The Journey to Evidence-Based Programming: Changing the Face of Social Services

Lisa Williams-Taylor, Ph.D.
CHILDREN'S SERVICES COUNCIL OF PALM BEACH COUNTY

In 1986 and again in 2000, the citizens of Palm Beach County recognized the widening gap between the growing needs of children and their families and the limited resources available to meet those needs. In response, they voted overwhelmingly to create a special district of local government -- the Children's Services Council of Palm Beach County -- to specifically support services for the county's most valuable: Our children.

Our Vision
Children's Services Council of Palm Beach County aspires to be an innovative leader creating a community where children and families reach their full potential.

Our Mission
The mission of the Children's Services Council of Palm Beach County is to enhance the lives of children and their families and to enable them to attain their full potential by providing a unified context within which children's needs can be identified and resolved by all members of the community. In order to achieve its mission, the Council will plan, develop, fund and evaluate programs and promote public policies which benefit Palm Beach County's children and families.

Our Goals
We help children begin life healthy, enter school free from abuse/neglect, eager and ready to learn; and thrive in quality afterschool programs so they can reach their potential and succeed in life.

Acknowledgments
With special thanks to Karen Deringer, Gabrielle Bargerstock, Tana Ebbole, Tanya Palmer, Tom Sheehan, Carol Rodriquez and Grace Watson at the Children's Services Council and Dr. Alan Brown for their review and comment. Also, thanks to Marlene Passell for editing the publication.

The Council:
Rod Macon, FP&L, Chair
Judge Ronald Alvarez, Juvenile Court Judge, Vice Chair
Dari Bowman, Child Advocate, Secretary
Perry Borman, Circuit 15 Administrator, Department of Children & Families
Vincent Goodman, Retired Educator
Carrie Hill, School Board Member
Art Johnson, Superintendent of Schools
Jeff Koons, County Commissioner
Gregory A. Langowski, Executive Director, Palm Beach County Republican Party
Willie A. Miller, Jr., Aide to County Commissioner Priscilla Taylor

Chief Executive Officer
Gaetana D. Ebbole
# Contents

Executive Summary .................................................................................................................. 1  
Preface .................................................................................................................................................. 3  
Chapter 1: Introduction .................................................................................................................. 5  
Chapter 2: Return on Investment .................................................................................................. 15  
Chapter 3: Considerations in Moving Toward Evidence-Based Programming ...................... 25  
Chapter 4: Barriers to Implementing Evidence-Based Programs ............................................. 35  
Chapter 5: Case Study: The Children's Services Council of Palm Beach County ......................... 47  
Chapter 6: Advocacy: Changing One's Community ................................................................... 69  
Conclusion ........................................................................................................................................ 101  
Appendix A ...................................................................................................................................... 103  
Appendix B ...................................................................................................................................... 107  
Appendix C ...................................................................................................................................... 109  
Appendix D ...................................................................................................................................... 110  
Appendix E ...................................................................................................................................... 131  
Appendix F ...................................................................................................................................... 133
Executive Summary

- Not all programs are equal in terms of quality and effectiveness. The term evidence-based has become the newest buzzword in social services and education. Many programs are calling themselves evidence-based because of the prominence and the funding attached to them. For this reason, it is imperative to review and understand the actual research that supports the program.

- Every organization that rates evidence-based programs uses different categories and criteria to decide levels of effectiveness. There is no consistency among raters. It is always important to inquire what criteria were used to make the designation as evidence-based or promising.

- When deciding on a program for an organization, key points need to be considered. These include: (1) making sure that the program meets the needs of the population and fits well with the mission and vision of the organization that will implement the program; (2) having a clear understanding of funding and other resources needed, such as staff education and training; (3) ensuring that the program can be implemented with fidelity and sustained long term; and (4) planning for continued evaluation and monitoring.

- If an evidence-based program exists that meets the needs of an organization and the population to be served, it should be the first choice of the organization. However, evidence-based should not be a “yes/no” concept. We should be thinking about evidence-based as a continuum. Most programs fit somewhere on that continuum and need to plan how they wish to move up in order to demonstrate effectiveness.

- While there are not enough evidence-based programs at this time to meet all of the needs within a community, funders and government should be thinking about how to move unproven programs along the continuum in order to increase accountability and ultimately introduce more evidence-based programs to the field. Some important elements of this continuum should include making sure that the program has a strong theoretical foundation, quality data collection, fidelity to the model, and an evaluation component (process and outcome).

- Implementing evidence-based programs may mean discontinuing programs that do not work. These types of changes require planning and educating the public, providers, and funders, among others. Being aware and prepared for negative pushback that may occur and the politics involved will be critical to the process.
Readiness for implementing evidence-based programs does not happen overnight. A solid infrastructure is needed. The process to move an organization toward evidence-based programming can take years. The needed components include:

- Leadership support
- A strong understanding of the needs of the community
- Outcome-focused decision-making
- Alignment of population needs (risk/protective factors), program outcomes, and the mission/vision of the organization
- A strong accountability/data infrastructure to ensure fidelity, monitoring, and evaluation of programs
- A supportive professional development system to meet the needs of the providers

Some evidence-based programs cost much more than the average program, but it is important to consider the value of the research that shows that they work. These programs eliminate the possibility of investing in a program that may show no outcomes three to five years after implementation. Thus, although initial costs may be more, money is guaranteed well spent.

Some of the most common barriers to implementing evidence-based programs include the lack of rigorously tested programs, resistance, the requirements to implement programs with fidelity, cost, professional development concerns, lack of organizational readiness, and political forces.
In these economic times it is increasingly critical that funders, providers, community stakeholders, and others carefully consider how to spend resources. Among some groups, social services have gotten a reputation for wasteful spending. While it is true that changing behavior and improving the lives of children and families is no small task, part of the damaging reputation is well deserved. Billions of dollars have been spent year after year with little demonstrated change. For example, the poverty rate has remained relatively unchanged over the past 40 years despite billions of dollars spent on programs. So, where does the field go from here? How can we ensure the biggest bang for the buck or return on investment? The answer is quite simple. Funds must first be spent on programs that have evaluations that show that they are successful. Second, because there are simply not enough programs available to date that meet the criteria for evidence-based, funding must be set aside for evaluation of locally developed, innovative programs in order to move programs along a continuum of evidence-based.

This book will walk the reader through evidence-based literature to date and provide a step-by-step look at the process of moving toward evidence-based programming.

- Chapter One provides an introduction to the concept of evidence-based programs, including the criteria for determining what works, an historical look over time, and why moving in this direction is so important.

- Chapter Two discusses return on investment in general and reviews cost-benefit studies completed on specific programs.

- Chapter Three reviews the steps for searching for effective programs and what should be considered when choosing programs that meet targeted outcomes and populations. It also includes a brief description of implementation and how to ensure long-term sustainability.

- Chapter Four reviews several of the barriers to implementation that are often cited by providers and critics. Examples include the lack of evidence-based programs, the uniqueness of communities and maintaining fidelity, the costs associated with these programs, staffing requirements, lack of organizational readiness, and political forces.
Chapter Five is a case study of how a local funding agency, the Children’s Services Council of Palm Beach County (CSC), began its journey toward evidence-based programming. This chapter includes the process that CSC went through to choose programs, as well as how it began thinking about moving locally developed programs along the evidence-based continuum.

Chapter Six provides a review of top-tiered programs that should be considered by communities nationwide.

Lastly, the appendices include a glossary of evidence-based terms, an example of a program logic model, a pathways framework, information pertaining to national rating systems and organizations, and a sample of a screening and assessment tool.
Chapter 1: Introduction

The need to implement proven programs, in other words, programs with track records showing their effectiveness, has never been greater. With shrinking budgets and increased pressure to demonstrate accountability, organizations and funders are being asked to show that they are implementing programs that will work, also known as evidence-based programs, and provide a return on investment. While the term evidence-based has become universally used, there is no common definition or universal set of criteria to determine whether a program fits into this category. However, most definitions of evidence-based do include common elements such as having a strong theoretical foundation, quality data collection and procedures, and it must show evidence of effectiveness. This concept of evidence of effectiveness is where there are major variations in criteria and definitions.

Evidence-Based Criteria
Many national experts would agree that for a program to show effectiveness there must be a strong research design testing the outcomes. This means using an experimental/randomized controlled trial (RTC) or well-matched quasi-experimental design. The experimental design is often referred to as the “gold standard” in research and the more RTCs a program has, the more confidence in the outcomes. A comparison point is the medical field. A drug goes through two randomized clinical trials before it is approved by the Federal Drug Administration and put on the market. It is important to note that this type of study is needed in order to say that a program caused specific outcomes to occur. Without an evaluation that compares a group that received the program or intervention with another group that did not (random assignment), it is impossible to determine whether or not the program/intervention caused some notable change in those receiving the intervention. This type of evaluation causes some ethical concerns because it requires withholding treatment from some individuals. Conversely, some argue that it is unethical to provide services that have not been proven successful. One research strategy is to place those eligible on a waiting list so that if you find that the program works, they are the first to receive it after the initial study is completed.

**Experimental Study (RCT)**

A research design where participants are randomly assigned to either an experimental group (treatment) or control group (no treatment/placebo). This allows the researchers to examine whether the intervention/treatment caused the outcomes or effect to take place (causal inference).

**Quasi-Experimental Study**

This research design is very similar to and almost meets criteria for an experimental design, but is unable to control all potential factors and does not include random assignment of participants. The key is to use a well-matched comparison group to eliminate as many outside factors as possible.

Well-matched comparison group studies are the next best design. This design is similar to the RCT, but does not control all potential factors and does not include random assignment of participants. The key is to use a well-matched comparison group to eliminate as many outside factors as possible. In other words, eliminating the major contributors that may impact results bolsters one’s confidence that the outcomes achieved are the result of the activities of the program, not some outside influences and that the changes did not happen by chance. For example, if you just measure children before and after they receive treatment then you cannot say that the gains they made
would not have occurred despite the intervention. Likewise, you may determine that a program is unsuccessful, but if the participants receiving services were compared to another group, it may well show a significant impact.

Another important piece of information to gather when examining evidence-based programs is whether the rigorous study itself has any flaws. Knowing what the limitations are and how the flaws in methodology may have impacted the findings is critical to deciding whether or not one can believe the results. This must be explored before one can examine the findings of a study. For example, determining whether an evaluation has an adequate sample size (meaning a sufficient number of research subjects received the intervention) in order to determine statistical significance should be considered. On the other hand, it is just as important not to have too many research subjects since an effect of little scientific importance can be noted as statistically significant by chance alone. It is also important to determine whether the researchers utilized appropriate measurement tools and analyses.

For a program to be considered evidence-based, researchers must show that the outcomes are positive and meaningful (measured by significance levels and effect sizes), sustainable over time, show that it has been replicated in more than one area or with different populations and that the outcomes continue to be positive. These four criteria are critical to labeling a program as evidence-based (See Glossary of Terms: Appendix A for more information on terms used in this section).
The term *evidence-based* is rather new and originated in the medical field. The evidence-based movement was thrust forward after a 1927 landmark decision to create the Food and Drug Administration, which is responsible for testing the safety of medical treatments (Leff, 2002). The second major influence was in the use of randomized, controlled studies. It was in 1948 that the first such study took place – researching the efficacy of streptomycin in treating tuberculosis. By the 1960s, the number of randomized controlled trials reached into the hundreds, and today there are tens of thousands occurring every day (Dodge, 2006).

During the mid-1970s, there was a general pessimistic view that “nothing works.” This resulted from research showing that many social programs were not having an impact. This view originated in the study of corrections, where a review of over 200 evaluations looking at treatment programs showed no or very little impact. However, the ideology began to change by the early 1990s when evaluation research demonstrated that, in fact, several prevention and treatment programs did produce significant results. As a result, federal, state and private organizations began to support the efforts to continue doing these types of evaluation studies. By the late 1990s, the emphasis on program outcomes became so diffuse that many practitioners found they could not secure funding unless they conducted or cooperated with outcome evaluation studies. It became common practice by the early 21st century that, in order for a program to acquire federal grant dollars, an evaluation plan must also be submitted. A recent focus of prevention and intervention programs has been identifying and replicating effective research-based programs. Today, these are typically referred to as evidence-based programs. In fact, while these terms were at one point used interchangeably, now evidence-based is used as a step above research-based, meaning that the research needs to be of a certain caliber – rigorous.

**History of Evidence-Based Programming**

While at one point, the terms research-based and evidence-based were used interchangeably, now evidence-based is used as a step above research-based, meaning that the research needs to be of a certain caliber – rigorous.

- **Psychology** - In the field of psychology, which does not have a governmental body examining the efficacy of treatments, it is the responsibility of those in the field to research effective programs. While some researchers began earlier, it was not until the 1990s that this idea began to expand. The Alcohol, Drug Abuse and Mental Health Reorganization Act of 1992 helped create the Substance Abuse and Mental Health Services Administration (SAMHSA), whose role is to assist in disseminating research
and effective programs/services addressing problem behaviors. In 1999, the American Psychological Association established a taskforce for the primary purpose of promoting scientific treatments, also termed *empirically supported treatments* (Dodge, 2006, p.477). The task force wanted to advocate for improving patient outcomes by using research and evidence, much like what happened years earlier in the medical field. It was during this time that a backlash began with some psychologists pushing against these treatments. They believed that “it infringes on their autonomy and dehumanizes clients” (Dodge, 2006, p.477). There was concern that clients vary too much in regard to disorders, co-morbidity, personality, race, ethnicity, and culture to use a one-size-fits-all, “cookie-cutter” approach (Levant, 2005). This is still a concern of providers and some individuals in the field. However, with continued research showing that many of these programs have been successful with various ethnicities and racial groups, as well as that clinical practice should be based on evidence, the tides are slowly changing.

**Education** - In education, the No Child Left Behind Act of 2002 was the first major move by the education field to promote evidence-based programs. This law affects children in kindergarten through high school and stresses accountability for results and emphasizes implementing programs and practices based on scientific research. The Department of Education uses specific criteria noted in the Act to define evidence-based programs, which include:

- “Research that involves the application of rigorous, systematic, and objective procedures to obtain reliable and valid knowledge relevant to education activities and programs;
- Data analysis adequate to test and justify the general conclusions drawn;
- Measurements or observational methods that provide reliable and valid data across evaluators, observers, multiple measurements and observations, and studies;
- Evaluated using experimental or quasi-experimental designs in which individuals, entities, programs, or activities are assigned to different conditions and with appropriate controls;
- Experimental studies are presented in sufficient detail and clarity to allow for replication or, offer the opportunity to build systematically on their findings; and
- The research has been accepted by a peer-reviewed journal or approved by a panel of independent experts through a comparably rigorous, objective, and scientific review” (No Child Left Behind Act, 2002, pp.1964-65).
The Journey to Evidence-Based Programming

• **Violence Prevention** - Prevention science has been the last discipline to welcome evidence-based programs; until recently, the focus has been on intervention. In 1996, the Center for the Study and Prevention of Violence in Colorado began examining various youth programs to determine which one’s worked to reduce crime and violence. Blueprints for Violence Prevention is a project of the center. The main goal is to review and recommend top programs with evidence of effectiveness. Some of the core criteria examined included research design, effects, sustainability of outcomes, and replicability.

As can be seen from the previous examples, the areas of substance use, mental health, and juvenile justice have been working toward using evidence-based programs for the past 15 years, but a systematic review of programs in the primary prevention areas, such as maternal health and early care and education, is in its infancy.

**Why is Evidence-Based Programming Important?**

While human service funders finance many successful programs and providers deliver many effective services, as a field, we also know that funders overfund programs with little evidence of effectiveness, underfund effective programs, and fund programs that continue fragmentation of services. Human services have a history of investment in untested programs, based on questionable assumptions, delivered with little consistency or quality control, and without effective evaluation.

