studies were "stepped into" the regressions prior to treatment measures.	Juvenile delinquency treatment.	Varied. The sample size ranged from 25 or fewer to 801 and more participants.	The study notes that "more structured and focused treatments (e.g., behavioral, skill-oriented) and multimodal treatments seem to be more effective than less structured and focused approaches (e.g., counseling). Some of the approaches, particularly "deterrence treatments" which included "shock incarceration and the "scared straight" program model" produced negative effects. (Effects were in regard to re-arrest or re-conviction)

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Sherman, L. W., Gottfredson, D. C., MacKenzie, D., Eck, J., Reuter, P., & Bushway, S. (1998). Preventing crime: What works, what doesn't, what's promising. Washington, DC: Department of Justice, National Institute of Justice.	Shock approaches	Review of state and local crime prevention programs	Crime prevention programs	Varies. Includes programs affecting individuals across the lifespan and in diverse communities.	In regard to programs administered to minor juvenile offenders after arrest, "scared straight" programs "fail to reduce the participants' reoffending rates and may increase crime" (p. 9; Research Brief). Scared straight programs described as "bringing... offenders to visit maximum security prisons to see the severity of prison conditions" (p. 9; Research Brief). Shock probation and shock parole programs, for which the intent is that "a short period of time incarcerated would 'shock' offenders into abandoning criminal activity," have also been found to be largely unsuccessful and in some cases linked to negative outcomes (Section 3.2, Full Report).

L. Table B3

VIII. Findings on Program Practices from Non-Experimental Research

	1. Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
GROUP SIZE					

	1. Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Jastrzab, J., Blomquist, J., Masker, J., & Orr, L. (1997). Youth Corps: Promising strategies for young people and their communities (Studies in Workforce Development and Income Security, No. 1-97). Cambridge, MA: Abt Associates, Inc.	Group size	Comprehensive national study of Youth Corps programs. Quantitative data are based on the data provided to the national Evaluation Information System during their 1993/1994 grant cycle for the Commission on National and Community Service.	Youth corps community service programs	Varies; typically, participants are "educationally or economically disadvantaged young people between the ages of 19 and 25." Participation currently comprises 26,000 youth, and typically lasts 4-5 months; program size ranges from 20 to several hundred. In-depth analysis conducted at 8 sites.	"In general, larger programs (50 or more participants) tended to have lower completion rates than their smaller counterparts."
Rosenthal, R., & Vandell, D.L. (1996). Quality of care at school-aged child-care programs: Regulatable features, observed experiences, child perspectives, and parent perspectives. Child Development, 67, 2434-2445.	Size of enrollment	Of 70 potential centers that were listed in a "resource and referral directory," served the target population, and were willing to participate in the study, 30 programs were selected for the evaluation. In this group, "The number of programs from any one affiliation was limited to four... [and] programs were selected to reflect variations in director-reported curriculum focus, number of children attending, and program activities. " The selected group programs were larger, enrolled more children in the target age group, and had more teachers (though not a significantly different child-staff ratio), than those programs not selected for the study..	After-school programs	180 third- through fifth-grade children enrolled in after-school programs in or near Madison, WI. Grade distribution: 94 in 3rd, 55 in 4th, 21 in 5th. Mean age: 9.2 years. 57% male; 90% Caucasian. These participants were selected from a group of 265 whose parents approved participation at one of the 30 day care sites (response rate of 57%).	Participants reported "poorer program climate" with larger enrollment (-.26**). For each site, one boy-girl pair from the three target grade levels was randomly selected for the study (i.e., 6 children per center). "When it was not possible to obtain this distribution, other children were selected in an attempt to obtain similar distributions by grade and gender."
ACTIVITIES: Diversity of					

	1. Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Herrera, C., & Arbreton, A. J. A. (2003). Increasing opportunities for older youth in after-school programs: A report on the experiences of Boys & Girls clubs in Boston and New York City. Philadelphia: Public/Private Ventures.	Participation in a diverse variety of activities	Staff-administered (though confidential) annual youth survey of a sample of the club's participants, over a three-year period. (n=645 at T1, 583 at T3) The sample was comprised of groups (gender, membership status, etc.) specifically targeted by P/PV.	Boys/Girls Club for high- and higher-risk youth.	12-18 year old youth in five Boston and three New York City clubs.	"Diversity of activity participation was the single most important variable in predicting [reports of adult and peer support, leadership experience, decision-making opportunities, interesting activities at the club, and academic and job-related benefits] from club participation--- even more important than frequency of attendance of length of membership." (Potential that this is related to other factors?: More kids in Boston than in NY participated in a wide variety (58% vs. 45%), reflecting (according to the authors) Boston's greater emphasis on teen programming, required participation in certain activities, and possibly more encouragement to participate.)
Pierce, K.M., Hamm, J.V., & Vandell, D.L. (1999). Experiences in after-school programs and children's adjustment in first-grade classrooms. Child Development, 70(3), 756-767.	Number and variety of program activities	Of 92 potential centers in the Madison, WI area, 38 programs were selected for the evaluation; this was based on some non-Caucasian enrollment, a minimum enrollment of three first-graders per program site, the selection of a roughly even distribution of for- and non-profit sites, and participation in/cooperation with the participant recruiting process. There was no difference in participant gender or minority status between participating and nonparticipating programs. Correlational study.	After-school programs	150 first-grade children who attended after-school programs in or near Madison, WI, at least three days per week were selected "using a conditional random sampling strategy so that approximately one half were boys. All minority-race children and all children living in single-parent homes were selected." Mean age: 6.5 years. 51% male; 87% Caucasian; 75% from two-parent families.	"Boys who attended programs offering a larger number of different activities had more internalizing [.39**] and externalizing [.26*] problems, and poorer grades in reading [-.40**] and math [-.37**]." (Regressions control for family structure.) It should also be noted that, "families who chose to participate... had better educated fathers... and families of boys were more likely to refuse participation than families of girls."

	1. Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Rosenthal, R., & Vandell, D.L. (1996). Quality of care at school-aged child-care programs: Regulatable features, observed experiences, child perspectives, and parent perspectives. Child Development, 67, 2434-2445.	Variety of program activities	Of 70 potential centers that were listed in a "resource and referral directory," served the target population, and were willing to participate in the study, 30 programs were selected for the evaluation. In this group, "The number of programs from any one affiliation was limited to four... [and] programs were selected to reflect variations in director-reported curriculum focus, number of children attending, and program activities. " The selected group programs were larger, enrolled more children in the target age group, and had more teachers (though not a significantly different child-staff ratio), than those programs not selected for the study..	After-school programs	180 third- through fifth-grade children enrolled in after-school programs in or near Madison, WI. Grade distribution: 94 in 3rd, 55 in 4th, 21 in 5th. Mean age: 9.2 years. 57% male; 90% Caucasian. Participants were selected from a group of 265 whose parents approved participation at one of the 30 day care sites (response rate of 57%).	"The presence of a greater number of different types of program activities was associated with staff having more frequent positive or neutral interactions with children" (.47**). (Unit of analysis was the program center (N = 30). Participants reported better overall climate (.17+) and more emotional support (.19*) with a greater variety of activities offered. For each site, one boy-girl pair from the three target grade levels was randomly selected for the study (i.e., 6 children per center). "When it was not possible to obtain this distribution, other children were selected in an attempt to obtain similar distributions by grade and gender."
ACTIVITIES: Interesting, enjoyable					
Arbreton, A.J.A., & McClanahan, W.S. (2002). Targeted outreach: Boys & Girls Clubs of America's approach to gang prevention and intervention. Philadelphia: Public/Private Ventures.	Interesting activities	The report is based on two different studies. One program is to prevent at risk youth from joining gangs (discussed here), the other is a program to intervene with youth already involved in gangs. Neither background or socio-demographic characteristics, nor referral source, predicted level of participation. The gang prevention program used a Gang Risk Factors scale to determine if a youth was at high risk of gang involvement. Points were assigned for such behaviors as "exhibiting gang signs and symbols", coming from a highly distressed or crisis ridden family, of having family members or friends who are gang members. To evaluate change a sample or youth answered a questionnaire when they were recruited and again approximately 12 months later. Results were compared to a group of youth who did not attend Clubs.	Program for prevention of gang involvement.	236 youth (avg. of 44 per site) at risk of gang membership. 48% aged 13 or older, 64% male, 51% African American, 29% Hispanic. Virtually all 'low income' (25% in public housing, 78% free/reduced-price lunch)	Frequency of participation was higher for youth in the prevention program who report "that the activities are interesting (.23**)." This factor did not predict the participation of youth in the intervention program (described elsewhere).

	1. Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Herrera, C., & Arbreton, A. J. A. (2003). Increasing opportunities for older youth in after-school programs: A report on the experiences of Boys & Girls clubs in Boston and New York City. Philadelphia: Public/Private Ventures.	Degree to which club programs were considered interesting.	Staff-administered (though confidential) annual youth survey of a sample of the club's participants, over a three-year period. (n=645 at T1, 583 at T3) The sample was comprised of groups (gender, membership status, etc.) specifically targeted by P/PV.	Boys/Girls Club for high- and higher-risk youth.	12-18 year old youth in five Boston and three New York City clubs.	MIXED FINDING: In Boston, "youth's attendance was only related to how interesting they found activities." In New York, attendance was not related to degree of interest in club activities. (p25).
Herrera, C., Sipe, C. L., McClanahan, W. S., Arbreton, A. J. A., & Pepper, S. K. (2000). Mentoring school-age children: Relationship development in community-based and school-based programs. Philadelphia: Public/Private Ventures.	Fun, social activities	Telephone interviews were conducted with mentors; "interviews and focus groups with youth, school and agency staff from eight exemplary programs." Mentor-reported measures.	Mentoring programs	669 mentors served school-aged children in community- and school-based programs (a subset of programs from authors' 1997 surveys). Mentors: varied in age, were 33-40% male; and 16-28% non-Caucasian.	"Relative to all the other variables we examined, the extent to which youth and mentors engage in social activities is the strongest contributing factor for both community-based and school-based programs in all three measures of positive relationship quality: closeness, emotional and instrumental supportiveness." Benchmark for most close and supportive relationships Some/a lot.
Metz, E., McLellan, J., & Youniss, J. (2003). Types of voluntary service and adolescents' civic development. Journal of Adolescent Research, 18(2), 188-203.	Type of voluntary service (engaging, agent for change)	Data drawn from two administrations of a survey--one in October 1998, the other in May 1999. Analyses controlled for background demographic characteristics, and personality dimensions (measured in a version of the adult Big Five instrument). Multivariate analyses of covariance were conducted. Social cause types of service gave students a chance to "improve or remedy an explicit social problem or a perceived injustice," as compared to "standard" service, such as tutoring, which consisted of "assisting other persons but without exposure to people in need or to issues of injustice or inequality in society" (p. 192).	Volunteering programs	428 students at a Boston area public high school. (Data were collected on 367 at Time 2)	From T1 to T2, students who performed social cause service increased significantly in social concern, while students who performed service not related to social causes decreased significantly in social concern. This finding was net of personality and demographic characteristics. Increases in intended civic activity (such as boycotting, demonstrating, or political campaigning) were also noted for the former group.

	1. Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
OPTIMIZING THE PARTICIPATION OF ADULT VOLUNTEERS					
Belgrave, F., Reed, Melba, C., Plybon, L., and Corneille, M. (2004). The Impact of a culturally enhanced drug prevention program on drug and alcohol refusal efficacy among urban African American girls. *Journal of Drug Education*, 34(4), 267-279	Cultural Enhanced Program Implementation		Drug abuse prevention program for urban African-American girls		

	1. Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Harvey, A. (2004). Africentric youth and family rites of passage program: promoting resilience among at-risk African American youths. *Social Work,* 49(1), pp.65-74	Culturally competent program evaluation	Participants were recruited from the juvenile justice system. Referred on the criterion that they did not have any history of drug use. IUR evaluated the MAAT Centers Rites of passage Proram over 3 years. The IUR collected data on 57 African American male adolescents. 17—1st cohort, 14—2nd cohort, and 27 ---3rd cohort. Family data was collected from 12 families. Pg 70	Substance abuse program	African American Adolescent boys between the age of 11.5 and 14.5 and their parents and other family members.	5 themes that was repeatedly identifies as contributing to the program success: holistic, family-oriented, africentric, strength-based, and had an indigenous staff. African based rituals increased the positive racial identity and cultural awareness of the youths and their parents.
Herrera, C., Sipe, C. L., McClanahan, W. S., Arbreton, A. J. A., & Pepper, S. K. (2000). Mentoring school-age children: Relationship development in community-based and school-based programs. Philadelphia: Public/Private Ventures.	Addressing the needs of the population. (Screening)	Telephone interviews were conducted with mentors; "interviews and focus groups with youth, school and agency staff from eight exemplary programs." Mentor-reported measures.	Mentoring programs	669 mentors served school-aged children in community- and school-based programs (a subset of programs from authors' 1997 surveys). Mentors: varied in age, were 33-40% male; and 16-28% non-Caucasian.	Screening was related to relationship development in community- but not school-based programs. What really matters is matching according to interest; "sharing similar interests is the second most important contributor to feelings of closeness and supportiveness on the part of the mentor." Benchmark for most close and supportive relationships: "four standard screening procedures plus additional screening."

	1. Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Herrera, C., Sipe, C. L., McClanahan, W. S., Arbreton, A. J. A., & Pepper, S. K. (2000). Mentoring school-age children: Relationship development in community-based and school-based programs. Philadelphia: Public/Private Ventures.	Training, pre- and post-match	Telephone interviews were conducted with mentors; "interviews and focus groups with youth, school and agency staff from eight exemplary programs." Mentor-reported measures.	Mentoring programs	669 mentors served school-aged children in community- and school-based programs (a subset of programs from authors' 1997 surveys). Mentors: varied in age, were 33-40% male; and 16-28% non-Caucasian.	Positive association with the strength of the relationships . Benchmark for most close and supportive relationships More than six hours of pre-match orientation/training; "at least two hours of post-match training or a minimum of monthly contact with program staff."
Karcher, M. (2005). The effects of developmental mentoring and high school mentors' attendance on their younger mentees' self-esteem, social skills, and connectedness. *Psychology in the Schools*, 42(1), 65-77.	Approach: Theory-based Mentors' attendance on their mentees' outcomes	Pre/post randomized experimental design was used to include an equivalent comparison group for both the high – and low-risk youth in the study. To determine the direct effects of six months of development mentoring, posttest between-group differences on self-reported measures of connectedness were computed using SPSS.	Developmental Mentoring	73 Caucasian, rural youth	The relationship between mentors; inconsistent attendance and mentees' decline in self-esteem and behavioral competence suggests that absent mentors may do more harm than good. Mentors' attendance predicted changes in mentees' social skills and self-esteem. Mentor attendance was a better predictor of mentee change than was mentee attendance, suggesting it was the experience with the mentor that best accounted for changes in the areas of self-management, self-esteem, and social skills.
King, W., Holmes, S., Henderson, M., & Latessa, E. (2001). The community corrections partnership: examining the long-term effects of youth participation in an Afrocentric diversion program. *Crime & Delinquency*, 47(4), pp. 558-572.		Reports an evaluation of an Afrocentric treatment program, for male, juvenile, felony offenders in one city. The evaluation uses a 2-group, quasi-experimental design to compare the 281 African American youths in the Afrocentric treatment program (called the Community Corrections Partnership) with a comparison group of 140 probation youths. Overall, the youths assigned to the Afrocentric treatment program performed slightly better than the probationers on 4 out of 15 measures of juvenile and adult criminality			

	1. Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
McGill, D. E., Mihalic, S., & Grotpeter, J. K. (2001). Blueprints for violence prevention: Big Brothers/Big Sisters program. Boulder: Center for the Study and Prevention of Violence, University of Colorado.	Supervision	Overview of various program evaluations.	Mentoring programs	Varies between program sites. Typically, youth, 6-18 years old, from single parent homes.	"Supervision was the program practice most associated with positive match outcomes--- those sites following national procedures for regular supervision had matches that were meeting at the highest rates."
Russell, S. and Lee, F. (2004). Practitioners' perspectives on effective practices for hispanic teenage pregnancy prevention. *Perspectives on Sexual and Reproductive Health*, 36(4), 142-149.	Culturally Effective Practices		Teenage Pregnancy Prevention		
Sipe, C. L. (1996). Mentoring: A synthesis of P/PV's research: 1988-1995. Philadelphia: Public/Private Ventures.	Staff-mentor contact/supervision	Observations based on a review of findings from 10 reports from 1988 to 1995 by P/PV	Mentoring programs	Mentoring pairs of adults and disadvantaged teenagers.	"When professional staff spend more time contacting volunteers, mentors have better "attendance" and develop better quality matches… Programs in which mentors were not contacted regularly by program staff reported the most "failed" matches--- those that did not meet consistently and, thus, never developed into relationships."
FAMILY PARTICIPATION/INVOLVEMENT					

	1. Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Dungan-Seaver, D. (1999). Afterschool programs: An analysis of research about characteristics of effectiveness. Produced for the McKnight Foundation. Retrieved March 16, 2003, from the World Wide Web: http://www.mcknight.org/display_file.asp?FileID=28	Family involvement	Review of several after-school studies.	Afterschool programs	Variety of programs and populations.	"Nearly everyone agrees that the most effective programs emphasize ongoing outreach to and communication with families," understand their needs and culture, and facilitate family engagement in the program.
Jekielek, S., Moore, K. A., & Hair, E. C. (2002). Mentoring programs and youth development: A synthesis. Washington, DC: Child Trends.	Involvement w/family	Review of numerous experimental, quasi-experimental, and non-experimental analyses.	Mentoring programs	Varied. Most often, disadvantaged youth.	This characteristic/approach was associated with "effective" mentoring (i.e., "improves youth outcomes, or it results in successful mentor-mentee relationships"). "when students perceived that their mentor knew their parents well, these youth had better GPAs and higher levels of college attendance than non-participants."
Silva, M. (2002). The effectiveness of school-based sex education programs in the promotion of abstinent behavior: A meta-analysis. Health Education Research, 17(4), 471-481.	Parental participation	Review of 12 experimental, quasi-experimental, or non-experimental (n=3) studies published in peer-reviewed journals from 1985-2000. To calculate effect sizes, the author included in analyses only one finding, and only the most successful intervention variation, if applicable, per study. Two studies involved parental participation.	Pro-abstinence sex education programs	Participants from the studies were "normal adolescent populations attending public or private schools in the US."	"Parental participation appeared to moderate the effects of sex education on abstinence as indicated by the significant Q test between groups ($QB(1) = 5.06$; $P = 0.025$)… Although smaller in magnitude ($d = 0.24$), the point estimate for the mean weighted effect size associated with programs with parental participation appears substantially larger than the mean associated with those where parents did not participate ($d = 0.04$). The confidence interval for parent participation does not include zero, this indicating a small but positive effect." This finding was also significant in weighted least-squares regression analysis [this was with % of females in the regression as well] ($B = 0.22$, $p < 0.05$, $SE = 0.09$). p477.