The reality is that there has been little success (or evidence of success) to date, despite a multitude of programs and billions of dollars. As a nation, we are, in effect, spending vast sums of money on programs that are ineffective at best and potentially harmful. A

> “…it is not an exaggeration to believe that we can save this nation billions of dollars over the coming decade. If we apply ourselves well, if we do good research, and if we apply that research to public policy…”

- Brian Baird
U.S. House of Representatives

As a nation, we are, in effect, spending vast sums of money on programs that are ineffective at best and potentially harmful. A program may appear on the surface to work and logically should work, but when formally evaluated it may show no results or may in fact be harmful to the population it serves.
program may appear on the surface to work and logically should work, but when formally evaluated it may show no results or may in fact be harmful to the population it serves. (e.g. the Scared Straight program). In fact, some of the most widely used crime, substance abuse and education interventions have been shown to be ineffective or harmful. For instance, according to Baron (2008):

- Widely used crime prevention programs, such as boot camps have been found ineffective and Scared Straight programs have been determined harmful to participants;
- The U.S. has made no significant progress in preventing drug/alcohol abuse since 1990 and commonly used programs in schools, such as Drug Abuse Resistance Education (DARE) and Project Alert, have been found to be ineffective in preventing substance use;
- The U.S. has made almost no progress in raising the K-12 educational achievement over the past 30 years according to the respected National Assessment of Educational Progress despite a 90% increase in public spending per student; Many programs aimed at increasing academic outcomes (select 21st Century Community Learning Centers for students in elementary school) have demonstrated no effect and have shown adverse effects on student behavior.

This is not only a problem for social services. Many fields have implemented programs and practices that when rigorously studied, have shown ineffective or harmful effects. Baron (2008) provides the following examples:

- Vouchers used for disadvantaged workers to help subsidize their employment – a well-designed randomized trial found large negative effects on employment;
- Intensive efforts to lower blood sugar of diabetics to normal levels – one trial found that it increases the risk of death;
- Hormone replacement therapy for post-menopausal women – one trial found that it increased risk of stroke and heart disease for many women;
- Dietary fiber to prevent colon cancer – a trial found the practice to be ineffective;
- Stents to open clogged arteries – a rigorous trial has shown the procedure is no better than drugs for most heart patients;
- Having babies sleep on their stomachs – later trials showed that it increased the risk of Sudden Infant Death Syndrome;
- Beta-carotene and Vitamin E supplements (“anti-oxidants”) to prevent cancer – trials have shown them to be ineffective and harmful in some cases;
• Oxygen-rich environment for premature infants – later research showed that it increased the risk of blindness; and
• Recent promising AIDS vaccine – research found it to double the risk of AIDS infection.

Thus, because so many programs are being implemented without sound research showing effectiveness, and because much of the conventional wisdom about “what works” is probably wrong, practitioners and funders must turn to evidence-based programming. According to Elliott (2007), the benefits of using evidence-based programs are simple:

• Ethics - avoiding harmful effects;
• Stronger and more consistent positive outcomes;
• Improving the well-being of our children; and
• Cost savings to taxpayers.

If prevention programming does not begin to move in this direction, vast sums of money will be wasted and, more importantly, the very people programs are meant to help may be harmed as shown in the above examples.

Baron, J. (2008). Personal communication on April 4, 2008 – Quoted from a presentation provided to the Children's Services Council of Palm Beach County.


Elliott, D., (2007). Personal communications on January 11, 2007. Quoted from a presentation delivered at the Senior Executive Policy Institute, a meeting hosted by the Children's Services Council of Palm Beach County.


Chapter 2: Return on Investment

Evidence-based programming helps ensure that agencies are spending resources on proven programs that work. We must be accountable to the families we serve, as well as to community stakeholders, funders, and taxpayers (Hyde, Falls, Morris & Schoenwald, 2003). Implementing evidence-based programs assists agencies and organizations in moving towards accountability. The fact remains that, over the last few decades, service providers and funders have spent billions of dollars on social programs and when you examine the data – there has been minimal improvement in lessening or eradicating social ills. For this reason, there is an increasing demand that funding be spent on programs with demonstrated outcomes, and that show a good return on investment. This means either implementing programs that have already been labeled evidence-based through a national process or moving locally developed or homegrown programs along a continuum of evidence-based (See Chapter 5 for an example). Communities must move in the direction of proving that their programs work (i.e. that they are effective for the children and families they serve).

Concerns & Considerations

One concern often raised is that implementing an already existing evidence-based program can be costly. While this is not always true (e.g. LifeSkills training [LST] and Teen Outreach Program [TOP]), some evidence-based programs do cost more than the average. However, it is important to consider the value of the research that shows that they work. These programs eliminate the possibility of investing in a program that may

“In times of shrinking budgets and increasing federal and state deficits, policymakers and practitioners must make efficient use of prevention resources by opting for programs that have the greatest likelihood of producing positive effects” (Kyler, Bumbarger, & Greenberg, 2005).
show no outcomes three to five years after implementation. Thus, although initial costs may be more, money is guaranteed well spent and will eventually show a return on the dollar.

As some funders have noted, if you look at what you have spent over the past five to ten years on programs that do not work, costs become less of a concern. Furthermore, the sustained effectiveness of evidence-based programs and what society will pay in later costs if we do not invest is critical to consider.

The costs of evidence-based programs and services are also routinely questioned because the outcomes from these services may not be immediate – sometimes, the highest return on investment comes many years down the road. For example, a cost benefit analysis of the High/Scope Perry Preschool Project, a high-quality preschool for disadvantaged children was completed when the children were 27-years-old. The program showed a $7.16 return for every dollar invested. The developers then conducted another cost-benefit analysis when the children were 40-years-old, which showed that for every dollar invested, there was a $17 savings in later costs. While many studies and cost benefit analyses have been completed on evidence-based or proven programs, questions and reservations continue. Furthermore, with limited funding, there is a debate about where to allocate funding. Research indicates that investing in prevention and early intervention programs yields the best outcomes for those at risk, as well as the best
return on the dollar. The earlier you impact a child or family when they are experiencing problems - by either promoting protective factors or mitigating risk factors - the best probable outcomes are achieved. However, funding will continue to be needed for those children still engaged in problematic or high-risk behaviors.

While the debate will continue, the information herein will provide a snapshot of the possible return on investment that can be expected if a decision is made to move forward with implementing proven programs. Although this review is not a formal return on investment or cost-benefit analysis and may not resolve all concerns, it will provide some context to the ongoing discussion and debate. It is meant to provide a glimpse at data and research so that relationships between outcomes, costs, and program impact can be inferred.

While an organization may target various outcomes, some of the most common outcomes discussed in the cost-benefit literature are teen pregnancy, prematurity, low birthweight, child maltreatment, and juvenile crime. The following data helps to show the need for evidence-based programs that target these specific outcomes. In addition, a sampling of cost-benefit and return on investment studies on specific programs that target these outcomes is reviewed.

Data

Teen Pregnancy: The annual public cost of teen childbearing (adolescents 19 years and younger) is approximately $9.1 billion nationally (Hoffman, 2006). This $9.1 billion includes the health care needed for the children ($1.9 billion), foster care ($2.3 billion), later incarceration of the male children ($2.1 billion), tax revenue losses from lower earnings of the mother, father, and child as an adult ($6.3 billion), and money for offsetting public assistance needs of the mother ($3.6 billion)\(^1\). In 2004 dollars, the average annual public sector cost of each child born to a mother 17 years and younger was $4,080 (Hoffman, 2006).

Prematurity: The annual cost for preterm deliveries is at least $26 billion (Behrman & Butler, 2007). This includes money needed for medical care, early intervention services, special educational needs, and losses in productivity, including both household and labor market productivity. According to the National Governors Association Center for Best Practices (2004), a hospital charge alone for premature babies averages $75,000 per child.

\(^1\)Costs are not mutually exclusive.
Low Birthweight: The annual cost for low birthweight babies is $6 billion, of which 75% is spent during the first year on hospitalizations (Bernstein, 2000). The initial hospitalization costs for low birthweight babies (<2,500 grams/<5.8 pounds) have been estimated at $20,600 for low birthweight and $52,300 for very low birthweight (<1,500 grams/<3.4 pounds) deliveries (Russell et al., 2007). The health care costs in the first year of life for low birthweight babies are, on average, $15,000 higher than those for normal weight babies.

Costs for very low birthweight babies can soar up to more than $100,000. Each normal birthweight that occurs instead of a very low weight birth saves $59,700 in the first year of care. “Even among babies weighing 2,000-2,100 grams, who have comparatively low mortality rates, an additional pound (454 grams) of weight is still associated with a $10,000 difference in hospital charges for inpatient services” (Almond, Chay & Lee, 2004). These costs are only those associated with the immediate first year of care for these infants and do not include the added risks that many low birthweight infants experience (i.e. “chronic health and developmental problems such as cerebral palsy, brain damage, chronic lung and liver disease, deafness, blindness, epilepsy, learning disabilities, and attention deficit disorder)” (O’Connor, 2004, p.2).

Child Abuse and Neglect: The annual cost (both direct and indirect) of child abuse and neglect in 2007 was approximately $103.8 billion (Wang & Holton, 2007). The direct costs of $33.2 billion include legal expenses, foster care, law enforcement expenditures, and money needed for health care services. The remaining $70.6 billion are indirect costs of the long-term impact, including special education needs, juvenile delinquency and adult criminal justice system expenditures, mental health and other health care expenses, and losses in productivity (Wang & Holton, 2007). While overall economic impacts are important, costs per child or case are equally important, but more difficult to measure given the range of impacts. For example, the costs for a child needing hospitalization was approximately $19,266 in 1999 and the costs for needing mental health services varied by type of abuse (e.g. physical abuse - $2,700; sexual abuse - $5,800) (Wang & Holton, 2007). Children may need single services or multiple services. However, one estimate indicated that the average direct cost per child abuse case was $17,319 in 1999 dollars and indirect costs was $40,143 (Conrad, 2006). Lastly, the total opportunity costs (i.e. lifetime income lost, lifetime federal tax payments lost) were approximately $299,910 per case in the 1980s and would be much higher now (Daro, 1988).

While monetary costs are notable, many of the long-term costs associated with abuse and neglect are intangible – “it is impossible to calculate the impact of the pain, suffering, and reduced quality of life that victims of child abuse and neglect experience” (Wang & Holton, 2007, p.2).
Juvenile Delinquency: In 1997 dollars, the estimated range of costs for a typical juvenile delinquent ages 14 to 17 was between $83,000 and $335,000 (Cohen, 1998). These costs include everything from medical costs to lost wages to pain and suffering endured by the victim (Cohen, 1998). This cost range includes $62,000 to $250,000 for victim costs and $21,000 to $84,000 for criminal justice related costs. The range is quite large because the number of crimes committed and types of crime vary considerably. For example, the cost of a robbery will be different than the cost of burglary or drug related crimes. Researchers reviewed various studies and found that the combined cost to the victim and criminal justice costs was $3,409 for burglary, $32,483 for armed robbery, $41,130 for a serious assault, $116,994 for rape and sexual assault, and $3,973,625 for a murder (Cohen & Piquero, 2007). The typical juvenile delinquent commits between one and four crimes annually between the ages of 14 and 17 (Cohen, 1998).

Programs

Some evidence-based programs have had a cost-benefit analysis or return on investment analysis completed. This information provides additional context regarding the importance of this type of programming and indicates that the initial price tag should not be singled out without giving weight to longer-term outcomes and savings. The following are a few examples of programs that are labeled as evidence-based because they have data that proves they make a difference, as well as data that shows that the value of benefits outweighs the costs. These programs target teen pregnancy, child maltreatment, academic success, and juvenile crime.

Teen Outreach Program: Aos and colleagues (2004) at the Washington Institute of Public Policy conducted a cost-benefit analysis of prevention and early intervention programs for the State of Washington. They assessed many teen pregnancy prevention models, the most commonly known include: Reducing the Risk Program, Teen Talk, Children's Aid Society - Carrera Project and Teen Outreach Program. Of all these models, the only program to show a return on investment was the Teen Outreach Program. All other models were either not effective (yielding $0 in benefits) or the costs outweighed the benefits.

The benefits of Teen Outreach Program (TOP) equate to about $801 per youth. The estimated cost of the program is $620 per youth. Thus, there is a $181 return or approximately $1.29 for every dollar spent. This is the return when examining teen pregnancy alone. The benefits are even higher when other outcomes are factored in: “One study of TOP participants across several sites found they had significantly lower levels of suspension, school dropout, and pregnancy” (Lehr, Johnson, Bremer, Cosio
& Thompson, 2004, p. 60). When these other outcomes are considered along with the long-term impacts, the potential net benefit for boys was $13,500 and $12,749 for girls (Harrison, Juris, Stern & Stern).

**Nurse-Family Partnership (NFP):** A study conducted by the Washington Institute of Public Policy (2009) explored the cost benefits of criminal justice programs for the State of Washington. They found that the total benefits of NFP on crime reduction alone is $20,756 per family (a decrease in crime reduction of the mothers in the program results in $8,189 in benefits and the reduction among the children later in life produces $12,567 in benefits. The estimated cost of the program is $6,336 per family. An analysis indicates that the program reduces criminal behavior of the mothers participating in the program by 38.2%. Longitudinal studies have been completed which show that the crime rates of children born to mothers participating in the program are lowered by 15.7% compared to those not participating.

NFP is not only effective in decreasing crime rates. It has shown positive effects in all of the following domains - cognitive/achievement, behavioral/emotional, child maltreatment, health, accidents, injuries, and crime. When these outcomes are taken into account, the program has an overall cost-benefit ratio of $2.88 for every dollar spent. However, Karoly et al. (1998) took this cost-benefit analysis a step further and estimated results separately for both high- and low-risk families. They found that the net benefits per child were almost 18 times greater for the high-risk sample than for the low-risk group ($34,148 versus $1,880), with a benefit cost ratio of 5.70 compared to 1.26, respectively. This is important and shows that targeting the program to the appropriate population is critical. NFP is solely for first-time, low-income, high-risk mothers. Implementing it as designed for the right population will be critical to getting the highest return on investment.

**Multidimensional Treatment Foster Care:** The Washington Institute of Public Policy (2009) reported that the total benefit of this program on crime reduction was $88,953 per youth involved in the program, one of highest returns. The program costs approximately $6,926 per participant. Based on prior research, the program is projected to decrease recidivism by 17.9%, which produces $95,879 in benefits ($69,519 in saving to victims and $26,360 in savings to taxpayers).

**Multisystemic Therapy:** Research from the Washington Institute of Public Policy (2009) shows that the benefit of this program is $17,694 per youth over time given its projected decrease in crime of 7.7%. The cost per youth is approximately $4,364 and it produces $22,058 in total benefits ($15,001 in savings to victims and $7,057 in savings to taxpayers). In the longest longitudinal follow-up study completed to date (13.7 years
after participation), researchers Schaeffer & Borduin (2005) found that, in a randomized clinical trial, participants had significantly lower recidivism rates (reduced incarceration by an average of 62.4 days per/year). Furthermore, in a study that explored the cost benefit with this population, researchers found that there was a cost savings of at least $50,000 for each youth (Merbler, Borduin, & Schaeffer, 2004).

**Functional Family Therapy:** The total benefit of this program for the State of Washington as per a study conducted by the Washington Institute of Public Policy (2009) was $49,776 per youth involved in the program (totals benefits minus cost per juvenile). The program costs approximately $2,380 per juvenile delinquent. Based on prior research, the program is projected to decrease recidivism by 18.1%, which produces $52,156 in benefits per participant ($35,470 in saving to victims and $16,686 in savings to taxpayers).
References


Chapter 3: Considerations in Moving Toward Evidence-Based Programming

Once a decision has been made to move in the direction of evidence-based programs, many issues must be considered. Some of most significant considerations are:

- Choosing the right type of program;
- Choosing a specific program; and
- Determining what is needed for implementation and long-term sustainability.