M. Table B4

IX. Findings on Dosage and Duration from Non-Experimental Research

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
DOSAGE					
Arbreton, A.J.A., & McClanahan, W.S. (2002). Targeted outreach: Boys & Girls Clubs of America's approach to gang prevention and intervention. Philadelphia: Public/Private Ventures.	Frequency of attendance	The report is based on two different studies. One program is to prevent at risk youth from joining gangs, the other is a program to intervene with youth already involved in gangs (discussed below). Neither background or socio-demographic characteristics, nor referral source, predicted level of participation. The gang prevention program used a Gang Risk Factors scale to determine if a youth was at high risk of gang involvement. Points were assigned for such behaviors as "exhibiting gang signs and symbols", coming from a highly distressed or crisis ridden family, of having family members or friends who are gang members. To evaluate change a sample or youth answered a questionnaire when they were recruited and again approximately 12 months later. Results were compared to a group of youth who did not attend Clubs.	Program for prevention of gang activity.	236 youth (avg. of 44 per site) at risk of gang membership were tracked from baseline to a follow-up survey after one year (81% of baseline pop; differs in characteristics by race). 48% aged 13 or older, 64% male, 51% African American, 29% Hispanic. Virtually all 'low income' (25% in public housing, 78% free/reduced-price lunch)	More frequent attendance is associated with: delayed onset of gang behavior; less contact with the juvenile justice system; stealing less, less likely to start smoking marijuana; higher grades, greater valuing of doing well in school; engaging in more positive after-school activities; increased levels of positive peer and family relationships. (p. 27)

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Arbreton, A.J.A., & McClanahan, W.S. (2002). Targeted outreach: Boys & Girls Clubs of America's approach to gang prevention and intervention. Philadelphia: Public/Private Ventures.	Frequency of attendance	The report is based on two different studies. One program is to prevent at risk youth from joining gangs (discussed above), the other is a program to intervene with youth already involved in gangs. Neither background or socio-demographic characteristics, nor referral source, predicted level of participation. Youth were determined to be in gangs according to self-report. To evaluate change a sample or youth answered a questionnaire when they were recruited and again approximately 12 months later. Results were compared to a group of youth who did not attend Clubs.	Program for intervention in gang involvement.	66 youth already in gangs (avg. of 34 per site) were tracked from baseline to a follow-up survey after one year (78% of baseline pop; differs in characteristics). 96% aged 13 or older, 74% male, 32% African American, 26% Hispanic, 24% Asian, 12% other race. Virtually all 'low income' (22% in public housing, 72% free/reduced-price lunch)	More frequent attendance is associated with: disengagement from gang-associated behaviors and peers; less contact with the juvenile justice system; greater expectations of graduating from high school or receiving a GED. (p. 27)
Cervantes, R., Ruan, K., Duenas, N. (2004). Programa shortstop: A culturally focused juvenile intervention for Hispanic youth. *Journal of Drug Education*, 34(4), 385-405.	Length of Intervention	The purpose of the study was to test the effectiveness of a culturally focused juvenile and substance abuse intervention program for first time Hispanic youth offenders.	Substance Abuse Intervention	352 youth and at least one parent/caregiver over the 5-year study period.	The analysis found a significant interaction between Length of Intervention and Legal Knowledge. Participants in 4 sessions vs 3 had a larger increase in legal knowledge

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Dynarski M., Moore, M., Mullens, J., Gleason, P., James-Burdumy, S., Rosenberg, L., et al. (2003).When schools stay open late: The national evaluation of the 21st Century Community Learning Centers program. Washington, DC: US Department of Education.	Dosage: Frequency of attendance	The middle school study was based on "comparison-student designs (with matching to identify comparison students)." Overall, the program was found to affect few participant outcomes, which may explain null findings in regard to program component effects. (More methodology information available on p.9 and 77 of the report.)	After-school programs	4400 middle school youth. "Based on nationally representative sample of after-school programs and participants and a matched comparison group" (34 school districts, 62 21st Century centers in the districts). (Avg. 57% minority, 66% of host schools considered high poverty)	Attendance was not related to any academic outcomes or student-reported delinquent behavior composite. (It was positively related to parental PTO attendance.) Note: the population of youth who attended most frequently appear to have many high-risk characteristics, potentially bucking the motivation-attendance pattern. Attendance categorized to: less than 20, 20-40, and more than 40 days over the school year.
Dynarski M., Moore, M., Mullens, J., Gleason, P., James-Burdumy, S., Rosenberg, L., et al. (2003).When schools stay open late: The national evaluation of the 21st Century Community Learning Centers program. Washington, DC: US Department of Education.	Dosage: Frequency of attendance	Data collection and random assignment of participants in the fall of 2000; subsequent data collection in the spring of 2001. Overall, the program was found to affect few participant outcomes, which may explain null findings in regard to program component effects. (More methodology information available on p.13 and 106 of the report.)	After-school programs	1000 elementary school children. Random assignment of students to treatment and control groups (14 school districts, 34 21st Century centers; these data from 7 districts. (Avg. 57% minority, 66% of host schools considered high poverty)	Attendance was not related to any academic outcomes except for receiving more help on their homework from a non-parent adult/less help from another child (which is necessarily built into program participation.) NOTE: the population of youth who attended most frequently appear to have several high-risk characteristics, potentially bucking the motivation-attendance pattern. Attendance categorized to: less than 20, 20-40, and more than 40 days over the school year.

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Fagan, J., & Inglesias, A. (1999). Father involvement program effects on fathers, father figures, and their Head Start children: A quasi-experimental study. Early Childhood Research Quarterly, 14(2), 243-269.	Dosage	Quasi-experimental; comparison group were elementary schools in geographically similar neighborhoods. Data on father involvement were collected from site sign-in sheets. Program was studied over 8 months. For intensity analyses, the participants were split between three different levels of participation, low being 0-4 hours of participation, high being over 21.5 hours.	Head Start father involvement program	146 interested fathers/father figures significantly involved in the raising of their children; children were enrolled in one of four urban public elementary schools with Head Start classrooms." 58% African American, 29% Latino-American; 68% biological fathers; "slightly more than one-half... participated in the labor force."	There was a significant association between "high-dosage [> 21.5 hours] participation in the intervention and increased father involvement with children at post-treatment." "The children of high dosage intervention fathers also showed higher mathematics readiness change scores," (e.g., applied problems skills).
Herrera, C., Sipe, C. L., McClanahan, W. S., Arbreton, A. J. A., & Pepper, S. K. (2000). Mentoring school-age children: Relationship development in community-based and school-based programs. Philadelphia: Public/Private Ventures.	Dosage	Telephone interviews were conducted with mentors; "interviews and focus groups with youth, school and agency staff from eight exemplary programs." Mentor-reported measures.	Mentoring programs	669 mentors in community- and school-based programs serving school-aged children (this, a subset of the programs used in authors' 1997 surveys). Between the two types of programs, mentors were 33-40% male; 16-28% non-Caucasian; and of various ages.	"Mentors who spend more time with their youth feel more close and supportive in their relationships." (Though not as important a predictor as social activities.) Benchmark for most close and supportive relationships Greater than 10hrs/month.

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Horn, K., Dino, G., Kalsekar, I., & Fernandes, A. (2004). Appalachian teen smokers: not on tobacco 15 months later. *American Journal of Public Health*. 94(2), pp.181-184.	Length of exposure to intervention		Substance Abuse (Tobacco) Intervention		
Jekielek, S., Moore, K. A., & Hair, E. C. (2002). Mentoring programs and youth development: A synthesis. Washington, DC: Child Trends.	Dosage	Review of numerous experimental, quasi-experimental, and non-experimental analyses.	Mentoring programs	Varied. Most often, disadvantaged youth.	This characteristic/approach was associated with "effective" mentoring (i.e., "improves youth outcomes, or it results in successful mentor-mentee relationships"). (From a few programs): youth whose mentors contacted them most frequently experienced myriad positive outcomes compared to less-contacted youth (better school attitudes, reaction to drug use situations, less absence, etc.); youth who spoke with their mentors rarely "did not experience benefits from program participation, and may even have experienced harm" (lower self-esteem).
Mehran, M., & White, K. R. (1988). Parent tutoring as a supplement to compensatory education for first-grade children. Remedial & Special Education (RASE), 9(3), 35-41.	Dosage	Children were "rank ordered according to their total reading scores on the CTBS." Starting with the two children with the lowest scores, pairs were made of children with similar scores; one from each pair was randomly assigned to the treatment group.	Compensatory education program	38 mothers of children identified for Chapter 1 compensatory education programs" (e.g., at-risk for having academic difficulty) in a small western city.	"When analyses were limited to those pairs of children in which the parent of the experimental group child had participated more completely in the tutoring program, there were both immediate and long-term statistically significant differences between the groups." This, compared to no lasting differences between the groups when intensity was not controlled.

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Rotheram-Borus, M. J., Koopman, C., Haignere, C., & Davies, M. (1991). Reducing HIV sexual risk behaviors among runaway adolescents. Journal of the American Medical Association, 266(9), 1237-1241.	Dosage	Nonrandomized control, with 78 youth at the intervention shelter compared with 67 runaways at a nonintervention shelter. Data collection at baseline, 3, and 6 months. Recruitment occurred over 2 years. The 145 participants who participated in subsequent interviews were 77% of the original participant group. "A 20-session intervention rotated in a 3-week sequence, with runaways joining the sequence at various points."	HIV intervention program	145 runaway youth (78 treatment), ages 11-18, at publicly-funded residential shelters in New York City. 36% male; 63% black, 22% Hispanic, 8% white. Sample groups at the two sites did not differ significantly in demographic characteristics.	"As the number of intervention sessions increased, runaways' reports of consistent condom use increased significantly (at 3 months, unique R2 = .06, $p < .05$; at 6 months, unique R2 = .09, $p < .05$), and their reports of engaging in a high-risk pattern of sexual behavior decreased significantly (at 3 months, unique r2 = .03, p = .06; at 6 months, unique R2 = .04, $p < .05$). "Participants received 3-30 sessions; median of 11 (X = 12.8, SD = 6.3)
Tobler, M. S., Roona, M. R., Ochshorn, P., Marshall, D. G., Streke, A. V., & Stackpole, K. M. (2000). School-based adolescent drug prevention programs: 1998 meta-analysis. Journal of Primary Prevention, 20(4), 275-337.	Dosage	Meta-analysis of 207 school-based drug prevention programs whose data were reported between 1978 and 1998.	School based drug prevention programs	Varied. Adolescents.	"Higher intensity interactive programs with 16 or more hours of lessons had greater impact than lower intensity efforts (average delivery of 6 hours), although length of programming did not influence outcomes for non-interactive programs."
Vandell, D.L., & Shumow, L. (1999). After-school child care programs. The Future of Children: When school is out, 9(2), 64-80.	Intensity of involvement	Review of several large-scale studies on after-school programs.	Afterschool child care programs	Variety of programs and populations: Ecological Study of Afterschool Care, Study of After-School Care and Children's Development, Child Development Project, and Boston After-School Time Study. Among all the programs, 1038 children were involved.	INTERESTING: "Findings from the Child Development Project, which included middle-income and low-income children, suggest that children's development suffers from both too few and too many activities… Children who spent moderate amounts of time (one to three hours per week) in activities and lessons during grades one and three were more socially competent in grade six. Children who participated in no activities and those who were more extensively involved did not fare as well." p70
DOSAGE & DURATION					

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Eisen, M., Pallitto, C., Bradner, C., & Bolshun, N. (2000). Teen risk-taking: Promising programs and approaches. Washington, DC: Urban Institute.	Duration and dosage	Review of 51 "successful" prevention programs. Minimum criteria were: "treatment group is matched with a comparison… group; there is at least one follow-up review at three months after program completion; initial sample size consists of 100 adolescents or more; the study retains at least half of the participants at the final follow-up review; there is a statistically significant improvement for at least one target behavior in at least one target group." Secondary review of 21 high-quality evaluations. For inclusion in this analysis, additional criteria were: "Each program collected data about participants' problem behavior prior to the program beginning; the follow-up review period is longer (at least 12 months or the full school year); no fewer than 150 individuals are included in the treatment and comparison; the program retains more participants--- 67% in each group by the final follow-up date."	Prevention programs	Varied. Adolescents.	Common element of program success (one of six), culled from the review of higher-quality studies: "substantial duration and intensity are necessary." "The most effective programs are generally more intense in terms of the number of sessions and the length of intervention. Of the programs examined, 14 programs include over 10 hours of intervention and 2 have over 100 hours of intervention. Half the programs take place over 10 sessions, and a few are taught over an entire school year or more."
Lipsey, M. W. (1992). Juvenile delinquency treatment: A meta-analytic inquiry into the variability of effects. In T. Cook, H. Cooper, D.S. Cordray, H. Hartmann, L.V. Hedges, R.J. Light, et al, Eds., Meta-analysis for explanation: A casebook (p. 83-127). New York: Russell Sage Foundation.	Duration and dosage	Meta-analyses of experimental and quasi-experimental studies (N = 443) conducted between 1950 and 1987. One of the criteria for the inclusion of studies was that the study had to be experimental based on random assignments or if it was quasi-experimental, the study had to provide: 1) pre- and post-outcome measures, and 2) "some evidence of matching between the two groups prior to treatment" (or measures that show the similarity of the two groups prior to treatment). The study used "a weighted (stepped) multiple regression in which the contribution of each case (study) to the analysis is weighted by the inverse variance of the effect size" in order to adjust for different sample sizes. All the measures related to methodologies of the studies were "stepped into" the regressions prior to treatment measures.	Juvenile delinquency treatment.	Varied. The sample size ranged from 25 or fewer to 801 and more participants.	As both duration and dosage increased, participants experienced larger effect size for reducing re-arrest/reconviction. There was a "modest positive relationship between effect size and the duration, frequency," and "amount or intensity of treatment (dosage)."
DURATION					

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Botvin, G. J., Baker, E., Filazzola, A. D., & Botvin, E. M. (1990). A cognitive-behavioral approach to substance abuse prevention: One-year follow-up. Addictive Behaviors, 15, 47-63.	Duration	Schools were randomly assigned to either "(a) prevention program implemented by older peer leaders, (b) prevention program implemented by regular classroom teachers, (c) prevention program with booster sessions implemented by older peer leaders, (d) prevention program with booster sessions implemented by regular classroom teachers," and control. All program leaders were trained using the same procedures.	Substance abuse prevention	Whole sample: "998 eighth-graders from 10 suburban New York junior high schools" were followed over two years. Sample was "predominately [80%] white and…. from middle-class families." 49% of the sample was male; and participants came from largely intact families *83%).	"Where prevention effects were found, the booster condition was superior to the non-booster condition." Monthly, weekly, daily, and an index of cigarette smoking were significantly lower for the booster peer group than the non-booster peer group. Note: booster sessions reinforced the prevention program material from the previous year.
Jekielek, S., Moore, K. A., & Hair, E. C. (2002). Mentoring programs and youth development: A synthesis. Washington, DC: Child Trends.	Duration	Review of numerous experimental, quasi-experimental, and non-experimental analyses.	Mentoring programs	Varied. Most often, disadvantaged youth.	This characteristic/approach was associated with "effective" mentoring (i.e., "improves youth outcomes, or it results in successful mentor-mentee relationships"). Positive outcomes found for 6 or more months, none for three to 6 months. Short relationships may actually harm youth; negative outcomes (e.g., less academic confidence, lower self-worth) were found for youth in relationships that dissolved quickly (i.e., less than three months). There was limited research on this, but "one study did show evidence that all the program impacts disappeared when the intervention ended (with the exception of persistent decreased marijuana use 6 months after intervention end).
Sipe, C. L. (1996). Mentoring: A synthesis of P/PV's research: 1988-1995. Philadelphia: Public/Private Ventures.	Duration	Observations based on a review of research findings from 10 reports from 1988 to 1995 by P/PV	Mentoring programs	Mentoring pairs of adults and disadvantaged teenagers.	"Generally about six months of regular meetings were required" for trust to develop in the pair.

N. Table B5

X. Findings on Overall Program Quality from Non-Experimental Research

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
GENERAL QUALITY					
Faw, L., Hogue A., and Liddle, H.A., (2005). Multidimensional implementation evaluation of a residential treatment program for adolescent substance abuse. *American Journal of Evaluation*, 26(1), 77-94.	Approach: Theory-based	The authors applied contemporary methods from the evaluation literature to measure implementation in a residential treatment program for adolescent substance abuse. The study attempted to apply Mowbray's model to an implementation evaluation of a residential treatment program for adolescent substance abuse. Structure and process were conceptualized as the two main categories within which elements of the program theory would fit	Adolescent Substance Abuse	43 Consecutive admissions to the ATP, Between the ages of 13 and 17, dually diagnosed meeting *DSM-IV* criteria, currently living with one parent, provided informed consent	On average, 50% of the treatment services prescribed by the program theory on a weekly basis. The milieu of the program was rated by adolescents as highly therapeutic. Pyschometric suggest therapeutic milieu can be measured reliably in adolescents.
Pathways to Success for Youth: What Counts in After-School. Massachusettes After-School Research Study (MARS).	Quality Indicators and Youth Outcomes	Collected pre and post data on a variety of academic and youth development outcomes of afterschool program participation that have been documented in previous studies. Used "multilevel modeling" to control for student's pre-test scores as well as a set of demographic characteristics. The goal of this approach was to separate out the influence of program quality/program charactersitsc from other possible causes of the change in SAYO outcomes.	Youth Outcomes	Varied Adolescents in afters-school time programs.	The analysis indicated that, while change in SAYO-T scores provided by teachers varied by child background characteristics, they did not vary by program characteristics or program quality. The study could not establish links between the changes in outcomes and the quality of the after-school program the children attended.

	Study Design		Program and Sample Description		Findings
Source	**Manipulable Feature**	**Methodological Approach**	**Domain**	**Population**	**Differences in Effectiveness**
Spielberger, J., Horton, C., Michels, L., & Halpern, R. (2004). New on the shelf: teens in the library. Findings from the evaluation of public libraries as partners in youth development. *An Initiative of the Wallace Foundation*. Final Report, July 31, 2004.		Evaluation data were gathered from a variety of sources, including (1) annual site visits involving program observations and interviews with youth, library staff, and representatives of partner organizations; (2) annual surveys of youth, library staff, and community informants; and (3) the quarterly collection of information on youth participants, program activities, staff and youth trainings, community partnerships, and program expenditures. In addition, three sites were selected for an intensive study of youth participation, and four were used for in-depth examination of cost and financing issues. As noted above, a policy map (Whalen & Costello, 2002) also was developed during the first half of the Initiative to lay out the policy implications of the potential connections between libraries and the national youth development movement. (Additional information about the research methodology can be found in Appendix B.)			