These last stages (implementation and sustainability) are time consuming and difficult – but absolutely critical to success. They include assessing readiness for change and managing the change process – both willingness to change and ability to implement new programs. Without dedicated and quality implementers, there is no program.

Choosing the Right Type of Program: Outcomes & Target Population

When deciding the type of program you wish either to fund or implement, it is important to have a clear understanding of what outcomes and population you wish to target. Programs may target and achieve results for a single outcome or multiple outcomes. Are you looking for a violence prevention program, substance use prevention program, or both? During this process it is important to explore not only the end goals you are trying to achieve, but also the short-term and intermediate outcomes. You should begin thinking about what needs to be in place in order to get to the end goal. Program selection should be done using sound data and system logic model or pathways document (See Appendix B for an example), which will specify what needs to be done in order to achieve the targeted outcomes or goals. These types of documents lay out target population, activities needed to impact outcomes, as well as immediate, intermediate, and long-term outcomes. For example, if the end goal is decreasing child maltreatment, the following figure outlines some outcomes/indicators that you may want to target, which will impact what programs you chose. Program logic models are also important to review when working to align programming and desired outcomes/goals (See Appendix C for an example).
The Journey to Evidence-Based Programming

**GOAL:** Decreasing child maltreatment

Choose a program that targets

- Understanding of children's needs and child development
- Parenting skills and education regarding appropriate discipline
- Parental stress and mental health conditions
- Social isolation
- Parent-child relationships, including bonding and attachment
- Substance use or violence within family

*Source: Centers for Disease Control and Prevention*

The next consideration is the target population. Is the program trying to reach youth, adults, or everyone? Data will help inform what the needs are in the community, who most needs the services and the level of services needed. A decision must be made whether you are targeting populations based on risk and protective factors, that is the factors that make one either less or more vulnerable or whether you wish to provide more general education or awareness to an entire population (i.e. universal, selective, or indicated populations). According to the U.S. Department of Health and Human Services (2008) *universal* prevention programs address the entire population regardless of risk; *selective* prevention programs target subsets of the total population that are considered at increased risk, and *indicated* prevention programs provide services to those individuals who have risk factors that place them at very high-risk. For example, if your outcome is reducing infant mortality, you may choose to invest in a program that targets a specific population in a particular area where the infant mortality rate is higher.
Types of Prevention Programs

<table>
<thead>
<tr>
<th>Type</th>
<th>Target</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universal</td>
<td>General population of children and families</td>
<td>Less expensive – does not require specific recruitment strategies</td>
<td>Lack of attendance or involvement</td>
</tr>
<tr>
<td>Selective</td>
<td>Offered to at-risk youth and families-based on risk factors</td>
<td>Targeted to those who are likely to need a prevention or intervention program</td>
<td>Criteria is often broad and not easy to determine who should receive services</td>
</tr>
<tr>
<td>Indicated</td>
<td>Children already manifesting identified risk factors</td>
<td>Enhances the appropriate targeting of limited prevention – more intensive</td>
<td>Resources – requires highly trained, sensitive staff that are paid more for service delivery</td>
</tr>
</tbody>
</table>

Source: Brown, A. (2007). Personal communications. Quoted from a training delivered to provider staff hosted by the Children's Services Council of Palm Beach County.

Choosing a Specific Program

Finding the right program is not an easy task and requires various steps of research at different stages of the process. After reviewing data and making a decision about the target population and outcomes, you should begin to search for the program that fits best. There are many resources available to help with the search for an evidence-based program; however, not all resources are equal. There are many organizations rating programs and each has different criteria (See Appendix D for examples). For example, Blueprints Violence Prevention uses *model* and *promising* to rate programs, the Office of the Surgeon General uses *model, promising* and *does not work*, and the Prevention Research Center for the Promotion of Human Development (Penn State University) uses *effective* and *promising*. Each of these organizations not only uses different groupings, but when they use the same term “promising,” criteria is vastly different. As shown by the following example, for a program to be labeled as *promising* by the first two organizations,
a rigorous evaluation must have been completed, but not for the third organization. This type of discrepancy among definitions causes confusion for funders and service providers and can lead to miscommunication among parties discussing evidence-based programming. A critical question to ask when someone says a program is evidence-based is: “What criteria are you using to determine that label?”

**Example of Criteria for a Rating of Promising**

- Blueprints Violence Prevention – Programs that show a deterrent effect using either an experimental or quasi-experimental design;

- Office of the Surgeon General – Rigorous experimental design (experimental or quasi-experimental); Significant deterrent effects on: Violence or serious delinquency (Level 1) or any risk factor for violence with an effect size of .10 or greater (Level 2); and either replication or sustainability of effects;

- Prevention Research Center for the Promotion of Human Development (Penn State University) – Programs that appear promising, but are not proven, meaning they lack a controlled design, contain very small samples, or have findings that are indirectly related to mental health outcomes.

The following are a few recommended organizational websites to explore when reviewing programs, but one must still consider the criteria:

<table>
<thead>
<tr>
<th>Coalition for Evidence-based Policy</th>
<th><a href="http://evidencebasedprograms.org/wordpress">http://evidencebasedprograms.org/wordpress</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Blueprints for Violence Prevention</td>
<td><a href="http://www.colorado.edu/cspv/blueprints">http://www.colorado.edu/cspv/blueprints</a></td>
</tr>
<tr>
<td>Promising Practices</td>
<td><a href="http://www.promisingpractices.net/default.asp">http://www.promisingpractices.net/default.asp</a></td>
</tr>
<tr>
<td>OJJDP Model programs Guide</td>
<td><a href="http://www2.dsgonline.com/mpg/">http://www2.dsgonline.com/mpg/</a></td>
</tr>
<tr>
<td>Strengthening America's Families</td>
<td><a href="http://www.strengtheningfamilies.org/">http://www.strengtheningfamilies.org/</a></td>
</tr>
</tbody>
</table>
Other considerations when exploring programs is the purpose of the reviewing organization and/or website, the inclusion and exclusion criteria it uses, and level of detail provided. Depending on the purpose of the organization, different programs will be reviewed. For example, some organizational websites allow for researchers or organizations to submit programs for review, while others have no formal procedure and programs may be reviewed for advocacy or educational purposes only. In addition, some websites are more general and include a broad array of programs, while others are focused on specific types of programs – early childhood, violence prevention, among others. Another important consideration is the level of information reviewed by the website. Some are very comprehensive and cover many of the elements critical to decision-making (i.e. research and details regarding level of effectiveness, cost of implementation, sustainability, technical assistance, and training requirements, among others), while others offer minimal information (See Terzian, Moore, Williams-Taylor & Nguyen, 2009 for more detail).

Also important to making an informed decision is whether a program can be replicated. It is important to know who the program has been most successful with and whether the research shows positive outcomes with multiple groups of individuals. In some instances, a program may have been researched on a very limited population or results may show that it only works well for specific groups (e.g. males or Hispanic children only). For example, while there is a teen pregnancy prevention program used repeatedly with all teens across the country (Draw the Line/Respect the Line), a closer look at the current research reveals that the program is ineffective for females. Another pregnancy prevention program, Safer Choices, has had a greater impact on Hispanic students than on other racial/ethnic groups regarding delay in sexual activity. These are important considerations when choosing a program for your target population – especially given limited funding.

The last important consideration when choosing a program to replicate is whether a program can be disseminated. For example, in 2004, a federal workgroup was created consisting of the U.S. Department of Justice, the U.S. Department of Education, and the U.S. Department of Health and Human Services. The Hierarchical Classification Framework for Program Effectiveness (2004) was drafted. The workgroup created a rating system for programs (i.e. Effective, Effective with reservation, Promising, Inconclusive

---

Evidence-based Programs are more than a list of programs. The decision to choose a specific program model is only the first step on the journey.

Planning for ensuring that programs are implemented with fidelity, as well as for long-term sustainability is vital.

---

repeatedly with all teens across the country (Draw the Line/Respect the Line), a closer look at the current research reveals that the program is ineffective for females. Another pregnancy prevention program, Safer Choices, has had a greater impact on Hispanic students than on other racial/ethnic groups regarding delay in sexual activity. These are important considerations when choosing a program for your target population – especially given limited funding.

The last important consideration when choosing a program to replicate is whether a program can be disseminated. For example, in 2004, a federal workgroup was created consisting of the U.S. Department of Justice, the U.S. Department of Education, and the U.S. Department of Health and Human Services. The Hierarchical Classification Framework for Program Effectiveness (2004) was drafted. The workgroup created a rating system for programs (i.e. Effective, Effective with reservation, Promising, Inconclusive
evidence, Ineffective, and Insufficient evidence), as well as a rating system for readiness to disseminate. The dissemination criteria included training and support materials, technical assistance, information materials, and quality control. These are all important components of dissemination and the ability of an organization to implement a program with fidelity. These should be taken into consideration before making a decision about program implementation. In summary, there are many issues to consider when choosing programs and having a list of programs to choose from is only the start.

There are many issues to consider when choosing programs and having a list of programs to choose from is only the start.

What’s Needed?: Implementation with Fidelity & Long-term Sustainability

In deciding what program to move forward with, it is also important to consider what it will take to implement the program with fidelity and sustain it over time. First, it is important to understand why fidelity is important. Fidelity is defined as the degree to which program implementers provide services or a program as designed by the developer. It is usually measured by adherence to the program dosage, quality of delivery, and participant’s acceptance of the program (Rohrbach, Grana, Sussman & Valente, 2006, p. 308).

Too often, an agency will choose a program and then decide that it is too expensive to implement as designed or that it has to be modified to fit the population in their community. If a program is modified from its original design, the outcomes cannot be guaranteed. Any changes must be approved by the program developer. If any major changes are implemented, even if approved by the developer, the program should be re-evaluated.

Fidelity – The extent to which, during the delivery of the intervention, a person adheres to the program model and core components as originally designed by the program developer.

Each program will have a set of core components, defined as the content of a program that must be in place in order for the program to work - the essentials. For instance, think of a program as a car. If the manufacturer chose to leave out the engine, it would
not work. The engine is an essential or core component. However, they could choose to leave out the floor mats – not a foundational piece to the functioning of the car. Another common problematic practice is taking select components of different evidence-based programs and piecing them together to create a new program. This does not mean that you are implementing an evidence-based program and, without a rigorous evaluation, you cannot determine which, if any, outcomes will be achieved. Susan Philliber (2009), a well-known researcher in the social sciences, describes it as “taking random ingredients from six different cake recipes, mixing them together and expecting to get a cake.”

Now that the importance of fidelity has been discussed, other key components of successful implementation will be reviewed. The first is readiness for change and managing the change process. Readiness for change is defined as “a developmental point at which a person, organization, or system has the capacity and willingness to engage in a particular activity” (Fixsen, Blase, Horner & Sugai, 2009). There is general readiness for change, but there is also readiness in terms of ability to implement and sustain the program. Rather, there needs to be an assessment of the agency’s willingness to implement (readiness for change), as well as ability to implement (agency readiness).

It is important to listen to challenges and concerns of potential implementers and address them as they come up. In order to ensure fidelity and success, individuals must understand why the change is occurring, must feel part of the movement – not that it is happening to them, and must be fully committed.

Change for many people is difficult, so building one’s readiness for change is imperative. There will be some resistance, as many in a community will feel that the services they provide work already, so why change to evidence-based programs. Service providers are often passionate about their work and being forced to implement something new may feel like they are being told that they were doing something wrong (See Chapter 4 for more detail). For this reason, the change process begins with general education on what it means to be evidence-based and building the understanding that in many cases, we just do not know whether the programs being implemented work because they have not been rigorously evaluated. It is important to listen to challenges and concerns of potential implementers and address them as they come up. In order to ensure fidelity and success, individuals must understand why the change is occurring, must feel part of the movement – not that it is happening to them, and must be fully committed. The inclusion of providers and the community and recognizing everyone’s expertise and experience are key to accomplishing the change process (See Chapter 4 for more details). Lastly, it is important to remember that resistance can be constructive. According to Brown (2010), once you work through the resistance and address concerns, many resistors become some of the best champions for the work.
Agency readiness is another important component to successful implementation and sustainability. There is the question of staff capacity and fiscal soundness of the agency. Another question to consider is the agency’s relationship with other stakeholders and partners in the community. Is this relationship strong or are there problems that need to be addressed.

Professional development and training on the new models is also critical to implementation and sustainability. Providers must be trained and fully understand what it means to implement with fidelity and how changes can and often do impact outcomes. Lastly, funding is critical. A person must not only consider how much it will cost to implement the program (e.g. training, materials, site resources, salaries, etc.), but how much it will cost to sustain it over many years. If this is not considered carefully, the people being served will ultimately suffer because the program will end abruptly. This is so important that some evidence-based programs need guaranteed sustainable multi-year funding before the developers allow the program to be implemented (e.g. Nurse-Family Partnership).

To determine whether you are ready to implement a selected evidence-based program, the following are some of the general tasks/steps to engage in beforehand (Nurse-Family Partnership, Incredible Years, 2003):

- **Step 1**: Learn about the program model and implementation requirements;
- **Step 2**: Assess whether the program is needed;
- **Step 3**: Determine whether the program aligns with the organization's mission/vision, philosophy, and goals;
- **Step 4**: Assess how the program will function (e.g. referrals, recruitment, marketing);
- **Step 5**: Determine funding for start-up and sustainability;
- **Step 6**: Examine each of the agencies in your community with a good reputation for successfully working with the target population. Determine which one might be the best able to operate the program, and which one’s might best play supporting roles;
- **Step 7**: Assess your capacity to recruit and support staff knowing the high level of expertise and qualifications needed for quality implementation (e.g. training, technical assistance, supervision);
- **Step 8**: Plan for ongoing monitoring and fidelity checks;
- **Step 9**: Create a plan for program evaluation;
- **Step 10**: If needed, determine what is needed in order to participate as a member of a national network of program implementing agencies and coordinating with the national office; and
- **Step 11**: Complete a formal application – some evidence-based programs require an application to become a formal implementation site.
References

Brown, A. (2007). Personal communications. Quoted from a training delivered to provider staff hosted by the Children’s Services Council of Palm Beach County.


Chapter 4: Barriers to Implementing Evidence-Based Programs

In researching, as well as implementing evidence-based programs, it is apparent that there are many concerns among funders and providers about whether moving in this direction is possible. Some of the most common concerns are that there are not enough evidence-based programs to meet all of a community’s needs, program fidelity, cost, staffing needs/requirements, and organizational barriers.

Lack of Programs

Not enough programs currently meet evidence-based criteria, namely having been rigorously tested. Some of the reasons for this are that we are still in the infancy stage of moving toward evidence-based programming in the social sciences, as well as financial and time barriers.

It is important to remember that it was only in the early 1990s that the research and evaluation field really began to flourish for the social sciences. Only after several programs produced significant results did the government begin to provide concentrated funds for outcome evaluation studies. For this reason, the prevention field is still learning about what does and does not work.

In addition, after programs have been researched they can be submitted to one of the organizations rating evidence-based programs (Appendix D). However, not everyone is aware of this process, so some programs do not come to the attention of funders, providers, and national rating groups. Many of the programs known today have motivated marketing staff that dedicate their time to increasing awareness about the program. If resources are not available, some effective programs may stay under the radar.

Financial resources are also needed to complete outcome evaluation studies. Experts in the field agree that to become evidence-based, a program must be subjected to rigorous evaluations – two are needed to test ability to be replicated. Furthermore, to demonstrate sustainability, participants need to be followed for at least one year. This, too, increases the financial costs.

So, while not enough evidence-based programs exist at this time to meet all of the needs within a community, funders and government should begin thinking about how to move unproven programs along a continuum of effectiveness in order to increase

---

2 Co-authored with Jennifer Estrada, LMHC, contract manager and CSC evidence-based committee member.
accountability and ultimately introduce more evidence-based programs to the field. The more programs we can determine are evidence-based, the more impact we can guarantee for the children and families served.