Appendix C

Features of Developmental Settings		
9-Year-Olds	**Features**	**16-Year-Olds**
Children are paired with "buddies" — other participants in the after-school program that are also in their class. They spend the ten minutes of free time in between school and the program with their buddy, either in the cafeteria or on the playground — two designated areas where parent volunteers relax with children until buses have left and the space used for the program is ready.	**Physical and Psychological Safety**	In an effort to reduce conflicts occurring after school, a core of volunteer parents and staff from a local community center are ready to greet students, providing "coverage" and creating an opportunity to build rapport between youth and neighborhood adults between the last school bell and students' travel home or to after-school programs.
The Culture Club after-school program offers a consistent schedule. The activity board for Tuesday lists: Snacks: 2:50; Games Around the World: 3:15; Homework Heroes: 4:15; Discovery Workshops: 5:00. Whatever day it is, participants know they will spend the afternoon with adults they trust, will get to move around the center and do something interesting.	**Appropriate Structure**	From 2:30–4:00 the teen center offers a variety of options, including computers, open gym, a quiet area for reading/studying or an informal volunteer-led activity like sketching. At 4:00, teens meet in small groups to work on their community service projects. At 5:30, some stay to talk with specific staff, a small group prepares to leave for their street outreach shift, and others prepare the lounge for open mike night.
Staff development and retention is a strong focus. The director maintains open communication and consistently follows through on compensation and recognition, training, scheduling and adequate program resources. Staff support allows the center to provide children with consistent adults ready to focus on their needs in the program.	**Supportive Relationships**	Club members get a lot of support from each other in Express Yourself — a ritual the group initiated when the program began. At every meeting, participants set aside time to share issues on their mind. Adult staff provide a consistent presence, modeling listening, supporting an environment of psychological safety, and following up with individual youth as needed.
Staff members help children make "a plan" for how they want to use their time, helping children who want to join an activity with other children to do so, and structuring activity spaces and materials ahead of time to minimize conflicts and encourage inclusion.	**Opportunities to Belong**	Teens set and monitor the program rules based on their principles of inclusiveness and mutual support. New young people are greeted by peer staff members who talk to them about what goes on at the center, and set the tone for making "The Spot" a place where everyone can belong.
"Everybody cares for the center" is a motto practiced every day at South End Neighborhood Club. During snack time, participants rotate being "on" for snack duty, helping with serving the snack and doing light clean-up afterwards.	**Positive Social Norms**	"Family meetings" provide a space for teens and staff to set goals and norms, plan activities, make decisions, solve problems and reflect. Staff facilitate discussions, some scheduled and some ad-hoc, always modeling active listening skills, a structured problem-solving process, and a focus on positive program culture.
On Thursdays, several participants in the local after-school club travel to a nearby retirement community to listen to the stories of seniors and spend time with them. The youth will collect these stories into a book that will be distributed at the community fair.	**Support for Efficacy and Mattering**	Fifty cents of every purchase of coffee from a youth-run coffee delivery service in downtown Nashville goes to support youth programs — young people earn income, learn the skills of running a small business, and contribute to their community to provide expanded opportunities for their peers.
Students involved in the Readers for Life literacy program can spend a half-hour reading to a guide dog in training, building reading skills while avoiding the social pressure of reading aloud to other people. This reading time also helps the guide dog get used to human contact.	**Opportunities for Skill Building**	Poetry slams are popular events at the Zone. Youth form groups that review and practice poetry together. To gain skills, interested youth join weekend workshops taught every other month by their peers or college students. Slams are held every few months, and there are opportunities to join a competitive slam team.
Program staff comprised of graduate student teachers and neighborhood parents make regular links with teachers and parents through face-to-face meetings and "passports" that each student carries between school, after school and home.	**Integration of Family, School and Community Effort**	Staff at this employment program function as part educator, part guidance counselor and part life planner. They move freely between where youth live, hang out, and go to school, and are respected in all worlds. They help teens develop individualized plans, connect with the services they need, and make the most of their internship experiences.
A school-coordinated neighborhood after-school initiative allows participating students to join after-school activities in one of four sites. Vans, purchased through foundation grants and jointly operated between sites, pick up participants and take them safely home.	**Basic Care and Services**	An after-school drama troupe creates original theater to address physical and mental health issues impacting young teens in their neighborhood. At each performance, they make sure that related health information is available and help connect their peers to community resources.

Note. From "Policy Commentary #2: High School After-School: What is it? What Might it Be? Why is it Important?" by The Forum for Youth Investment, 2003, Washington, DC: Author. Copyright 2003 by The Forum for Youth Investment. Reprinted with permission. Examples developed by Forum staff. List of features adapted by Forum for Youth Investment from "Community Programs to Promote Youth Development," by the National Research Council and the Institute of Medicine, (J. Eccles and J. A. Gootman, Eds.), 2002. Washington, DC: National Academy Press. (Available online at www.nap.edu/catalog/10022.html.)

Appendix D
Presentation of Findings

Table D1

XI. Conclusions Regarding Program Features, as Expressed by Source of Information[11]

Manipulable Feature	Negative	Negative-Neutral	Neutral	Positive-Neutral	Positive
Staff & Staffing					
Staff training: More participant-or program-specific training vs. less				N	Ex, PW
Committed staff with a combination of training, experience, and a positive personality					PW
Staff management and supervision					PW
Ongoing staff support and development					PW
Staff is aware of the developmental needs specific to the participant group					PW
Peer leaders vs. adult leaders in substance use programs				Ex	
Nurse home-visiting vs. paraprofessional home-visiting				Ex	
Lower participant-staff ratio (or higher staff-participant ratio) vs. higher participant-staff ratio (or lower staff-participant ratio)			PW[12]	N	Ex; PW (for younger children)
Staff "matching" program population according to demographic characteristics.			N		
Staff cultural competence and sensitivity					PW
Positive staff-participant interaction				N (older children, teens)	N (young children); PW
Staff stability/retention vs. staff turnover				N	PW
Higher vs. lower, or more vs. less satisfying, staff wages				N	PW
Conceptual Approaches					
Fidelity to program design ("strength of implementation")				N	
Participant-centered approaches				N	PW

[11] Ex = experimental findings (Chapter II); N = non-experimental findings (Chapter III); PW = provider wisdom (Chapter IV).

[12] Provider wisdom suggests that both staff-child ratio and group size, for school-age and older children, should necessarily vary by particular program characteristics. See Chapter IV for more information.

Manipulable Feature	Negative	Negative-Neutral	Neutral	Positive-Neutral	Positive
Degree of structure in programs				N	
Degree of flexibility in programs		N (young boys)		N (older children and teens)	PW
Clear and consistent message				N	PW
Multi–component, "whole person" approach				N	PW
Strength-based philosophy				N	PW
Information-only/didactic approach compared to interactive approach			Ex, N, PW		
Interactive approach compared to non-interactive approaches				Ex	N, PW
"Shock" approaches		N			
Program Practices					
Video instruction plus interactive and/or skills training vs. video instruction only				Ex	
Case management as compared to cash incentives					Ex
More versus less intensive case management/ Multisystemic Therapy					Ex
Mediation instead of litigation in child custody disputes					Ex
Smaller vs. larger group size			PW[2]	N	Ex, PW (younger children)
Screening, training and supervising adult volunteers				N	
Diversity of activities offered		N (younger children)		N (older children)	PW
Interesting, engaging, enjoyable activities				N	PW
Behavior management and discipline					PW
Safe, stable atmosphere					PW
Targeting participation of hard-to-reach populations					PW
Conducting program evaluations					PW
Providing incentives for participation					Ex, PW
Accessibility of program to participants					PW
Family participation				Ex, N	PW
Community collaboration					PW
Dosage & Duration					
Dosage: Higher vs. lower levels				Ex, N	PW
Duration: Higher vs. lower levels				N	PW
Dosage + duration: Higher vs. lower levels				Ex	N, PW
Overall Program Quality					
Overall program quality					N, PW

Table D2

XII. Conclusions Regarding Program Features, by Age and Program Context/Setting, as Expressed by Source of Information[13]

	Age			Institutional setting				
Manipulable Feature		**Middle childhood (6-12)**	**Teens and Youth (13 and older)**	**Classroom-based program**	**Mentoring program or interpersonal counseling**	**Parenting/ parent-child program**	**Center-based or after-school program**	**Home-based program**
Staff & Staffing								
Staff training: More participant- or program-specific training vs. less		Positive (Ex; PW); Positive-neutral (N)	Positive (Ex; PW)	Positive (Ex; PW)	Positive (Ex)		Positive (Ex; PW); Positive-neutral (N)	
Staff education: More formal education vs. less								
Committed staff with a combination of training, experience, and a positive personality		Positive (PW)	Positive (PW)				Positive (PW)	
Staff management and supervision		Positive (PW)	Positive (PW)	Positive (PW)			Positive (PW)	
Ongoing staff support and development		Positive (PW)	Positive (PW)				Positive (PW)	
Staff is aware of the developmental needs specific to the participant group		Positive (PW)	Positive (PW)	Positive (PW)			Positive (PW)	
Peer leaders vs. adult leaders in substance use programs			Positive-neutral (Ex)	Positive-neutral (Ex)			Positive-neutral (Ex)	
Nurse home-visiting vs. paraprofessional home-visiting			Positive-neutral (Ex)			Positive-neutral (Ex)		

[13] Ex = experimental findings (Chapter II); N = non-experimental findings (Chapter III); PW = provider wisdom (Chapter IV).

	Age			Institutional setting				
Manipulable Feature		**Middle childhood (6-12)**	**Teens and Youth (13 and older)**	**Classroom-based program**	**Mentoring program or interpersonal counseling**	**Parenting/ parent-child program**	**Center-based or after-school program**	**Home-based program**
Lower participant-staff ratio (or higher staff-participant ratio) vs. higher participant-staff ratio (or lower staff-participant ratio)		Positive-neutral (N); Mixed (PW)[14]	Mixed (PW)[2]	Positive (Ex)		Positive-neutral (N)	Positive-neutral (N); Mixed (PW)[2]	
Staff "matching" program population according to demographic characteristics.		Neutral (N)	Neutral (N)	Neutral (N)	Neutral (N)	Positive-neutral (N)	Neutral (N)	
Staff cultural competence and sensitivity		Positive (PW)	Positive (PW)				Positive (PW)	
Positive staff-participant interaction		Positive-neutral (N); positive (PW)	Positive-neutral (N); positive (PW)		Positive-neutral (N)		Positive-neutral (N); positive (PW)	
Staff stability/retention vs. staff turnover		Positive (PW)	Positive (PW)				Positive (PW)	
Higher vs. lower, or more vs. less satisfying, staff wages		Positive (PW)	Positive (PW)				Positive (PW)	
Fidelity to program design ("strength of implementation")		Positive-neutral (N)	Positive-neutral (N)	Positive-neutral (N)	Positive-neutral (N)		Positive-neutral (N)	
Participant-centered approaches		Positive-neutral (N); Positive (PW)	Positive-neutral (N); Positive (PW)		Positive-neutral (N)		Positive (PW)	
Degree of structure in programs								

[14] Provider wisdom suggests that both staff-child ratio and group size, for school-age and older children, should necessarily vary by particular program characteristics. See Chapter IV for more information.

	Age			Institutional setting				
Manipulable Feature		**Middle childhood (6-12)**	**Teens and Youth (13 and older)**	**Classroom-based program**	**Mentoring program or interpersonal counseling**	**Parenting/ parent-child program**	**Center-based or after-school program**	**Home-based program**
Degree of flexibility in programs		Positive-neutral (N)	Positive (PW)		Positive-neutral (N)		Positive-neutral (N); Positive (PW)	
Clear and consistent message		Positive-neutral (N)	Positive-neutral (N); Positive (PW)	Positive-neutral (N)			Positive-neutral (N); Positive (PW)	
Multi–component, "whole person" approach		Positive-neutral (N); Positive (PW)	Positive-neutral (N); Positive (PW)	Positive-neutral (N)			Positive-neutral (N); Positive (PW)	
Strength-based philosophy			Positive (PW)				Positive (PW)	
Information-only/didactic approach compared to interactive approach		Neutral (Ex; N; PW)	Neutral (N; PW)	Neutral (Ex, N)			Neutral (Ex, N; PW)	
Interactive approach compared to non-interactive approaches		Positive (Ex; N; PW)	Positive (N, PW)	Positive-neutral (Ex); Positive (N)			Positive(N; PW)	
"Shock" approaches			Negative-neutral (N)				Negative-neutral (N)	
Video instruction plus interactive and/or skills training vs. video instruction only			Positive (Ex)	Positive (Ex)		Positive (Ex)	Positive-neutral (Ex)	
Case management as compared to cash incentives			Positive (Ex)				Positive (Ex)	
More versus less intensive case management / Multisystemic Therapy		Positive (Ex)					Positive (Ex)	
Mediation instead of		Positive (Ex)			Positive (Ex)	Positive (Ex)		

	Age			Institutional setting				
Manipulable Feature		**Middle childhood (6-12)**	**Teens and Youth (13 and older)**	**Classroom-based program**	**Mentoring program or interpersonal counseling**	**Parenting/ parent-child program**	**Center-based or after-school program**	**Home-based program**
litigation in child custody disputes								
Smaller vs. larger group size		Positive-neutral (N); Mixed (PW)[2]	Positive-neutral (N); Mixed (PW)[2]				Positive-neutral (N); Mixed (PW)[2]	
Screening, training and supervising adult volunteers		Positive-neutral (N)	Positive-neutral (N)		Positive-neutral (N)			
Diversity of activities offered		Mixed (N);[15] Positive (PW)	Positive (PW)				Mixed findings by age (N);[3] Positive (PW)	
Interesting, engaging, enjoyable activities		Positive-neutral (N); Positive (PW)	Positive-neutral (N); Positive (PW)	Positive-neutral (N)	Positive-neutral (N)		Positive-neutral (N); Positive (PW)	
Behavior management and discipline		Positive (PW)	Positive (PW)				Positive (PW)	
Safe, stable atmosphere		Positive (PW)	Positive (PW)				Positive (PW)	
Targeting participation of hard-to-reach populations			Positive (PW)				Positive (PW)	
Conducting program evaluations		Positive (PW)	Positive (PW)				Positive (PW)	
Providing incentives for participation		Positive (Ex)	Positive (PW)				Positive (PW)	
Accessibility of program to participants		Positive (PW)	Positive (PW)				Positive (PW)	
Family participation		Positive-neutral (Ex; N)	Positive-neutral (N)	Positive-neutral (Ex)	Positive-neutral (N)		Positive (Ex); Positive-neutral (N)	

[15] Negative-neutral for younger school-age children (1st grade), positive-neutral for older school-age children (3rd grade; N).

	Age			Institutional setting				
Manipulable Feature		**Middle childhood (6-12)**	**Teens and Youth (13 and older)**	**Classroom-based program**	**Mentoring program or interpersonal counseling**	**Parenting/ parent-child program**	**Center-based or after-school program**	**Home-based program**
Community collaboration		Positive (PW)	Positive (PW)	Positive (PW)		Positive (PW)	Positive (PW)	
Dosage & Duration								
Dosage: Higher vs. lower levels		Positive-neutral (N); Positive (PW)	Positive-neutral (Ex; N)	Positive-neutral (Ex; N)	Positive-neutral (Ex; N)	Positive-neutral (Ex; N)	Positive-neutral (Ex; N)	Positive-neutral (Ex)
Duration: Higher vs. lower levels		Positive-neutral (N); Positive (PW)	Positive-neutral (N)	Positive-neutral (N)	Positive-neutral (N)		Positive-neutral (N)	
Dosage + duration: Higher vs. lower levels		Positive (Ex; PW); Positive-neutral (N)	Positive-neutral (N); Positive (PW)	Positive-neutral (N)	Positive (PW)	Positive-neutral (N)	Positive (Ex; PW); Positive-neutral (N)	Positive (Ex); Positive-neutral (N)
Overall Program Quality								
Overall program quality		Positive (PW)	Positive (PW)				Positive (PW)	

Appendix E

Valuable Resources for Program Design and Implementation

Bruner, C., Greenberg, M., Guy, C., Little, M., Schorr, L., & Weiss, H. (2002). *Funding what works: Exploring the role of research on effective programs and practices in government decision-making* (Resource brief #10). Des Moines, IA: National Center for Service Integration.

Catalano, R. F., Berglund, M. L., Ryann, J. D., Lonczak, H. S., & Hawkins, J. D. (1999). *Positive youth development in the United States: Research findings on evaluations of positive youth development programs.* Washington, DC: U.S. Department of Health & Human Services, Office of the Assistant Secretary of Planning and Evaluation.

Family Support America (former name: Family Resource Coalition): http://www.familysupportamerica.org/content/home.htm

Harms, T., Jacobs, E. V., & White, D. R. (1996). *School-age care environment rating scale.* New York: Teachers College Press.

Harvard Family Research Project: Home page: http://www.gse.harvard.edu/hfrp/ Evaluations: http://www.gse.harvard.edu/hfrp/projects/afterschool/mott/mott6.html

Institute for Youth Development: www.youthdevelopment.org

National Research Council, & Institute of Medicine. (2002). *Community programs to promote youth development*. Washington, DC: National Academy Press.

Oregon Social Learning Center: www.oslc.org

Parents as Teachers: http://www.patnc.org/

Pathways to Outcomes & The Pathways Mapping Project: www.PathwaysToOutcomes.org

Slavin, R. E., Karweit, N. L., & Wasik, B. A. (Eds.). (1994). *Preventing early school failure: Research, policy, and practice*. Needham Heights, MA: Allyn and Bacon.

U.S. Department of Education: Upcoming database of educational program evaluations.

Watson, B. H. (2002). *Community change for youth development: Ten lessons from the CCYD initiative*. Philadelphia: Public/Private Ventures.

References

Arbreton, A. J. A., & McClanahan, W. S. (2002). *Targeted outreach: Boys & Girls Clubs of America's approach to gang prevention and intervention*. Philadelphia: Public/Private Ventures.

Barton, W. H., & Butts, J. A. (1990). Viable options: intensive supervision programs for juvenile delinquents. *Crime and Delinquency, 36*(2), 238-256.

Battistich, V., Schaps, E., Watson, M., & Solomon, D. (1996). Prevention effects of the Child Development Project: Early findings from an ongoing multisite demonstration trial. *Journal of Adolescent Research, 11*(1), 12-35.

Beckett, M., Hawken, A., & Jacknowitz, A. (2001). *Accountability for after-school care: Devising standards and measuring adherence to them*. Santa Monica, CA: RAND.

Blau, D. M. (2000). The production of quality in child-care centers: Another look. *Applied Developmental Science, 4*(3), 136-148.

Botvin, E. M., Renick, N., & Baker, E. (1983). The effects of scheduling format and booster sessions on a broad-spectrum psychosocial approach to smoking prevention. *Journal of Behavioral Medicine, 6*, 359-379.

Botvin, G. J., Baker, E., Filazzola, A. D., & Botvin, E. M. (1990). A cognitive-behavioral approach to substance abuse prevention: One-year follow-up. *Addictive Behaviors, 15*(1), 47-63.

Burns, B. J., Farmer, E. M. Z., Angold, A., & Costello, E. J. (1996). A randomized trial of case management for youths with serious emotional disturbance. *Journal of Clinical Child Psychology, 25*(4), 476-486.

Burrow, F. B. (2001). The effect of parental involvement on social skills training for children with and without attention deficit hyperactivity disorder. *Dissertation Abstracts International: Section B: The Sciences & Engineering, 62*(4-B), 2048.

Campbell, D., & Stanley, J. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand-McNally.

Catalano, R. F., Berglund, M. L., Ryann, J. D., Lonczak, H. S., & Hawkins, J. D. (1999). *Positive youth development in the United States: Research findings on evaluations of positive youth development programs.*: U.S. Department of Health & Human Services, Office of the Assistant Secretary of Planning and Evaluation.

Connell, D. B., Turner, R. R., & Mason, E. F. (1985). Journal of School Health. *Summary of the findings of the school health education evaluation: Health promotion effectiveness, implementation, and costs, 55*(8), 316-323.

Currie, J. (2000). *Early childhood intervention programs: What do we know?* Paper presented at the Brookings Institution's Roundtable on Children, Washington, DC.

Dishion, T. J., & Andrews, D. A. (1995). Preventing escalation in problem behaviors with high-risk young adolescents: Immediate and 1-year outcomes. *Journal of Consulting and Clinical Psychology, 63*(4), 538-548.

Dungan-Seaver, D. (1999). *Afterschool programs: An analysis of research about characteristics of effectiveness*. Produced for the McKnight Foundation. Retrieved March 16, 2003, from the World Wide Web: http://www.mcknight.org/display_file.asp?FileID=28

Dynarski, M., Moore, M., Mullens, J., Gleason, P., James-Burdumy, S., Rosenberg, L., & Garrett, S. (2003). *When schools stay open late: The national evaluation of the 21st Century Community Learning Centers program*. Washington, DC: US Department of Education.