The following is a general strategy for upgrading the quality of programs in one’s community (Elliott, 2007):
- Whenever possible, implement top-tiered evidence-based (proven) programs;
- Include an outcome evaluation component in funding for non-evidence-based (unproven) programs;
- Discontinue programs evaluated and found ineffective; and
- Sustain and build capacity for evidence-based programs with mainstream funding

The following is an important addition to this general strategy:
- Determine a strategy for moving non-evidence-based (unproven) programs further along a continuum of effectiveness (See Chapter 5).

**Uniqueness of One’s Community and Maintaining Fidelity**

Adaptation or change of the program design is usually done because it is perceived that it will make the program more acceptable to the local environment. However, “there is strong evidence that some programs only work when implemented with a high degree of fidelity, and other research suggests that closer adherence to core components results in stronger participant outcomes” (Mihalic, Fagan & Argamaso, 2008). For example, Rohrback et al. (2006) has reported that many times school programs are adapted and components are eliminated to make them more feasible. The LifeSkills Training Program is often adapted by the teachers implementing it by adding a scare tactic component. This approach has been shown to have no effect and may in fact be harmful. Another example was a pilot implementation project of the Olweus Bullying Prevention Program conducted in 15 schools. This was done before the program was to be implemented statewide in Arizona to over 200 schools. In the sample of 15 schools, researchers found that five schools implemented the program with very good fidelity to the program model. All of them achieved the same or better outcomes than the original program research produced. Researchers also found five schools whose implementation had significant problems (e.g. site coordinator left, school did not embrace the program, lack of commitment, and implementation was poor). These schools, while producing some positive outcomes, produced mainly weak results compared with what is possible when
the program is implemented with fidelity. Finally, the five remaining schools experienced some implementation problems, but succeeded in implementing parts of the program successfully. These sites produced mediocre outcomes - not as good as the schools that implemented with fidelity, but better outcomes than the schools that were judged to have failed to implement well (Chadwick, 2009).

While these examples illustrate the importance of implementing with fidelity, providers continue to report the need to adapt programs in order to implement them within their community. However, Elliott and Mihalic (2004) report that little research supports the need to adapt programs, but they acknowledge that language and cultural adaptations may be the exceptions (p. 51). For example, language and the translation of evidence-based programs may impact program fidelity; thus, it is important to discuss these changes with the program developer. Indigenous beliefs and culturally traditional practices also play a factor in the implementation of a program. Family context of a population should be considered. For example, is the family system patriarchal or matriarchal, is open communication acceptable, how important is consultation of elders or protecting family reputation (Kader, 2007).

Thus, while language and culture must be considered, Elliott and Mihalic (2004) report that every program does not necessarily need separate treatments for different

“Proponents of adaptation have a tendency to substitute program sustainability for program effectiveness. Local adaptation may well increase the likelihood of sustaining a program, but if it renders the program ineffective, this is not a desirable outcome. Both fidelity and sustainability are necessary to an effective prevention effort” (Mihalic, Fagan, & Argamaso, 2008, p.2).

sexes or racial/ethnic groups, especially when the program is geared toward children and adolescent populations. Evidence-based programs are geared toward promoting protective factors and decreasing the impact of risk factors. In general, children that have the same protective and risk factors should respond similarly to these types of interventions. Remember that these programs have already been researched with various populations across the country. Thus, while one's approach in working with the family may be different (with the developer's approval), the core components of the program should not, in most cases, need to be altered.

In summary, it is understandable and appropriate to question the generalizability of evidence-based programs, but it should not be assumed that the program will not work due to the uniqueness of one's population given the lack of evidence to support this belief. Continuous evaluation and exploration of outcomes of these programs will provide feedback as to whether they are working within a unique population.
Fidelity and Process Evaluations

There are many factors that impact one’s ability to implement with fidelity, including, but not limited to, staff training, implementer support and comprehensive monitoring and oversight after implementation (Mihalic et al., 2008). Given that many factors impact whether implementation is successful (i.e. with fidelity), as well as its importance and influence on outcomes, process evaluations become essential. These types of evaluations assess whether a program is implemented as planned or as intended (Maxfield & Babbie, 2005, p. 438). The premise is that as long as the evidence-based program is implemented as designed, you should get the desired outcomes. For this reason, the process evaluation replaces the need for a rigorous outcome evaluation when using evidence-based programs. This means that while you will still need to monitor outcomes using a non-experimental evaluation (e.g. pre-post test), an RCT or quasi-experimental evaluation is not needed.

According to Carroll and colleagues (2007), a process evaluation to test implementation fidelity should include adherence to the intervention, exposure, quality of delivery, participant responsiveness and program differentiation.

- **Adherence** refers to whether the service or intervention is being delivered as it was designed or written (i.e. with all core components being delivered to the appropriate population, using the right protocols, techniques, and materials, that staff are trained appropriately, and that it is delivered in the locations or contexts as prescribed).
- **Exposure**, also referred to as dosage, includes any of the following: the number of sessions implemented, length of each session, and/or the frequency with which program techniques were implemented.
- **Quality of program delivery** includes the manner in which a program is delivered (e.g. skill in using the techniques or methods prescribed by the program, enthusiasm, preparedness, and attitude).
- **Participant responsiveness** is the extent to which participants are engaged by and involved in the activities and content of the program; and
- **Program differentiation** identifies the core or essential components of a program. A component analysis is thus recommended to determine fidelity to these core parts of the program.

These are the five key factors that should be addressed in a process evaluation to assess whether the program is being delivered in a way that will produce the expected outcomes.

As you can see from these examples, fidelity is a complex concept that requires commitment and support from those implementing the program (from leadership to
front line implementers), as well as from funders and/or grantors. Given the need for intensive monitoring, sometimes this is considered a barrier to implementing evidence-based programs.

Cost

The initial costs of implementing an evidence-based program can be high because often programs need to be purchased from their developers. However, this is not always the case (e.g. LifeSkills Training, Teen Outreach program). Additional expenses include formal staff training and the monitoring and process evaluation of the program.

Thus, while evidence-based programs may initially be more expensive than using untested programs developed by agency staff, one of most important considerations is the concept of “pay now or pay later.” It may cost more up front for these types of programs, but you know that you will be getting real undisputable results. Furthermore, many of these programs have gone through in-depth cost-benefit analyses and have shown a return on the investment years later. So, while these programs may cost more money upfront, the return is much higher later on.

Staffing

Several barriers can arise with regard to staffing. These include resistance to change, lack of understanding and misperceptions about evidence-based programs, complexity of the intervention, staffing qualifications and training, and misunderstandings about data.

- **Resistance to change, lack of understanding and misperceptions** –

  “Each time we ask someone to change, we ask him or her to take a journey into incompetence” Tynette Hills (Quoted in Talan & Bella, 2007)

  When implementing anything new, sometimes there is resistance. Some individuals may see change as meaning that what they were doing was no good or they were doing something wrong. They often will not understand why they are being asked to change, which may lead to resentment or total disinterest. For this reason, it is important that individuals be educated about evidence-based programming, why a shift or movement is occurring, and how the change will benefit them as well as the people they wish to serve. Working with staff to develop a shared vision in regard to the program’s goals and how they fit into the organization’s vision and mission can
provide staff with a sense of direction and shared purpose. Without a shared vision, staff may implement programs differently, causing confusion and failed outcomes (Small et al., 2005, p. 34-37).

Clarifying misperceptions about evidence-based programs can assist in changing attitudes, as well as impacting fidelity and overall success of implementation. The idea that the new practice or program will take an inordinate amount of time or that it is a “cookbook” or one-size-fits-all approach should be dispelled through education and support to the practitioner. Practitioners should be made aware that evidence-based programs are only appropriate for use when they address individual client needs and their own clinical expertise (Melnyk, 2002). By definition, evidence-based programs require more than one study. They are programs that work with various populations and have shown effectiveness at various sites across the nation. Through research, one can see for whom the program works best, as well as if there are populations for whom it is less effective. Practitioners must buy into the idea that the new practice will be better than what they have been delivering.

Training alone will not change the way practitioners deliver their service. Adequate time must be allowed so that providers can also overcome the fear of practicing differently than their peers (Melnyk, 2002). It is vital that leadership pay attention to this fear and resistance because it can impact fidelity and quality of implementation. Practitioners may regress to what they are more comfortable doing thus either deliberately or unintentionally sabotaging fidelity. Also, individual resistance to change could be masking individual readiness or influences from other organizational barriers to implementation (Hyde et al., 2003).

- **Complexity of the intervention** – If the service is complex or requires more time from the practitioner they are not likely to feel positive or eager about implementing it. This can be a deterrent for overworked employees. Complexity can also impact fidelity. “Complex interventions have greater scope for variation in their delivery, and so are more vulnerable to one or more components not being implemented as they should” (Carroll et al., 2007, p.5).

- **Appropriately qualified and trained staff** – Employing qualified staff is a critical piece to evidence-based programming. Effective program components alone are not enough to ensure program success. Staff who are motivated, have a sense of ownership and possess the necessary skills, experience and credentials are an integral part of a successful program (Small et al., 2005, p. 34-37).

For these reasons, supporting staff with training, coaching, and technical assistance is critical. First, staff must be educated and trained about evidence-based programming in general before being asked to adopt a specific program. Providing
formal and informal trainings on evidence-based programming, access to articles and publications on the concept, as well as programs, and allowing staff to visit currently implemented programs can also help engage them in the process and movement towards evidence-based programming (Hyde et al., 2003).

When moving towards training on a specific program model, ongoing support must be provided. This includes allowing time to learn and discuss the new program, time to participate in planning, time to practice the new skills and for implementing the new program, as well as providing reinforcement of the learned skills and feedback opportunities. Support through training and ongoing technical assistance will ensure that staff members are more likely to implement the program with fidelity (Hyde et al., 2003). Programs with well-trained staff are more likely to have positive outcomes. Research indicates that effective training consists of presenting information in a consistent and engaging manner, providing demonstrations of the important elements of the program, and affording staff opportunities to practice key skills in the training setting (e.g. role playing). It also entails ongoing coaching and mentoring throughout the implementation of the program to ensure that learned material and core components of fidelity are integrated and applied into one’s everyday practice as taught in the workshops/training. This is often referred to as “transfer of learning.” Ongoing technical assistance can also address unforeseen problems as they arise (Small et al., 2005, p.34-37).

- **Misunderstandings about data** – Data collection should not only be used for measuring practitioner performance, but also adherence to the program model, assessing client outcomes and for overall quality improvement (Metz, Blasé & Bowie, 2007). All parties involved in the implementation of the program (e.g. practitioners and funders) should know if the program is working. The concept of data collection has gotten a bad reputation. This may be called the “gotcha effect.” Often practitioners feel that data is collected as a way of telling them what they are doing wrong. Instead, data should be used as a way of giving regular ongoing feedback, both positive and negative. It can give practitioners a gauge of how they are doing as well as provide feedback on the impact their efforts are having on the community they serve. Furthermore, practitioners should be trained on how to use data themselves for quality improvement purposes.
Organizational Readiness and Other Barriers

While many organizations may want to begin implementing evidence-based programs, some of the negative factors often cited as barriers include problems with readiness, resistance to change, bureaucratic processes, and financial issues.

- **Organizational Readiness** – An assessment of an organization's readiness is important to determining whether the commitment and resources needed are available before program implementation. Assessments can include identifying potential barriers or problems to implementation and how to overcome these obstacles. Having these discussions early in the process can help create buy-in among staff which translates to less resistance and fear (Small et al., 2005, p. 34-37). Additional areas that should be considered are prioritizing the outcomes to be achieved, obtaining the support of critical stakeholders, researching and identifying which program(s) are a “good fit”, identifying what technical assistance is needed, program consultation and decision-making about how to ensure that the program or programs selected will be implemented with fidelity (Metz, 2007).

  The Readiness for Organizational Learning and Evaluation Scale (ROLE) is a tool to measure readiness. Respondents can rate their organizational readiness on six dimensions: culture, leadership of the organization, systems and structures, communication, working in teams, and the capacity to conduct program evaluations and use data in decision making (Preskill & Torres, 2000). Another tool, the General Organizational Index (GOI) measures the operating characteristics related to capacity that are necessary for implementing and sustaining evidence-based programming. The tool measures ten broad areas including, program philosophy, training, supervision and program monitoring (North Carolina Evidence Based Practices Center, 2002). According to the North Carolina Evidence-Based Practices Center (2002), the rationale for using such a tool is that programs that score well in these areas generally do an excellent job at implementing evidence-based programs.

- **Organizational Resistance to Change** - The implementation of a new program cannot flourish in a change-averse culture in which clients, staff or the organization as a whole resists change. Resistance may be caused by philosophical differences, leadership style of top level managers and boards, length of time the leadership has been in existence, the context in which the organization operates, and the financial stability and history of the organization. The ability of the supervisors to implement change and the level of managerial skills and knowledge of how to implement the practice themselves also has an effect. In addition, the question should be raised - do senior leadership, clinical leadership
and front-line supervisors endorse and believe in the proposed change? All three levels of support are necessary (Hyde et al., 2003).

- **Bureaucratic Process** - The way in which the program or an organization is set up may impede new practices. Bureaucratic processes are often used as excuses for not implementing a new practice or program. The rules of doing business for that organization may pose problems for the implementation of a new practice or program. Potential ways to combat this include identifying where this is true for the organization and keeping an open mind about changing it. Too often we hear that “It’s been done like that for years” and “It is too difficult to change.” While this may be true in some cases, it probably can be combated, it will just take time, patience, and at times, thinking outside the box.

- **Financial Issues** - A major barrier is the way in which services are funded and practitioners are paid. For many agencies that receive support from multiple funders, there may be lack of support for moving in this direction due to a lack of understanding regarding the expense. Financial mechanisms and payment rates or limitations can also drive or hinder changes in service delivery. Reimbursement rates for practitioners and whether the old method of employing practitioners was more cost effective than the new method can be a hindrance. Such payment limitations can be imposed at the state or local level (e.g. Medicaid Title IV-E, TANF or local funders). Lastly, the program's compensation rate, productivity requirements and financial support can also cause problems (Hyde, et al., 2003).

**Political Forces**

There are often advocates that have reasons to support programs that have little or no evidence of effectiveness. Programs exist that have incredible support of legislators and other policymakers, yet have never been asked to show outcomes. In some cases, even if the program has been researched and has shown no effect, there is continued support. Many of these programs have been implemented nationally for decades. For example, in one state, researchers were asked to create a portfolio of programs that work and recommend eliminating those with no research to support them. In doing so, the researchers began making recommendations and began experiencing these types of responses - “That program's been around for years, we can't just eliminate it” and “My grandson attended that program and it worked for him” (Brown, 2007). Political forces impact whether change occurs, but the important point is to continue educating and advocating for evidence-based programs. Change takes time and persistence.
References

Brown, A. (2007). Personal communications. Quoted from a training delivered to provider staff hosted by the Children's Services Council of Palm Beach County.


Elliott, D. (2007) Personal communications on January 11, 2007. Select strategies are quoted from a presentation delivered at the Senior Executive Policy Institute, a meeting hosted by the Children's Services Council of Palm Beach County.


The Journey to Evidence-Based Programming
Chapter 5: Case Study:  
The Children’s Services Council of Palm Beach County

Children's Services Council (CSC) of Palm Beach County was created in 1986 through a countywide referendum. It is an independent special district of local government, which levies property tax dollars to provide funding for programs and services serving children and families. In the last few years the pressure for no-nonsense spending has instigated the movement toward (1) ensuring that funds are provided to primary prevention and early intervention programs with evidence that they work, (2) funding locally developed programs that have the potential to become evidence-based, and (3) funding programs that demonstrate a return on investment.