Eisen, M., Pallitto, C., Bradner, C., & Bolshun, N. (2000). *Teen risk-taking: Promising programs and approaches*. Washington, DC: Urban Institute.

Emery, R. E., Laumann-Billings, L., Waldron, M. C., Sbarra, D. A., & Dillon, P. (2001). Child custody mediation and litigation: Custody, contact, and coparenting 12 years after initial dispute resolution. *Journal of Consulting and Clinical Psychology, 69*(2), 323-332.

Epstein, L. H., Valoski, A., Wing, R. R., & McCurley, J. (1990). Ten-year follow-up of behavioral family-based treatment for obese children. *Journal of American Medial Association, 264*(19), 2513-2523.

Fagan, J., & Forst, M. (1996). Risks, fixers, and zeal. *The Prison Journal, 76*(16), 5-21.

Fashola, O. S. (1998). *Review of extended-day and after-school programs and their effectiveness* (Report No. 24). Washington, DC: Center for Research on the Education of Students Placed at Risk.

Fraser, M. W., Walton, E., Lewis, R. E., Pecora, P. J., & Walton, W. K. (1996). An experiment in family reunification: Correlates of outcomes at one-year follow-up. *Children and Youth Services Review, 18*(4-5), 335-361.

Glazerman, S., Levy, D. M., & Myers, D. (2002). *Nonexperimental replications of social experiments: A systematic review* (Interim report/discussion paper). Washington, DC: Corporation for the Advancement of Policy Evaluation, Mathematica Policy Research, Inc.

Golan, M., Fainaru, M., & Weizman, A. (1998). Role of behaviour modification in the treatment of childhood obesity with the parents as the exclusive agents of change. *International Journal of Obesity, 22*, 1217-1224.

Gottfredson, D. C., & Gottfredson, G. D. (2002). Quality of school-based prevention programs: Results from a national survey. *Journal of Research in Crime and Delinquency, 39*(1), 3-35.

Greenwood, P. W., Deschenes, E. P., & Adams, J. (1993). *Chronic juvenile offenders: Final results from the Skillman After-care Experiment*. Santa Monica, CA: The RAND Corporation.

Hall, G. E., & Hord, S. M. (1987). *Change in schools*. Albany, NY: State University of New York Press.

Hallfors, D., & Godette, D. (2002). Will the 'Principles of Effectiveness' improve prevention practice? Early findings from a diffusion study. *Health Education Research, 17*(4), 461-470.

Halpern, R., Spielberger, J., & Robb, S. (2001). *Evaluation of the MOST (Making the Most of Out-of-School Time) initiative: Final report* (Discussion paper). Chicago: Chapin Hall Center for Children.

Henggeler, S. W., Melton, G. B., Brondino, M. J., Scherer, D. G., & Hanley, J. H. (1997). Multisystemic therapy with violent and chronic juvenile offenders and their families: The role of treatment fidelity in successful dissemination. *Journal of Consulting and Clinical Psychology, 65*(5), 821-833.

Herrera, C., & Arbreton, A. J. A. (2003). *Increasing opportunities for older youth in after-school programs: A report on the experiences of Boys & Girls clubs in Boston and New York City*. Philadelphia: Public/Private Ventures.

Herrera, C., Sipe, C. L., McClanahan, W. S., Arbreton, A. J. A., & Pepper, S. K. (2000). *Mentoring school-age children: Relationship development in community-based and school-based programs*. Philadelphia: Public/Private Ventures.

Hollister, R. G., & Hill, J. (1995). Problems in the evaluation of community-wide initiatives. In C. H. Weiss (Ed.), *New approaches to evaluating community initiatives: Concepts, methods, and contexts*. Washington, DC: The Aspen Institute.

Howes, C., & Norris, D. J. (1997). Adding two school age children: Does it change quality in family child care? *Early Childhood Research Quarterly, 12*, 327-342.

Innocenti, M. S., & White, K. R. (1993). Are more intensive early intervention programs more effective? A review of the literature. *Exceptionality, 4*(1), 31-50.

Jastrzab, J., Blomquist, J., Masker, J., & Orr, L. (1997). *Youth Corps: Promising strategies for young people and their communities* (Studies in Workforce Development and Income Security, No. 1-97). Cambridge, MA: Abt Associates, Inc.

Jekielek, S., Cochran, S., & Hair, E. C. (2002). *Employment programs and youth development: A synthesis* (Report for the Edna McConnell Clark Foundation). Washington, DC: Child Trends, Inc.

Jekielek, S., Moore, K. A., & Hair, E. C. (2002). *Mentoring programs and youth development: A synthesis* (Report for the Edna McConnell Clark Foundation). Washington, DC: Child Trends, Inc.

Jemmott, J. B., III, Jemmott, L. S., & Fong, G. T. (1992). Reductions in HIV risk-associated sexual behaviors among black male adolescents: Effects of an AIDS prevention intervention. *American Journal of Public Health, 82*(3), 372-377.

Jemmott, J. B., III, Jemmott, L. S., & Fong, G. T. (1998). Abstinence and safer sex HIV risk-reduction interventions for African American adolescents. *Journal of American Medical Association, 279*(19), 1529-1536.

Jemmott, J. B., III, Jemmott, L. S., Fong, G. T., & McCaffree, K. (1999). Reducing HIV risk-associated sexual behavior among African American adolescents: Testing the generality of intervention effects. *American Journal of Community Psychology, 27*(2), 161-187.

Jucovy, L. (2002). *Same-race and cross-race matching* (Technical Assistance Packet #7). Philadelphia: Public/Private Ventures.

Kamb, M. L., Fishbein, M., Douglas, J. M., Jr., Rhodes, F., Rogers, J., Bolan, G., Zenilman, J., Hoxworth, T., Malotte, K., Iatesta, M., Kent, C., Lentz, A., Graziano, S., Byers, R. H., & Peterman, T. A. (1998). Efficacy of risk-reduction counseling to prevent Human Immunodeficiency Virus and Sexually Transmitted Diseases. *Journal of American Medial Association, 280*(13), 1161-1167.

Kirby, D. (1999). Reflections on two decades of research on teen sexual behavior and pregnancy. *Journal of School Health, 69*(3), 89-94.

Kirby, D. (2001). *Emerging answers: Research findings on programs to reduce teen pregnancy*. Washington, DC: National Campaign to Prevent Teen Pregnancy.

Kirby, D. (2002). Effective approaches to reducing adolescent unprotected sex, pregnancy, and childbearing. *Journal of Sex Research, 39*(1), 51-57.

Land, K. C., McCall, P. L., & Williams, J. R. (1990). Something that works in juvenile justice: An evaluation of the North Carolina Court Counselors' Intensive Protective Supervision randomized experimental project. *Evaluation Review, 14*(6), 574-606.

Leigh, W. A., & Andrews, J. L. (2002). *The reproductive health of African American adolescents: What we know and what we don't know*. Washington, DC: Joint Center for Political and Economic Studies.

Lipsey, M. W. (1992). Juvenile delinquency treatment: A meta-analytic inquiry into the variability of effects. In S. Garrett (Ed.), *Meta-analysis for explanation: A casebook* (pp. 83-127). New York: Russell Sage Foundation.

Love, J. M., Schochet, P. Z., & Meckstroth, A. L. (1996). *Are they in any real danger? What research does-- and doesn't-- tell us about child care quality and children's well-being* (Child Care Research and Policy Papers). Princeton, NJ: Mathematica Policy Research, Inc.

MacKenzie, D. (1998). Criminal justice and crime prevention (Chap. 9). In S. Bushway (Ed.), *Preventing crime: What works, what doesn't, what's promising*. Washington, DC: US Department of Justice.

McCartney, K. (1984). Effect of quality of day care environment on children's language development. *Developmental Psychology, 20*(2), 244-260.

McCurdy, K., Gannon, R. A., & Daro, D. (2003). Participation patterns in home-based family support programs: Ethnic variations. *Family Relations, 52*(1), 3-11.

McGill, D. E., Mihalic, S., & Grotpeter, J. K. (2001). *Blueprints for violence prevention: Big Brothers/Big Sisters program*. Boulder: Center for the Study and Prevention of Violence, University of Colorado.

Metz, E., McLellan, J., & Youniss, J. (2003). Types of voluntary service and adolescents' civic development. *Journal of Adolescent Research, 18*(2), 188-203.

Michelson, E., Zaff, J. F., & Hair, E. C. (2002). *Civic engagement programs and youth development: A synthesis* (Report for the Edna McConnell Clark Foundation). Washington, DC: Child Trends, Inc.

Mihalic, S., Fagan, A., Irwin, K., Ballard, D., & Elliott, D. (2002). *Blueprints for violence prevention replications: Factors for implementation success*. Boulder: Center for the Study and Prevention of Violence, University of Colorado.

Miller, B. M. (2003). *Critical hours: Afterschool programs and educational success*. Quincy, MA: Nellie Mae Education Foundation.

Moore, K., Vandivere, S., & Redd, Z. (2003). *Cumulative measures of stress, risk, turbulence and community connectedness: Item choice, reliability and validity*. Washington, DC: Child Trends.

Moore, K. A., & Zaff, J. F. (2002). *Building a better teenager: A summary of "what works" in adolescent development* (Research brief). Washington, DC: Child Trends.

Morrow, L. M., & Young, J. (1997). A family literacy program connecting school and home: effects on attitude, motivation, and literacy achievement. *Journal of Educational Psychology, 89*(4), 736-742.

Mostellar, F. (1995). The Tennessee study of class size in the early school grades. *The Future of Children: When school is out, 5*(2), 113-127.

Murray, D. M., Richards, P. S., Luepker, R. V., & Johnson, C. A. (1987). The prevention of cigarette smoking in children: Two- and three-year follow-up comparisons of four prevention strategies. *Journal of Behavioral Medicine, 10*(6), 595-611.