A Brief History

The CSC’s founding Council members believed that the success of the CSC would be based on the following guiding principles:

- Services and funding would be dedicated to primary prevention and early intervention;
- Programs and services must be innovative and based on high quality research; and
- Programs would be held fiscally and programmatically accountable for achieving child well-being outcomes and impacts.

Each of these founding principles has driven how CSC does business over time. In changing times, CSC has had to make adjustments in order to remain true to these founding principles. For example, as new research became available, like the newest brain research and its impact on child development, CSC took this information and incorporated it into its decision-making process for funding service delivery systems and systems of care. Thus, CSC has had to be flexible and amenable to change because of its focus on being data-driven and using research and data for continuous learning and improvement and to inform decision-making.
1st Guiding Principle: Services and funding would be dedicated to primary prevention and early intervention.

Research indicates that investing in primary prevention and early intervention yields the best outcomes for those at risk, as well as the best return on the dollar. The earlier you intercede with a child or family when they are experiencing problems by either promoting protective factors or mitigating risk factors, the greater the likelihood that outcomes are achieved. Think about a patient that has a heart condition or cancer, the earlier the disease is detected and treated, the higher the rate of recovery – it is the same for those experiencing social problems. For example, for a pregnant mother, the earlier she enrolls in prenatal care (i.e. during the first trimester), the greater her chance of delivering a healthy baby. It is no different for a family under incredible levels of stress and at risk of abuse or neglect. The earlier they receive the support needed, whether that is basic needs, parenting support, or substance or mental health treatment, the greater the likelihood of mitigating the risk factors and decreasing child maltreatment.

In addition, an incredible amount of research has emerged in the area of brain development over the past decade. One of the critical movements in the field has been the increased attention to the fact that brain development does not begin at birth, but in-utero and that there is a link between birth outcomes and later risk-taking behavior and poor long-term outcomes. Research has also demonstrated that a child's brain develops more rapidly during the first five years of life than at any other developmental period. Early childhood experiences greatly influence and form the foundation for later success in life. It is during the first five years that the foundation for physical, emotional, social, behavioral, and cognitive development occurs. Furthermore, research indicates that investing in prevention and early intervention yields the best return on the dollar. For example, if funding were spent on programs that prevented teen pregnancy, poor birth outcomes, child abuse and neglect, and juvenile delinquency, the impact and money saved on later programs working to mend the damage to our most vulnerable would be enormous. Using local data and return on investment research, the following data supports the need for CSC’s investment in prevention and early intervention programs and services:
• **Teen Pregnancy** – According to the Florida Department of Health (2009), in Florida, there were 8,495 births to mothers 10-17 years of age in 2007. In Palm Beach County, there were 500 births to these younger mothers. The average annual public sector cost of a child born to a mother 17 years and younger has been estimated at $4,080 (Hoffman, 2006).³

  In Florida that equates to approximately $35 million per year in public costs.
  In Palm Beach County that equates to approximately $2 million per year in public costs.

• **Prematurity** – In Florida, there were 33,818 premature births (born at less than 37 weeks) in 2007. In Palm Beach County, there were 2,304 babies born pre-term (Florida Department of Health, 2009). Hospital charges alone for these infants average $75,000 per child (National Governors Association Center for Best Practices, 2004).

  In Florida that equates to approximately $2.5 billion per year in hospital charges alone.
  In Palm Beach County that equates to approximately $173 million per year in hospital charges alone.

• **Low Birthweight** – According to the Florida Department of Health (2009), in Florida, 20,767 babies were born low birthweight (11.5% of all births), including 3,886 that were very low birthweight. In Palm Beach County, 1,480 babies were born low birthweight in 2007 (10.6% of all births). Of these babies, 259 were very low birthweight. The hospitalization costs for low birthweight babies have been estimated at $20,600 for low birthweight and $52,300 for very low birthweight deliveries (Russel et al., 2007).

  In Florida that equates to approximately $551 million per year in hospitalization charges alone.
  In Palm Beach County that equates to approximately $39 million per year in hospitalization charges alone.

³ Average costs are based on 2004 dollars
Child Abuse & Neglect – In 2007-08, the most recent year for which data is available, 53,484 children in Florida were victims of abuse or neglect (verified or some indication) (U.S. Department of Health and Human Services, 2009). In Palm Beach County, 4,293 children were abused or neglected in 2008-09 (1,499 verified and 2,794 some indication) (Department of Children and Families, 2009). While arriving at an average cost is difficult because types of abuse and services needed vary greatly, the best estimate indicates that the average cost is $57,462\(^4\) per child in direct and indirect costs (Conrad, 2006).

\[\text{In Florida that equates to approximately $3 billion per year in costs.}\]
\[\text{In Palm Beach County that equates to approximately $247 million per year in costs.}\]

Juvenile Delinquency – According to the Florida Department of Juvenile Justice, 74,692 children ages 14 to 17 were referred for delinquency in 2007-08 across the state; in Palm Beach County, the number was 5,220. The estimated range of external costs for a typical juvenile delinquent aged 14 to 17 was between $83,000 and $335,000 depending on the severity of the crime (Cohen, 1998).\(^5\)

\[\text{In Florida that equates to costs of approximately $6.2 billion to $25 billion per year.}\]
\[\text{In Palm Beach County that equates to costs of approximately $433 million to $1.7 billion per year.}\]

From this snapshot of statewide and Palm Beach County data, you can see the need to invest in programs that have proven results in these target areas. Governmental agencies such as CSC are responsible for making sure that money is spent cautiously and thoughtfully. Investing in evidence-based programming affords us the opportunity to do just that.

\(^4\) Average cost is based on 1999 dollars.

\(^5\) Average cost is based on 1997 dollars.
2nd Guiding Principle: Programs and services must be innovative and based on high quality research

Since its formation, the Council has required that programs and services be based on quality, research and have strong theoretical frameworks. While CSC does not want to stymie innovation, evidence that supports program models is required. One way that CSC ensures that the programs it funds do this is through the use of logic models. Each program is required to have a logic model that clearly outlines the activities and outcomes of the program.

National evaluators came to Palm Beach County in the early 1990s and trained CSC staff and program personnel in logic model development. The purpose is so that each funded program would have a clearly articulated model that has a sound theoretical foundation and research supporting the target population, activities, and proposed outcomes. Since then, CSC has taken on this role and now provides ongoing training to program staff and community partners on logic model development. In addition, staff also trains them on creating a theory of change for their programs – another strategy for ensuring a focused vision on outcomes.

3rd Guiding Principle: Programs would be held fiscally and programmatically accountable for achieving child well-being outcomes and impacts

CSC created a comprehensive system to measure accountability. The fiscal and research and evaluation departments work together to ensure that programs are held accountable. The fiscal department ensures that funding is spent for the outcomes the program staff proposes to achieve. The research and evaluation department then analyzes and reports on outcomes and impact levels attained. While many programs, both locally and nationally note what they are trying to achieve, their actual achievements may differ considerably.

Recently, CSC created a separate auditing department to ensure an additional level of accountability. Lastly, contract managers throughout the organization monitor the contracts to ensure that the agencies are implementing programs as per their agreements. CSC uses taxpayers’ money to fund many child and family programs, so it is essential that they use it on the best programs – those achieving outcomes and those that provide a return on investment.

One of the major strategies to make sure that CSC was funding the implementation of the best possible programs to ensure accountability was to move towards implementing
evidence-based programs. This endeavor required an entire system adjustment with various processes working independently and dependently to create the needed infrastructure to support a system of evidence-based programming.

The Infrastructure

Implementing evidence-based programs within a system is a process that requires many support systems working in collaboration. These include having a strategic policy agenda, a comprehensive research and evaluation system, an accountability infrastructure, a comprehensive professional development system, and policies related to agency relations – a support specifically needed for funders.

While this infrastructure is recommended for any organization wanting to move towards evidence-based programming, the following information will provide a snapshot of how CSC used the components to build the framework needed to ensure success.

The Infrastructure

Strategic Policy Agenda – A critical piece of the process was the support from CSC’s board and executive leadership to move in the direction of implementing evidence-based programs. This support was needed in part because a great deal of time and resources were needed during the process of choosing the right program, implementing the program and sustaining it over time.

According to Elliott and Mihalic (2004) when it comes to replicating evidence-based programs, most failures are the result of inadequate site preparation and/or capacity. Staff is not always ready for the complexity of implementing such a program and it can take upwards of six to nine months to get a site ready for implementation (Elliott & Mihalic, 2004). Moreover, when an agency decides to implement an evidence-based program, a need for some organizational change is almost always needed (Fixsen et al., 2005, p. 64). Thus, there is an increased need for commitment through a sometimes difficult period of transition. This decision-making and support is part of the overall strategic policy agenda.

Research and Evaluation System – An organization’s targeted outcomes should be the driving force behind choosing the types of evidence-based programs to implement. CSC created both a system logic model and worked with Lisbeth Schorr and others to devise a pathways document (Appendix B) to help focus efforts within the organization, as well as communicate with external organizations. The measurable conditions and interventions that lead to the primary long-term outcomes/goals were the inspiration behind what types of programs were researched. In other words, the logic model assisted the evidence-based work by helping to maintain a focused approach in researching what evidence-based programs have been implemented both nationally and internationally that achieve the outcomes CSC wanted to achieve. In addition, CSC is now engaged in creating a pathways tracking portfolio. A committee is exploring the best measures and assessment tools for each of the outcomes and measurable conditions. The portfolio identifies 19 measurable conditions as core indicators toward achieving CSC’s goals – healthy births, fewer children being abused and neglected in the first five years of life, and children being eager and ready to learn. This work enables the organization to monitor progress, review trends and patterns and visually display these trends over time.

Lastly, in addition to our system logic model, the research and evaluation department works with each individual program staff to create a program logic model and theory of change. This assists agencies in understanding why they need to implement their programs as designed and helps them see where they fit within the system.
Accountability Infrastructure⁶ – Accountability is the underlying foundation of evidence-based programming. To be able to prove effectiveness, CSC needed a sound data collection infrastructure in order to evaluate progress in meeting goals and outcomes. CSC also wanted to ensure that the funded programs were meeting their contractual requirements and providing the best services. Therefore, CSC needed a system that could track client-level changes and the sustainability of outcomes for the clients within various programs. In addition, CSC needed a comprehensive data system in order ensure programs were being implemented with fidelity.

For some families, the amount of support needed to achieve positive outcomes may be minimal, while others may require multiple services delivered over the course of several years. It is incumbent upon CSC to ensure resources committed are used efficiently and effectively, from the short-term perspective of the outcome associated with an individual family, as well as relative to the longer-term perspective of what happens over time within a cohort of children.

This second perspective requires that CSC conduct evaluations analyzing the synergistic impact of multiple services in achieving our outcomes and goals. CSC is committed to investing in evidence-based programs, and while considerable advances have been made in the number of programs that have been identified as evidence-based, gaps remain that must be filled by locally designed programs and services. We have much to learn about the impact of locally designed programs and services and what the return on investment is for high-need families if multiple services are needed. CSC will be analyzing data longitudinally to understand the impacts of our investments. CSC will examine whether children who participated in our funded programs are on grade level by third grade, graduate from high school, and do not become parents prior to their 18th birthday. To conduct this level of impact analysis, it will be necessary to match individual level data from client information systems with other administrative data sets. This will allow CSC to explore impacts longitudinally, ensuring that both the programs we fund and CSC itself remain accountable to the children and families served by programs and services funded by CSC.

Professional Development System⁷ – In order for programs to become more effective or remain effective in providing services to children and families, they must have professional development support. Even if they are implementing a program that is not defined as evidence-based, they will need professional development support to help move the program along the continuum of effectiveness.

⁶ Quoted in part from information provided by Tanya Palmer, lead person working on the accountability infrastructure.

⁷ Quoted in part from information provided by Carol Rodriquez, lead person working on the professional development system.
At CSC, we have created a system with specific requirements and trainings to help with the movement towards evidence-based programming. For example, CSC believes that there is a need for specific training to assist providers in collecting quality data. No matter what level or type of training is provided, one important piece to consider is that professional development is ongoing. There is a need for quality assurance on a continual basis to ensure that the evidence-based models are implemented with fidelity. Ongoing learning happens in organizations where structures and practices are in place to ensure that what is learned in training is utilized in practice. Examples of such structures are executive level support, accreditations and certifications, reflective practice, peer mentoring, quality and consistent one on one supervision (reflective and developmental in nature), and observation with feedback to staff interacting with clients on how well they are implementing the model as it was designed. CSC has found that there are other training areas imperative to evidence-based programming. These include trainings on: (1) developing a logic model and theory of change, (2) evidence-based programs, (3) the importance of fidelity; (4) quality data collection techniques; and (5) evaluation and measuring success.

- **Logic model and theory of change** - This assists agencies in addressing what outcomes they wish to achieve, what activities should be initiated, and what supports are necessary to achieve those outcomes. Creating these program documents helps with communication because everyone is keeping the goal in mind. Everyone has a clear picture of where they are going and why they are implementing the program in a certain way in order to achieve the outcomes. Lastly, logic models help everyone see where they fit in the big picture – how what they accomplish coupled with other program achievements will ultimately meet the overall goals of CSC.

- **Evidence-based programs** - Staff needs an understanding of what an evidence-based program is and why they may be asked to implement a program other than what they have been doing. Sharing with staff the research that supports evidence-based programs helps mitigate the fear and discomfort around change. Staff understanding, acceptance and willingness to change is one of the most important pieces to making sure that the program will be implemented with fidelity.

- **Fidelity** - Replication requires fidelity (i.e. that all of the core elements of a program are implemented as designed by the program developer). Understanding this and how changing a model can negatively impact results are vital to ensure success.
• **Quality data collection and evaluation** - Staff must also have knowledge about quality data collection and evaluation. Poor data collection impacts program success at multiple levels. Rather, poor data may indicate success when the program is not meeting its outcomes or may indicate failure when the program is achieving results. Furthermore, quality data collection is important for quality improvement and making informed decisions.

Overall, these staff development opportunities, among others (e.g. cultural competence, client engagement, understanding risk and protective factors) will assist programs in meeting goals, being accountable, and providing the best possible programs for the children and families.

**Agency Relations**\(^8\) – One important decision for a funding organization is in determining the type of relationship it should have with the agencies it funds to implement programs and services. The funder must establish what is needed in order for it to be successful in achieving outcomes. This also means determining what the agencies need and want in regard to support, both financially and otherwise. The level of support and what type of relationship model the funder wants to replicate is important to ensuring successful replication and fidelity to the design of evidence-based programs.

CSC has determined that its relationship with the agencies it funds cannot be strictly a contractor/service provider relationship. Rather, this relationship must incorporate (1) support for areas such as professional development to optimize the skills and competencies agencies’ staff need in order to implement the program models; (2) adequate funding that affords staff the resources needed to implement programs with fidelity; (3) support for the technology and data system(s) CSC needs to manage its systems of care and to evaluate the impact of the program and systems; (4) a plan for flexibility (without “losing” key CSC requirements) to deal with changes as evidence-based programs are implemented; (5) a high level of communication between CSC and its agencies; (6) provisions for a more collaborative approach between CSC and the agencies it funds to implement its programs; and (7) requirements for agencies to be high-performing, data driven, committed to continuous learning and improvement, and have a mission aligned with CSC’s core goals and outcomes. Specifically such agencies are required to:

\[^8\] Quoted in part from information provided by Tom Sheehan, lead person on the agency relations committee.
• Deliver the services and interventions needed to achieve the goals/outcomes CSC is seeking;
• Exhibit fidelity in running the program or service contracted for;
• Be data-driven – provide timely, accurate data to CSC and use data for continuous improvement of the agency. ("Atmosphere must be such that agencies are willing to share/discuss data honestly");
• Participate in system activities, where applicable;
• Have their own theory of change and logic model;
• Meet terms of contract (e.g., deliverables produced, timeframe met); and
• Participate in continuous improvement activities.

**Evidence-Based Programming at CSC**

Given the emerging movement toward evidence-based programming, CSC decided to form a committee in 2006 to begin researching programs. The committee was comprised of various individuals from across the agency, including researchers, system managers, contract managers, and strategic planners. The workgroup also worked with experts from outside organizations and universities, such as the Coalition for Evidence-based Policy, Blueprints, and Arizona State University.

This committee was guided by three main principles: (1) Primary prevention and early intervention systems of care and programs promote significant positive outcomes; (2) Evidence-based programs have demonstrated the most effective and long-lasting outcomes; and (3) Accountability is imperative to doing business.

The committee had two main objectives: (1) research nationally and internationally recognized programs, as well as programs that have not been formally rated, but appear to be promising given their research backing, and (2) construct criteria to rate CSC’s currently funded homegrown or locally developed programs to see where they might fall on a continuum of effectiveness. This second part was critical because CSC knew that not enough evidence-based programs existed to meet every need within Palm Beach County. Therefore, the discussion could no longer be “evidence-based or not.” It needed to be expanded so all programs and providers could see where they fit and what needed to be done in order to move their program towards becoming evidence-based.

---

9 Core members as of 2010 included Lisa Williams-Taylor (Lead), Betty Scott, Lance Till, Kim Lu and Jennifer Estrada.
Objective 1: Researching nationally and internationally recognized programs

The first step was using the logic model and pathways documents to help guide the process (See Appendix B). The committee explored more than 500 programs in total. While the primary goal was to determine which programs would be possible funding opportunities for CSC, a second priority emerged. This included exploring which programs CSC would advocate for within the community (i.e. the school district, the Department of Children and Families, the criminal justice and juvenile justice systems, and various other coalitions). This work occurred in phases – (1) Adopt a definition and criteria for evidence-based, (2) Explore the various organizations rating programs as evidence-based, (3) Collect information on programs and eliminate programs, (4) Create a screening system; and (5) Create a rating system and re-evaluate programs.

Phase I and II: Adopt a definition for evidence-based and explore the various organizations rating programs as evidence-based

Phases one and two occurred simultaneously. The committee began the process by examining who in the field had been rating programs and examining the criteria they used in determining “what works” to assist CSC in developing its own definition. The committee found that there was no consistency among raters. Staff analyzed the scope/criteria/definitions used for measuring the effectiveness of evidence-based programs using the information compiled by Sharon Mihalic from Blueprints, Center for the Study and Prevention of Violence. She noted 12 organizations that rate programs. The committee then found 13 additional organizations. Each of these organizations use different criteria, definitions, and labels (See Appendix D), meaning that many issues and concerns had to be considered before moving forward. The committee also found that organizations varied considerably in general terminology usage, such as evidence-based, research-based, best practices, and evidence-informed, among others. Furthermore, many organizations have created systems for identifying and classifying programs as evidence-based, and these too vary considerably. Evidence-based is often used synonymously with research or science-based programming. Other terms commonly used are promising programs, model programs, effective programs, and exemplary programs. Each of these terms has a different meaning depending on the organization using them. There are at least 25 organizations that have created criteria to rate program effectiveness, which
leads to further confusion. A comprehensive understanding of the criteria used by each organization is imperative for those researching program effectiveness – **choosing a program labeled as evidence-based by an organization that has lax criteria, may lead to funding or implementing a program that does not get the best results.**

In conducting a literature review, it became important that the committee decide on the terminology CSC would use as an organization. The committee accomplished this task by researching various organizations and then deciding which definitions and criteria were credible and rigorous. In the end, the committee utilized the information provided by many organizations to help inform decision-making and would recommend using the following:

- Blueprints for Violence Prevention (2007)
- Prevention Research Center for the Promotion of Human Development - Penn State University (2000 Report)
- Promising Practices
- Substance Abuse and Mental Health Services Administration
- What Works Clearinghouse
- Strengthening America's Families (1999)
- Office of the Surgeon General
- Child Welfare League of America

The committee decided to use the term evidence-based and decided that CSC would discuss evidence-based in terms of programs, practices, and curricula. We believed it was important to note that programs may use various curricula and practices in their delivery of services – some of which may have been researched independently and some which have not.

- **Evidence-based program (EBP)** – A program comprised of a set of coordinated services/activities that demonstrate effectiveness based on research. While the criteria for rating as such depend upon the organization or agency doing the rankings, it should include (1) researched using a rigorous design; (2) demonstrated positive effects; (3) having shown sustainable outcomes; and (4) replicable outcomes.
Evidence-based curriculum – Curriculums that utilize interventions that have shown benefit to consumers. These curriculums have been (1) researched using a rigorous design; (2) have demonstrated positive effects; (3) sustainable outcomes; and (4) replicable outcomes.

Evidence-based practice – An approach, framework, collection of ideas or concepts, adopted principles and strategies supported by research.

For example:

Phase III: Collect information on programs and eliminate programs

The committee began to collect information on each of the 500 recognized programs, including target population, research design, outcomes achieved, replication information, sustained effectiveness, and dissemination capability. As the group did this, the elimination phase began. Using CSC’s targeted outcomes and target population, many programs that did not meet CSC’s needs were eliminated. For instance, if a program targeted high school athletes using steroids, it would not be explored further under this first review because this was not CSC’s target population and the organization does not focus on substance use. That being said, at a later stage some of these programs were re-examined for their potential benefit to the community. In this regard, our role was one of advocacy for evidence-based programming in the larger community (e.g. recommending programs to criminal justice, substance abuse and mental health agencies). Additionally, programs were eliminated if not feasible for our specific community and if they did meet many of the evidence-based criteria. At the end of this process, 80 programs required further review.
Phase IV: Create a screening system
Each of the 80 programs went through a screening process that included at least 2-3 reviewers (See Appendix E). Multiple reviewers assured that the screening procedure was thorough and if there were inconsistent findings, meetings were held to discuss findings and validate results. At the end of the process, a matrix of results was created, which included information on alignment of outcomes with our organization, research design, sustainability, replicability (at different sites and with different populations), and whether the evaluations/studies were submitted for peer review. This assisted CSC in narrowing the number of programs to examine in depth.

Phase V: Create a rating system and reassess programs
The final phase included a more in-depth look at the programs. The committee created a data-gathering instrument that assisted in the process. This instrument included the following:

General Description
- Description of program
- Location of program services
- Activities of the program
- Core components (i.e. program cannot operate without)
- Dosage & frequency
- Caseload requirements
- Theory & logic model

Target Population
- Risk & protective factors the program addresses
- Age ranges
- Gender
- Race & ethnicity
- Location it has been implemented
- Income levels of participants
- Eligibility criteria
- Language it is available in
- If it has a fatherhood component
- If it has a social-emotional wellness component
Research Design
- Design(s) of the evaluations
- Sample size
- Methodology (qualitative, quantitative or mixed methods)
- Longitudinal/ follow-up (length of time)
- Measures used (tools used across research studies and reliability/validity)

Outcomes
- Desired outcomes
- Achieved outcomes (including significance level and effect sizes)
- Sustainability of outcomes
- Unintended outcomes (positive or negative)

Program Infrastructure
- Is there a national office & what is the data collection system
- Assessments required
- How is fidelity measured
- Clear Guidelines/ Expectations (including expandability)
- Dissemination capabilities
- Levels of support & technical assistance
- Application process

Professional Development
- Training required
- Credentialing/qualifications of staff
- Coaching offered

Cost
- Cost
- Return on investment studies

While examining the research on the 80 programs, the committee also evaluated each study on its methodological and statistical integrity. Studies were examined by design and plausible alternative explanations of chance, bias, and/or confounding variables. A summary was completed that included design, hypothesis, results, author's conclusions, and reviewer's conclusions. This work is ongoing.
1. **Study design** - In general, descriptive data or post-hoc subgroup analyses are not useful to test hypotheses. Analytic studies (i.e. case-control, cohort, or randomized trials) are needed to test hypotheses. Furthermore, for small to moderate effects, randomized evidence is necessary because the amount of uncontrolled and uncontrollable confounding variables may be as large as the effect sizes being sought (Hennekens, 2009).

2. **Plausible alternative explanations for the observed findings (i.e. chance, bias, and/or confounding)** - If chance, bias, and/or confounding factors are not plausible alternative explanations for the observed findings then one may safely conclude that the findings are the result of the intervention, not outside influences. The following is a sample of what was examined:

   - Chance – statistical significance
   - Bias – attrition, knowing who is receiving the intervention, follow-up rates
   - Confounding – sample size, randomization (Hennekens, 2009).

For those studies determined to have sound methodology, the committee further explored the outcomes in terms of effect sizes. This work is currently underway. CSC felt this scope of work was needed because many peer reviewed articles are not scrutinized as carefully as they should be prior to publication. This will allow the committee to be certain in its recommendations to both CSC and to the community. Once this work is completed, the committee will be able to rate the programs examined in terms of being evidence-based. A preliminary point system has been constructed to assist the committee in making recommendations. A program will need to meet all four of the following criteria to be considered evidence-based:

   - Design (worth 100 points) – must have been researched using an experimental or quasi-experimental design
   - Effect sizes (worth 200 points) – must show at least small effects
   - Replication (worth 40 points) – must show positive findings across studies at various sites or with multiple populations
   - Sustainability (worth 60 points) – must show sustained outcomes for <1 year

The minimum point value to be considered evidence-based is 250. Points increased based on type of design and size of effects. In CSC’s view, all evidence-based programs will

---

10 Completed in conjunction with Charles Hennekens, M.D., Dr.Ph., Patricia Hebert, Ph.D., Gabriel Suciu, Ph.D., and Karen Dodge, Ph.D.
need to be evaluated on research design. To be labeled evidence-based, a program must have had a rigorous study (experimental [RCT] and/or quasi-experimental designs). If a program has more than one RCT they will receive 100 points, one RCT equals 75 points, a quasi-experimental design equals 50 points, and a non-experimental study will receive 0 points. The programs will then be categorized for easy review based on the number of points received. It is important to note, that while this process will increase the level of objectivity, all programs do not necessarily fall neatly into one category or another. For example, if a program meets the criteria for being evidence-based, but there are mixed results – it may be placed in the promising category. In CSC’s opinion, a program that is evidence-based should be consistent in its findings across studies. In addition, if a program has been subjected to a rigorous study design, but if the study had serious flaws, the committee may have decided that the results were suggestive, but not conclusive, thus also placing it in the promising category. It is important to note that labeling a program as promising does not mean that it does not work – CSC just needs additional data or research before moving it up the evidence-based ladder and before promoting widespread dissemination.

The programs that received the highest ratings and that matched CSC’s logic model or pathways document were recommended to CSC leadership for implementation, as well as advocated for within our community.

**Objective #2: Criteria to rate our currently funded homegrown or locally developed programs**

The committees’ second objective was to explore a method of assessing programs that were not evidence-based. There are simply not enough evidence-based programs to meet all the needs of the community. For this reason, many programs are being funded that are not considered evidence-based and it was decided given CSC’s guiding principles, core beliefs around accountability and research, as well as belief in innovation, that CSC needed to assess where these programs would fall on a continuum of effectiveness. Thus, the committee created an assessment tool that would examine core areas supported by research that would impact a programs ability to become evidence-based. The tool has three core areas (See Appendix F for a sample of select assessment tool questions):

- Foundation/planning criteria *(e.g. theory of change, logic model)*;
- Operation/implementation criteria *(e.g. client/staff relations, supervision, fidelity, turnover, staff qualifications, data quality)*; and
- Evaluation criteria *(e.g. research design, intended outcomes, sustainability)*
The foundation and operations sections are dedicated to recognizing where the program excels and where additional support is needed, while the evaluation section determines where the program falls in regard to being rated as evidence-based. Research was used in the development of the tool and it was examined by outside reviewers for feedback.11 Once complete, the tool was piloted using four diverse programs to check for inter-rater reliability. This tool will now be utilized throughout CSC to assist with capacity and system-building and to comprehensively assess each program’s theoretical background, administrative practices, fidelity, and current documented outcomes.

It is crucial that all programs see where they fall on a continuum of becoming evidence-based. As a provider assesses their program according to the continuum, adjustments can be made along the way and they can see where efforts should be focused in order to improve chances for positive outcomes. The following model demonstrates how a locally developed program can begin moving towards becoming evidence-based. For example, a program must be based on research and have a strong theoretical foundation as a first step. A program should also have logic model or theory of change.

Everyone implementing a program must be on the same page as to the desired outcomes and activities needing to be completed with fidelity in order to make progress. During implementation, a program must construct databases to collect data. A program will never achieve the level of evidence-based unless the program can provide quality data to evaluators. In regard to evaluation, a program must first be assessed using a process evaluation – an evaluation that determines whether a program is being implemented with fidelity according to the program model. If service providers are not implementing a program as designed, you can not determine what is working and what is not working. Once it is determined that a program is being implemented with fidelity, you can begin outcome evaluations. A non-experimental or quasi-experimental study is a good place to start. It is not recommended to start with an RCT given the expense.

We know that one of the major problems facing local agencies is that they do not have the resources necessary to prove that their locally developed programs are effective because the type of studies that need to be conducted (randomized controlled studies or experimental designs) are very expensive. For example, on average, a three- to five-year evaluation study can cost several million dollars to fully research effectiveness. However, if a non-experimental study was utilized and yielded positive outcomes, then it may be time to invest in a more rigorous design and it is more probable that funding to do the research will be available. If a quasi-experimental study was used first, a determination should be made to conduct another study in order to begin showing how well it could be replicated. Once at least two evaluations take place, you can begin to discuss where the

11 Reviewers included Gail Chadwick, Vice President of Prevention Matters, LLC.; Gale Held, previous Director of the SAMHSA Model Programs Dissemination Project.
program would fall on the evidence-based continuum (rigorous evaluation with positive effects and significant outcomes, ability to replicate and sustainability).

A Continuum of Evidence-based Programming

Source: Children’s Services Council of Palm Beach County, 2009
References


CSC believes social service organizations have a responsibility to not only positively affect the children they serve, but to help improve the lives of those who they will never serve directly (Pizzigati, Stuck & Ness, 2002). While each organization will have its own mission and vision, target population, and outcomes it is responsible for achieving – each also has a responsibility to advocate that others in its community be data-driven and implement only the best programs available. Evidence-based programs are a good return on investment and more importantly, research shows they work for the children and families they serve.

If a community is exploring evidence-based programs and determining which ones may meet its targeted population and outcomes, the following noteworthy ones should be considered. This section provides detailed information on the programs rated highest in terms of evidence.12

12 Program snapshots completed by Johanna Dene-Sharp & Lisa Willaims-Taylor, Ph.D.
Multisystemic Therapy (MST)

Program Overview
Multisystemic Therapy (MST) was developed in the late 1970s to address shortcomings of traditional mental health services for youthful offenders. The goals of MST, as stated by the program are to 1) decrease rates of antisocial behavior and other clinical problems; 2) improve functioning (e.g., family relations, school performance); and 3) achieve these outcomes at a cost savings by reducing the use of out-of-home placements (e.g., incarceration, residential treatment, hospitalization).

MST is based on the social-ecological model, positing that youth antisocial activities result from a confluence of factors related to family, school, peers, and neighborhood. The program targets each of these factors during intervention. Sessions are held weekly in the youth's home or school. Each session lasts several hours and is facilitated by a Master's-level therapist. Families attend and collaborate in the treatment plan, which strives to empower families to enhance protective factors by improving parental effectiveness and increasing the social supports of extended family, friends, church, or other resources. At the same time, the program seeks to diminish risk factors by separating the youth from deviant peers. The program also seeks to improve educational or vocational performance.

Evaluation Methods and Program Outcomes
Multisystemic Therapy Services (2007) provides the following overview of research and outcomes.