National Council on Crime and Delinquency. (1987). *The impact of juvenile court intervention*. San Francisco, CA: Author.

National Research Council, & Institute of Medicine. (2000). *From neurons to neighborhoods: The science of early childhood development*. Committee on Integrating the Science of Early Childhood Development. J. P. Shonkoff & D. A. Phillips (Eds.), Board on Children, Youth, and Families, Commission on Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

National Research Council, & Institute of Medicine. (2002). *Community programs to promote youth development*. Washington, DC: National Academy Press.

National Youth Development Information Center. (2000). *A positive focus including all youth*. Retrieved November 27, 2002, from the World Wide Web: www.nydic.org/nydic/positive.html

Nelson, G. D., Poehler, D., & Johnson, L. (1988). Implementation of the teenage health teaching modules: A case study. *Health Education Research, 19*, 14-18.

Newman, R. P., Smith, S. M., & Murphy, R. (2001). *A matter of money: The cost and financing of youth development in America*. Washington, DC: Academy for Educational Development.

Nye, B., Hedges, L. V., & Konstantopoulos, S. (1999). The long-term effects of small classes: A five-year follow-up of the Tennessee Class Size Experiment. *Educational Evaluation and Policy Analysis, 21*(2), 127-142.

Olds, D. L., Robinson, J., O'Brien, R., Luckey, D. W., Pettit, L. M., Henderson, C. R., Ng., R. K., Sheff, K. L., Korfmacher, J., Hiatt, S., & Talmi, A. (2002). Home visiting by paraprofessionals and by nurses: a randomized, controlled trial. *Pediatrics, 110*(3), 486-496.

Peisner-Feinberg, E. S., & Burchinal, M. R. (1997). Relations between preschool children's child-care experiences and concurrent development: The cost, quality, and outcomes study. *Merrill-Palmer Quarterly, 43*(3), 451-477.

Peisner-Feinberg, E. S., Burchinal, M. R., Clifford, R. M., Culkin, M. L., Howes, C., Kagan, S. L., & Yazejian, N. (2001). The relation of preschool child-care quality to children's cognitive and social development trajectories through second grade. *Child Development, 72*(5), 1532-1553.

Perry, C. L., Telch, M. J., Killen, J., Burke, A., & Maccoby, N. (1983). High school smoking prevention: The relative efficacy of varied treatments and instructors. *Adolescence, 18*(71), 561-566.

Philliber, S. (1999). In search of peer power: A review of research on peer-based interventions for teens. In S. Philliber (Ed.), *Peer Potential: Making the Most of How Teens Influence Each Other. Washington, DC: The National Campaign to Prevent Teen Pregnancy*. Washington, DC: The National Campaign to Prevent Teen Pregnancy.

Pierce, K. M., Hamm, J. V., & Vandell, D. L. (1999). Experiences in after-school programs and children's adjustment in first-grade classrooms. *Child Development, 70*(3), 756-767.

Pittman, K. J., Irby, M., Tolman, J., Yohalem, N., & Ferber, T. (2001). *Preventing problems, promoting development, encouraging engagement: Competing priorities or inseparable goals?* Takoma Park, MD: The Forum for Youth Investment.

Project MATCH Research Group. (1998). Matching patients with alcohol disorders to treatments: Clinical implications from Project MATCH. *Journal of Mental Health (UK), 7*(6), 589-602.

Quinn, J. (1999). Where need meets opportunity: Youth development programs for early teens. *The Future of Children: When school is out, 9*(2), 96-116.

Redd, Z., Cochran, S., Hair, E. C., & Moore, K. A. (2002). *Academic achievement programs and youth development: A synthesis* (Report for the Edna McConnell Clark Foundation). Washington, DC: Child Trends, Inc.

Reid, W. J., & Bailey-Dempsey, C. (1995). The effects of monetary incentives on school performance. *Families in Society: The Journal of Contemporary Human Services, 76*(6), 331-340.

Reid, W. J., Bailey-Dempsey, C., Cain, E., Cook, T. V., & Burchard, J. D. (1994). Cash incentives versus case management: Can money replace services in preventing school failure? *Social Work Research, 18*(4), 227-236.

Rhodes, F., & Malotte, C. K. (1996). HIV risk interventions for active drug users. In S. C. Thompson (Ed.), *Understanding and Preventing HIV Risk Behavior: Safer Sex and Drug Use* (pp. 211-214). Thousand Oaks, CA: Sage Publications Inc.

Rodriguez, E., Hirschl, T., Mead, J., & Goggin, S. (1999). *Understanding the difference 4-H clubs make in the lives of New York youth: How 4-H contributes to positive youth development.* (Unpublished manuscript). Ithaca, NY: Cornell University.

Rosenthal, R., & Vandell, D. L. (1996). Quality of care at school-aged child-care programs: Regulatable features, observed experiences, child perspectives, and parent perspectives. *Child Development, 67*(5), 2434-2445.

Ross, J. G., Nelson, G. D., & Kolbe, L. J. (1991). Teenage health teaching modules evaluation (Special issue). *Journal of School Health, 61*(1).

Schorr, L. (1992). *Daring to learn from our successes: Services, supports, and systems that improve outcomes for children (Paper for discussion at the meeting of the Aspen Institute Domestic Strategy Group: Children and Families)*.Unpublished manuscript, Cambridge, MA.

Sherman, L. W., Gottfredson, D. C., MacKenzie, D., Eck, J., Reuter, P., & Bushway, S. (1998). *Preventing crime: What works, what doesn't, what's promising* (Research brief). Washington, DC: Department of Justice, National Institute of Justice.

Shonkoff, J. P., & Phillips, D. A. (2000). *Early childhood intervention: Views from the field*. Washington, DC: National Academy Press.

Silva, M. (2002). The effectiveness of school-based sex education programs in the promotion of abstinent behavior: A meta-analysis. *Health Education Research, 17*(4), 471-481.

Simpkins, S. (2003). Does youth participation in out-of-school time activities make a difference? *The Evaluation Exchange, IX*, 2-3, 21.

Sipe, C. L. (1996). *Mentoring: A synthesis of P/PV's research: 1988-1995*. Philadelphia: Public/Private Ventures.

Sontheimer, H., & Goodstein, L. (1993). An evaluation of juvenile intensive aftercare probation: aftercare versus system response effects. *Justice Quarterly, 20*(2), 197-227.

St. Pierre, R. G., Gamse, B., Alamprese, J., Rimdzius, T., & Tao, F. (1998). *Even Start: Evidence from the past and a look to the future*. Washington, DC: U.S. Department of Education, Planning and Evaluation Service.

St. Pierre, R. G., & Layzer, J. I. (1996). *Improving family literacy programs: Lessons from other intervention programs*. Cambridge, MA: Abt Associates.

St. Pierre, R. G., Swartz, J. P., Gamse, B., Murray, S., Deck, D., & Nickel, P. (1995). *National evaluation of the Even Start family literacy program: Final report*. Cambridge, MA: Abt Associates, Inc.

Sullivan-Temple, K., & Ravid, R. (1991). *Marvelous Me - Preschool edition: Enhancing self-concept development in preschool children.* Paper presented at the Annual meeting of the Mid-Western Educational Research Association, Chicago, IL.

Telch, M. J., Miller, L. M., Killen, J. D., Cooke, S., & Maccoby, N. (1990). Social influences approach to smoking prevention: The effects of videotape delivery with and without same-age peer leader participation. *Addictive Behaviors, 15*(1), 21-28.

The Forum for Youth Investment. (2003). *Out-of-school time policy commentary #2: High school after-school: What is it? What might it be? Why is it important?* Washington, DC: Author.

Tobler, M. S. (1986). Meta-analysis of 143 adolescent drug prevention programs: Quantitative outcome results of program participants compared to a control or comparison group. *Journal of Drug Issues, 16*(4), 537-567.

Tobler, M. S. (1992). Drug prevention programs can work: Research findings. *Journal of Addictive Diseases, 11*(3), 1-28.

Tobler, M. S., Roona, M. R., Ochshorn, P., Marshall, D. G., Streke, A. V., & Stackpole, K. M. (2000). School-based adolescent drug prevention programs: 1998 meta-analysis. *Journal of Primary Prevention, 20*(4), 275-337.

U.S. Department of Education, & U.S. Department of Justice. (2000). *Working for children and families: Safe and smart after-school programs*. Washington, DC: U.S. Department of Education.

Vandell, D. L., & Shumow, L. (1999). After-school child care programs. *The Future of Children: When school is out, 9*(2), 64-80.

Vandell, D. L., & Wolfe, B. (2000). *Child care quality: Does it matter and does it need to be improved?* Washington, DC: Office of the Assistant Secretary for Planning and Evaluation, US Department of Health and Human Services.

Walton, E., Fraser, M. W., Lewis, R. E., Pecora, P. J., & Walton, W. K. (1993). In-home family-focused reunification: an experimental study. *Child Welfare, 72*(5), 473-488.

Wasik, B. A., & Karweit, N. L. (1994). Off to a good start: Effects of birth to three interventions on early school success. In B. A. Wasik (Ed.), *Preventing early school failure: Research, policy, and practice*. Needham Heights, MA: Allyn and Bacon.

Webster-Stratton, C., Reid, M. J., & Hammond, M. (2001). Preventing conduct problems, promoting social competence: A parent and teacher training partnership in Head Start. *Journal of Clinical Child Psychology, 30*(3), 283-302.

Wiebush, R. G. (1991). *Evaluation of the Lucas County Intensive Supervision Unit: Diversionary impact and youth outcomes, final report*. Toledo, OH: Lucas County Juvenile Court.

Winett, R. A., Anderson, E. S., Moore, J. F., Taylor, C. D., Hook, R. J., Webster, D. A., Neubauer, T. E., Harden, M. C., & Mundy, L. L. (1993). Efficacy of a home-based Human Immunodeficiency Virus prevention video program for teens and parents. *Health Education Quarterly, 20*(4), 555-567.