The first controlled study of MST was published in 1986, and randomized trials of MST were initiated in the 1990's. To date, 18 published studies (16 randomized, 2 quasi-experimental) demonstrate effectiveness or promise in the treatment of youth offenders. Two of the first randomized studies were the Simpsonville, South Carolina and the Missouri Delinquency Project.

In the Simpsonville trial, 84 juvenile offenders were randomly assigned to MST or the usual services of the Department of Juvenile Justice (DJJ). Participants had an average of 3.5 arrests, 9.5 weeks of incarceration, were about 15 years old, and were 77% male. Fifty-six youth completed the program and were included in the post analysis. All 84 youth were included in the 59 week and 2.4-year follow-up. At program completion, MST participants had fewer reports of criminal activity than the control group. At the 59 week follow-up, MST participants showed 43% fewer arrests than DJJ participants. DJJ youth had three times as many weeks of incarceration on average than MST youth. After 2.4 years, 20% of DJJ participants had not been arrested, as opposed to 39% of the MST youth.
In the Missouri Delinquency Project, 176 chronic juvenile offenders and their families were randomly assigned to either MST or individual therapy (IT). Participants averaged 4.2 arrests, were about 15 years old, and were 68% male. One hundred twenty-six completed the program. At program completion, MST participants showed decreased behavior problems and improved family relations. At the four year follow-up, MST youth were arrested less often and for less serious crimes than IT youth. For example, the arrest rate for substance abuse related crimes was 4% for MST participants versus 16% for IT participants. MST youth who completed the program had a 22% arrest rate, those who only partially completed the program a 47% arrest rate, while IT program completers and partial completers showed a 71% arrest rate. The arrest rate for those youth who refused all treatment was 83%. In a 14 year follow-up, MST participants had 54% fewer arrests and 57% fewer days of confinement than IT participants. This difference applied to violent, drug-related, and non-violent offenders.

In a third randomized trial, treatment fidelity was examined. MST experts were not significantly involved in the actual program, but played more of a supervisory role, and results were less pronounced than in earlier studies. Several other large-scale, multi-site studies of MST programs have been conducted to examine quality assurance, with findings that show MST supervisory competence and therapist fidelity directly relating to successful program outcomes.

Substance abusing offenders were studied separately from other offenders in the earlier trials, and were found to have significantly lower substance use in a four-year follow-up than those in the control group. In a separate trial specifically targeting substance-abusing offenders, MST was more effective in reducing rates of subsequent drug use than community treatment post-treatment and in follow-up studies.

MST has also been shown to be effective in reducing recidivism, improving family relations, decreasing violence toward peers, and increasing academic performance for sexual offenders than traditional intervention methods. Finally, MST is demonstrating promise in addressing acute psychiatric emergences, parental maltreatment of youth, as well as helping to address youth health concerns that require ongoing monitoring.

**Program Contact Information**

MST Institute  
Molly Brunk, Ph.D., Director of Quality Assurance  
Email: msti@mstinstitute.org  
Website: www.mstinstitute.org
LifeSkills Training (LST)

SAMHSA (2009a) and Promising Practices Network (2009a) websites provide the following information.

Program Overview

Botvin’s LifeSkills Training (LST) is a substance abuse prevention curriculum offered to middle school students, with booster programs in high school. The program is designed to teach students social and self-management skills, including skills in resisting peer and media pressure to smoke, drink, or use drugs, and inform students of the immediate consequences of substance abuse. The curriculum is typically taught by school teachers for three consecutive years, starting in sixth or seventh grade, and utilizes facilitated discussion, structured small group activities, and role-playing scenarios. Teaching materials include handouts, multi-media presentations, and a lesson plan for teachers.

LST is based on the social influence and the enhancement models of prevention, and as such addresses risk and protective factors, as well as teaching personal and social skills to enhance resilience. Unlike traditional prevention programs, LST focuses on the immediate negative effects of substance use, hypothesized to be a more effective preventative strategy than emphasizing long-term health effects.

Evaluation Methods and Program Outcomes

Since 1980, LST has been evaluated in 30 studies based on at least 13 samples. The majority of the studies utilized randomized experimental designs. Seven individual evaluation studies have been conducted by external research groups and 10 replication studies have been conducted. The curriculum has been extensively evaluated with white, middle class suburban and rural youths, and evaluations have also been conducted with African American urban youth. The program was originally developed as a smoking prevention curriculum, and over 13 studies have demonstrated that participation in LST significantly reduces participants’ risk of becoming new or frequent smokers. Similar studies have shown that the program significantly reduces the risk of using alcohol and marijuana.

In three studies reviewed by SAMHSA’s National Registry of Evidence Based Programs and Practices (2009a) (Study 1, Study 2, Study 3), LST interventions produced outcomes in substance abuse prevention and normative beliefs about substance use and refusal skills. A fourth study (Study 4) produced outcomes in violence and delinquency reduction (Botvin, Baker, Dusenbury, Botvin & Diaz, 1995; Botvin, Griffin, Diaz & Ifill-Williams, 2001a and 2001b; Griffin, Botvin, Nichols & Doyle, 2003; Spoth, Randall, Trudeau, Shin & Redmond, 2008; Trudeau, Spoth, Lillehoj, Redmond & Wickrama, 2003; Botvin, Griffin & Nichols, 2006). Measures of substance use, normative beliefs about substance use, and
violence and delinquency were measured using the Life Skills Training Questionnaire (LSTQ).

Frequency and quantity of substance use was self-reported by study participants in studies one, two and three. All participants were middle school students. In studies one and three, the majority of the participants were white. In study two, the majority was African American and other minorities. In all three studies, the control group was comprised of individuals receiving a typical drug prevention program. Intervention groups participated in the LST intervention consisting of classroom time spent on materials and ongoing consultations. In study one, a third group who participated only in multi-media presentations of LST with no consultation was included.

Outcomes for study one included significant reductions in cigarette, alcohol, and polydrug use when compared to the control group. Additionally, the fidelity sample in this study showed significantly lower overall cigarette, alcohol, and polydrug use than the control group. This held as well for study two, both at posttest and at the one-year follow-up when compared to the control group. Participants in LST in study two also showed less binge drinking than the control group, at both the one- and two-year follow-up.

Initiation to drug use decreased as a result of participation in LST for participants in study two and study three. In study two, a subset of participants who were at high risk for initiation showed less smoking, drinking and polydrug use compared to the control at post-test. Participants in study three showed a lower rate of increase in substance initiation from pretest to posttest and one-year follow-up. Five years later, they also had lower cigarette and marijuana initiation than the control.

Middle school students who participated in LST demonstrated less violence and delinquency as measured by the LSTQ at the three-month follow-up than the control group. Effects were highest for participants who received at least half of the LST program.

**Program Contact Information**
National Health Promotion Associates
711 Westchester Avenue
White Plains, NY 10604
Phone: (914) 421-2525
Fax: (914) 421-2007
Email: lstinfo@nhpamail.com
Olweus Bullying Prevention Program

Program Overview
The Olweus Bullying Prevention Program was developed in Norway in the 1980s as a result of the highly publicized suicides of three youths who had been systematically bullied. The program’s first implementation in the United States was in the 1990s in South Carolina.

The program is administered in a school setting utilizing three components: school, classroom, and individual levels. Each component is part of a whole program in which adult awareness of and behavior toward bullying are the keys to successful reduction or elimination of bullying. A general school-level component attempts to create an environment in which bullying is not tolerated. An anonymous student questionnaire is distributed to assess the extent of the bully/victim problem in the school. Staff engages in training to recognize and confront bullying with negative consequences. Parents are involved in the formation and execution of the program when feasible. A classroom-level component includes reinforcement of the school-wide rules against bullying, holding regular meetings to increase knowledge and empathy, and holding informational meetings with parents. At the individual child level, interventions are carried out with children who bully, children who are bullied, and the involved parents.

Evaluation Methods and Program Outcomes
In the mid-1980s, using a quasi-experimental design, researchers conducted an evaluation of 2,500 children in elementary and middle schools in Norway. According to Olweus (1991), results showed significant reductions in self-reported delinquent behaviors and significant improvements in the social climate of the classroom. In addition, self-reported bullying and victimization showed 50% reductions. Greater reductions in bullying occurred within the schools that implemented all the essential components of the program than within the schools that implemented only part of the program. Furthermore, the results appeared to be cumulative, with improvements in some outcomes more marked 20 months after the intervention than after eight months.

A second Norwegian study from the late 1990s involved 3,200 junior high school students in 14 intervention groups and 16 control groups. According to researchers, in the implementation groups, bullying and victimization decreased 21% and 38%, compared to the control groups, where no significant changes in victimization were found, and a 35% increase in bullying (Olweus, 2004; Olweus, Limber & Mihalic, 1999). A similar study in Sweden in the late 1990s showed a 42% reduction in self-reports of victimization (33% for girls and 48% for boys) and a 52% reduction in self-reports of bullying others (64% for girls and 45% for boys).
Two studies were conducted in the United States in the mid- and late-1990s. In South Carolina, 18 middle schools implemented the program, resulting in large, significant decreases in boys' and girls' reports of bullying others and large, significant decreases in boys' reports of being bullied and social isolation (Limber, 2004). In 12 elementary schools in the Philadelphia area, the six schools that implemented the program with at least moderate fidelity showed significant reductions in self-reported bullying and victimization and significant decreases in adults' observations of bullying in the cafeteria and on the playground as compared to control schools (Black, 2003). Over four years of program implementation, bullying decreased by 45%.

Two additional studies were conducted in the United States in the early 2000s, one with students from 10 public middle schools, and another with students from three public elementary schools, also yielded some positive results. In the middle school study, among white students physical victimization decreased 37%. However there was no drop in victimization for minority students. In the elementary school study, reports of being bullied decreased 21% after one year and 14% after two years. Reports of bullying others decreased 8% after one year and 17% after two years. In addition, parents and students were significantly more likely to believe that adults at the school intervened to stop bullying (Bauer, Lozano & Rivara, 2007; Pagliocca, Limber & Hashima, 2007).

**Program Contact Information**
Marlene Snyder, Ph.D., Research Associate
Institute on Family & Neighborhood Life
Clemson University
Phone: 864-710-4562
Fax: 406-862-8971
E-mail: nobully@clemson.edu
Nurse-Family Partnership

Program Overview
The Nurse-Family Partnership (NFP) is a home visitation program targeting low-income, at-risk first-time mothers, most of whom are teenagers. NFP is implemented in rural, suburban, and urban areas, primarily through local, county or state health departments. Mothers’ enrollment in the program is voluntary.

The primary goals of NFP are to (1) to improve pregnancy outcomes by improving health behaviors; (2) to improve child health, development, and safety by promoting responsible and competent care for children; and (3) to promote the mother’s self-sufficiency and well-being by assisting with pregnancy planning, educational achievement, and employment (Promising Practices Network, 2009b).

Home visits to the mother by highly trained registered nurses begin early in the second trimester of pregnancy, and follow guidelines set by the program. Visits are weekly for the first month after program enrollment, and then every other week until the baby is born. These visits cover topics such as medical checkups, alcohol, drug, and tobacco use, and nutrition. For a short time after the birth (up until six weeks after delivery), the nurse visits weekly, providing information about infant care, providing a safe and nurturing environment for the child, child discipline, and health care. Then, until the child is 20 months old, the nurse visits every other week, assisting the mother with education and employment plans, issues of future pregnancies, choices about others who are involved with her child, and other positive health related behaviors. The nurse visits monthly during the last four months of the program, continuing to teach competent care of children and discussing maternal personal development.

The program was originally created to address some of the main factors associated with early antisocial behavior: neuro-developmental impairment of the fetus, dysfunctional care giving, and poor maternal life-course development. The theoretical underpinnings of the programs are human ecology theory, human attachment theory, and the theory of self-efficacy.

Evaluation Methods and Program Outcomes
NFP has been evaluated in three separate randomized controlled trials. An Elmira, New York study in 1977 targeted primarily white women; a Memphis, TN study in 1988 targeted primarily African American women, and a Denver, CO study in 1994 targeted women of different races and ethnicities. In the Elmira and Memphis studies, women were recruited from health clinics who presented with at least two of the three study requirements: first time pregnancy, unemployment, and less than 12 years of schooling.
Women were randomly assigned to treatment or control groups using stratification according to various demographic conditions. Women for the Denver study qualified if they were first-time mothers with no insurance and eligible for Medicaid.

Dropout rates for the studies varied between 15% and 21%. Longitudinal studies were conducted on each cohort, including a 13-year follow-up with the Elmira cohort. NFP demonstrates consistent program effects. Consistent program effects include: improved prenatal health, fewer childhood injuries, fewer unintended subsequent pregnancies, increased intervals between births, increased maternal employment, and “increased school readiness for children born to mothers with low psychological resources” (Nurse-Family Partnership, 2009).

Findings from the 15-year follow-up of the Elmira study demonstrate that child participants experienced fewer incidents of child abuse and neglect, and fewer self-reported arrests, convictions and parole violations. Mothers demonstrated less time on welfare, and fewer arrests and convictions. A nine-year follow-up of the Memphis cohort showed fewer childhood injuries and accidental ingestions and fewer days in the hospital for injuries at two years, and a lower mortality rate and better school achievement and readiness at nine years. The latter outcome presented only for a subset of children of mothers with low psychological resources. The mothers at the nine-year follow-up had spent less time on welfare, had fewer subsequent pregnancies, fewer subsequent underweight babies, and used fewer substances (Coalition for Evidence-Based Policy, 2009).

Program Contact Information
1900 Grant Street, Suite 400
Denver, CO 80203
Direct phone: 303-327-4240
Toll free: 866-864-5226
Fax: 303-327-4260
Email: info@nursefamilypartnership.org
Positive Parenting Program (Triple P)

Program Overview
Positive Parenting Program (Triple P) aims to prevent severe behavioral, emotional, and developmental problems in children by enhancing the knowledge, skills, and confidence of parents. The program is designed for preschoolers and primary school-aged children, as well as providing a distinct program for early teens.

The program offers five levels of intervention, each with greater involvement based on a continuum of increasing needs. Additionally, the reach of the intervention may be very broad - targeting an entire population, to very narrow - reaching only one family, depending on the level. Triple P offers two additional programs: one (Stepping Stone's) for parents of children with disabilities, and the other (Pathways) for parents at risk of child maltreatment.

Level 1 - Universal Triple P disseminates messages universally through printed materials, infomercials, radio time, press releases, or other media events. Self-sufficient parents can implement strategies on their own, and families with more need may be motivated to seek a higher level of intervention.

Level 2 - Selected Triple P is a strategy that is put into place at health facilities such as doctors' offices and clinics, that parents frequent but that are not associated with the stigma of a mental health intervention. Here a parent may ask for advice on a specific behavioral problem and receive printed materials or a short video. In these circumstances the particular problem is mild to moderate, and not complicated by other dysfunctions.

Level 3 - Primary Care Triple P targets the same population as Level 2, but the intervention is more intensive and involves parental active skills training.

Level 4 - This level targets families of children with multiple behavior problems in a variety of settings, but who do not meet the diagnostic criteria of a personality disorder. At this level, parents may participate in individual or group sessions with a therapist, or may pursue a self-directed program.

Level 5 – This level is for parents who have participated in a Level 4 intervention, but who continue to struggle with the child's behavior problem. This level is appropriate for families of children with severe behavioral issues who are additionally experiencing family adversity factors.

Evaluation Methods and Program Outcomes
Triple P has been evaluated in almost two dozen randomly controlled trials, from the
mid-1980s to the present. These studies demonstrate clinically meaningful outcomes for families, ability to be replicated, effectiveness of different levels of intervention, consumer acceptability, and effectiveness with a range of families.

For example, a randomized controlled study published in 2000 compared Standard Triple P and Enhanced Triple P in families with children with oppositional defiant disorder or conduct disorder and mothers with major depression (Sanders & McFarland, 2000). A group of families on a waiting list acted as the control. Forty-seven families participated, with an attrition rate of 21% for Standard Triple P and 13% for Enhanced Triple P. The study set out to measure child disruptive behavior, parent-child interaction, parenting confidence and parental adjustment. Both interventions reduced observed and reported child disruptive behavior and decreased parental depression. Both interventions significantly increased parenting confidence, with all results maintained at the six-month follow-up.

In another study published in 2000, researchers compared Standard, Enhanced, and Self-Directed Triple P, with a waitlist group as a control (Sanders, Markie-Dadds, Tully & Bor, 2005). Three hundred and five individuals participated, with an attrition rate of between 8% and 24%, depending on the group. For those in the Standard and Enhanced Triple P groups, self-reported dysfunctional parenting was significantly reduced. These two groups also showed significant improvements in mother-reported disruptive behaviors when compared to the control. In all three intervention groups mothers reported greater parenting confidence than the control.

In a study published in 2009, 18 counties in the southeastern U.S. were randomly assigned to receive Triple P or to act as a control with services as usual. A total of 649 service providers received training after randomization. Prior to treatment the two groups did not differ on substantiated child maltreatment, out-of-home placement, or hospitalization of emergency room visits. After two years of service, treatment counties differed significantly on measures of child maltreatment (Prinz, Sanders, Shapiro, Whitaker & Lutzker, 2009).

**Program Contact Information**
PO Box 12755, Columbia, SC 29211
Phone: (803) 451-2278
Fax: (803) 451-2277
Email: contact.us@triplep.net

---

Promoting Alternative Thinking Strategies (PATHS)

Program Overview
Promoting Alternative Thinking Strategies (PATHS) is a school-based curriculum designed to promote emotional and social competencies and reduce aggression and behavior problems in elementary school-aged children, with the longer-term goal of providing skills to find nonviolent solutions to social problems. A preschool component also exists. A secondary goal is to improve the educational process in the classroom. Children are taught to express and manage their feelings, to control impulses, improve decision-making, and other social and emotional skills.

PATHS is based on five conceptual models: the Affective-Behavioral-Cognitive-Dynamic Model of Development; an eco-behavioral systems orientation; neurobiology and brain structuralization/organization; developmental psychodynamic theory; and emotional intelligence (Blueprints, 2009).

The program ideally starts in kindergarten and continues through grade five, becoming increasingly complex as the children progress to higher grades. PATHS is administered in the classroom on a school-wide basis three times a week for 20 to 30 minutes at a time by teachers and counselors. It is taught using developmentally-based lessons in emotional literacy, self-control, social competence, positive peer relations, and interpersonal problem-solving skills. The three major components of the program are Readiness and Self-Control, Feelings and Relationships, and Interpersonal Cognitive Problem Solving. Teachers receive training in the PATHS curriculum in a two- to three-day workshop and bi-weekly meetings with a consultant. It is primarily classroom based, but includes generalization activities and information to be used outside of the classroom and with parents.

Evaluation Methods and Program Outcomes
SAMSHA’s National Registry of Evidence-Based Practices (2009b) examined several studies of the PATHS curriculum, the results of which were published between 1996 and 2007. The following information was excerpted from SAMSHA’s review.

All studies were experimental in design (using randomized controlled groups). These studies demonstrated improvement in protective factors and reduction of behavioral risks across a wide variety of elementary-school-aged children who experienced the PATHS curriculum. These findings also held across a variety of raters including teacher reports, self-reports, and child testing and interviewing.
A 2004 study of first through third graders in special education classes who received a modified version of the PATHS curriculum showed significant outcomes when compared to a control group in variables related to internalizing and externalizing behaviors and depression (Kam, Greenberg & Kusché, 2004). Second and third graders in a separate 2006 study showed significant positive differences with their control group peers with regard to internalizing and externalizing behaviors and neuro-cognitive capacity (Riggs, Greenberg, Kusché & Pentz, 2006).

In a separate 1999 study, first-graders that received the PATHS curriculum in a school described as high risk (due to relatively high levels of delinquency and juvenile arrests) differed significantly from the control group with regard to externalizing behaviors and learning environment (Conduct Problems Prevention Research Group, 1999). Finally, PATHS was tested in a preschool environment in a study published in 2007. Preschoolers that received the PATHS curriculum showed significantly higher levels of emotional knowledge, internalizing behaviors, and social-emotional competence, as measured by both teachers and parents (Domitrovich, Cortes & Greenberg, 2007).

**Program Contact Information**
Channing Bete Company, Inc.
One Community Place
South Deerfield, MA 01373-0200
Phone: (877) 896-8532
Fax: (800) 499-6464
E-mail: custsvc@channing-bete.com

Carol A. Kusché, Ph.D.
Director of Training
PATHS Training, LLC
927 10th Avenue East Seattle, WA 98102
Phone: (206) 323-6688
E-mail: ckusche@comcast.net
Program Overview
Big Brothers Big Sisters of America (BBBS) is a community-based mentoring program with the goal of improving the social, family and/or peer relationships, and academic achievement of participating youth.

Each youth program participant is matched with a Big Brother or Big Sister by an assigned case manager following an application by the parent or guardian. A case manager visits the youth's home to ensure parental/guardian support for the program, completes a home assessment, and learns as much as possible about the youth to ensure a mutually beneficial match. Working with the youth and parent or guardian, the case manager creates a series of goals for the relationship. The primary goal is typically to create a positive and supportive relationship between the youth and mentor, who has been rigorously screened to ensure youth safety. After the primary goal is accomplished, more specific goals may be pursued from suggestions by the case manager.

After the formation of the relationship between the youth and mentor, the case manager makes weekly contact for the first month, monthly contact through the first year, and every three months thereafter. The duration of the relationship is ideally at least one year, although the relationship may terminate voluntarily at any time if either party wishes. The targeted age range for the youth is six through eighteen. The youth typically has associated risk factors such as residence in a single-parent home or a history of abuse or neglect.

The mentor spends three to five hours a week with the youth. During this time activities can be chosen by the mentor and the youth. However the diligence and stability of the relationship is more important to the success of the program than the particular activities since it is based on the premise that a role model who is positive, caring, and supportive will positively impact a variety of behavioral outcomes. Thus, the program does not use a specific behavioral intervention nor does it target a specific behavior, such as academic improvement or drug use.

Evaluation Methods and Program Outcomes
During 1992 and 1993, BBBS conducted a randomized controlled study of over 1,000 youth, ages 10 to 16 served by various agencies across the country. Youths were randomly assigned to an intervention group, which agency staff attempted to match with a mentor, or a control group that was placed on a waitlist for the duration of the study (18 months). BBBS successfully provided 78% of the youths in the intervention group with a mentor. Sixty-two percent of the youths in the study were male, 56% were minorities, and 40% lived in households receiving food stamps and/or welfare.
All participants were interviewed at baseline. Eighty-four percent of the youth completed the follow-up interview and were included in a subsequent analysis. The analysis utilized a four-part framework, examining the effect on boys, girls, whites, and minorities.

Results of the study indicated that BBBS youth were 46% less likely to initiate illegal drug use, 27% less likely to initiate alcohol use (marginally significant), 32% less likely to hit someone, had a slightly higher grade point average, and were 52% less likely to skip a day of school.

Outcomes varied by group. For example, minority boys were 68% less likely to start using illegal drugs, while no significant differences were found for whites. Minority girls were 73% less likely to initiate illegal drug use, a difference that was marginally significant, while white girls did not differ from the control group. Minority girls were 54% less likely to initiate alcohol use than were control group girls, a difference that was marginally significant. Neither white boys or girls, or minority boys differed significantly from the control group in terms of alcohol use initiation. Girls who participated in BBBS attained significantly higher GPAs than did the comparison girls, with an average GPA of 2.84 versus an average of 2.67 (Promising Practices Network, 2009c).

**Program Contact Information**
Big Brothers Big Sisters National Office
230 North 13th Street
Philadelphia, PA 19107
Phone: (215) 567-7000
FAX: (215) 567-0394
Check and Connect

Program Overview
Check and Connect is a dropout prevention program directed to children and youth who are at high risk of school failure. The model was originally developed for urban middle school students, but is currently being replicated in grades K-12 in urban and suburban settings.

The key features of Check and Connect involve close monitoring of the at-risk student, and creating a relationship with him or her, and with the family that will allow intervention when needed. The “check” component involves facilitating the continuous assessment of student engagement levels with school, and to guide interventions when needed. The “connect” component involves two levels of intervention: basic and intensive. The basic intervention is the routine checking on student progress, relationship maintenance, and problem-solving. The intensive intervention is reserved for those students who are showing warning signs of withdrawal. In these cases, students are typically contacted each morning to ensure school attendance, and monitored more closely in terms of their attendance and school achievement. Over two-thirds of the participants in the program receive intensive intervention at one time or another.

The monitor/mentor is responsible for facilitating a student’s connection with school. The role of the monitor is predicated on the notion that the presence of an adult to motivate and foster the development of life skills is a protective factor in a child's life. The monitor works with the child and the family to help overcome the obstacles that prevent the child from fully engaging in the learning experience.

Evaluation Methods and Program Outcomes
Five studies since 1992 show the effectiveness of Check and Connect to reduce truancy and dropping out. The samples range from elementary-school-age cohorts to high school cohorts, were selected from urban and suburban areas, and involve both typical students and students with learning and/or emotional disabilities.

In two separate studies, children with learning and emotional disabilities were chosen from urban schools. The original pilot observed students in middle school and during the transition to high school between 1992 and 1995. The other study examined students in grades 9-12 between 1996 and 2001. In both studies, a majority of the participants were male, African American and eligible for free or reduced lunch.

Outcomes showed that more treatment students were attending school at the end of ninth grade than similar students randomly assigned to the control group (91% vs.
68%); more treatment students were on track to graduate within five years at the end of
ninth grade than similar students randomly assigned to the control group (68% vs. 29%);
treatment students were also significantly less likely to drop out of school than students
in the control group over the 4-5 year period (39% vs. 63%); treatment students were
more likely to be enrolled in an educational program or to have completed high school
by the end of the study than students in the control group (54% vs. 34%); and twice as
many youth in the treatment group who participated in the study for a fifth year went on
to complete high school compared to similar students in the control group (18% vs. 6%)
(Sinclair, Christenson, Evelo & Hurley, 1998; Sinclair, Christenson & Thurlow, 2005).

Two other studies of children with and without learning or emotional disabilities were
conducted, one between 1996 and 2002, and the other between 1997 and 2001, both
in suburban settings. The first targeted children ages 11 to 17 and the second targeted
children with referrals in K-6. Outcomes of these studies showed that the percentage of
treatment group students absent more than 15% of the time was reduced from 45% at
referral to 32% after two years; the percentage of treatment group students present at
least 95% of the time increased from 11% at time of referral to 34% after two years in one
study, and from 17% at referral to 40% after two years in the other. Lastly, the percentage
of students arriving to school on time increased from 42% at the time of referral to 86%
after two years (Lehr, Sinclair & Christenson, 2004; Sinclair & Kaibel, 2002).

In a study examining the effectiveness of Check and Connect to improve reading skills
in kindergarten, the program was shown to produce statistically significant differences
in early literacy and engagement for students receiving the program for two years
compared to the control group. The treatment also yielded meaningful changes in
teachers’ perceptions of children's behavior and academic competence (O'Shaughnessy
et al., 2004).

Program Contact Information
Ann Mavis
Institute on Community Integration
University of Minnesota
150 Pillsbury Drive SE
6 Pattee Hall
Minneapolis MN 55455
Phone: 612-624-1489
Email: checkandconnect@umn.edu
**Functional Family Therapy (FFT)**

**Program Overview**
Functional Family Therapy (FFT) is a family-based intervention program for youth ages 11-18 demonstrating behaviors such as delinquency, violence, substance use, conduct disorder, oppositional defiant disorder, or disruptive behavior disorder, with depression a frequent co-morbid condition. Younger siblings of referred adolescents may also be involved in the sessions. FFT is a short-term intervention providing, on average, eight to twelve one-hour sessions for mild cases and up to 26-30 sessions for more difficult situations. Sessions are typically spread over a three-month period. The program can be delivered in a variety of settings including home, school, or clinical and juvenile justice facilities.

FFT is organized into three distinct phases that build upon each other. The primary function of the first phase is to encourage sustained program participation by the youth and family. The therapist will attempt to reduce negativity and resistance and improve communication between all participants. The second phase, behavior change, involves developing and implementing individualized change plans, altering presenting delinquency behavior, and building relational skills. These changes may be accomplished through the use of communication training, specific tasks and technical aids, and/or basic parenting skills. In the third phase, generalization, the group works to maintain and generalize change and prevent relapses. The therapist also ensures that community resources are put into place to support the changes targeted in the sessions.

**Evaluation Methods and Program Outcomes**
Thirteen studies to date demonstrate the effectiveness of FFT in reducing recidivism and out-of-home placement. Varied assessments are utilized in the studies, including outcomes evaluations, quasi-experimental and experimental evaluations. The following information was excerpted from Child Trends (2004) and summarizes eight experimental evaluations.

Youth targeted for intervention are typically at-risk adolescents or adolescents engaging in delinquent behavior, and sometimes parents and siblings. Sample sizes ranged from 27 to 166 adolescents. Most of the evaluations took place in Salt Lake City, UT, with other evaluations in Pennsylvania, Indiana, and Sweden. Typically the FFT treatment group was compared with one to three control groups that received no treatment, received FFT along with individual therapy, or received alternative therapies or other social services. Follow-up for each evaluation ranged from program termination up to 30 or 40 months. Evaluations focused on a variety of factors, including negativity, positive communication, parent-child and family processes, self-esteem, relationship quality, family concept,
as well as behavior and mental health problems. The majority of studies also assessed recidivism, one of the larger aims of the FFT intervention.

In general, outcomes for FFT treatment groups differed significantly from control groups. FFT participation was associated with significant improvements in positive communication, a positive family concept and improved family interaction processes. Program participation was also better for FFT participants than with the control groups, as evidenced by greater parent involvement and decreased family dropout rates. Finally, studies found that FFT reduced maternal depression, as well as anxiety among adolescents.

Studies of FFT show that FFT participants were less likely than control groups to engage in future offenses. Recidivism rates among FFT participants ranged from 26% to 50%, whereas rates among control groups ranged from 47% to 88%. Finally, an evaluation that incorporated siblings of delinquents found that recidivism in siblings who received FFT was 20% compared with 63% of those who received an alternative family therapy.

**Program Contact Information**

Holly DeMaranville - FFT Communications Coordinator
1251 NW Elford Drive
Seattle, WA 98177
Phone: (206)369-5894
Email: hollyfft@comcast.net
Incredible Years (IY)

Program Overview
Incredible Years (IY) is a series of developmentally-based curricula targeting children two to ten years old who are at high-risk of developing, or have developed behavior problems. The program also targets these children's parents and teachers. The parent, child, and teacher curricula are to be used jointly with the objective of reducing behavioral and emotional problems and promoting emotional and social competence in children. The IY curricula is based on Gerald Patterson's social learning model which emphasizes the importance of the family and teacher socialization processes, especially those affecting young children.

The main focus of the BASIC Parent-Training Program - Early Childhood (the core of the parent curriculum) is to teach parents how to help children learn, how to give effective praise and incentives, play with children, use limit-setting, and handle misbehavior. Four supplemental parenting curricula address parents’ interpersonal skills, parenting older children, and encouraging school readiness. The parent curricula are delivered through a series of multi-media presentations to encourage group discussion and problem solving.

Two components make up the Child Training portion of IY: Dina Dinosaur Classroom Curriculum and Dina Dinosaur Child Training Program for small groups. The classroom curriculum is for four- to eight-year-old students and is taught to the entire classroom two to three times a week for 20 to 30 minutes per session over multiple years. The second component, Dinosaur Child Training, is targeted to four- to eight-year olds who are displaying conduct problems. Five to six children at a time participate in two-hour sessions once a week for 20 to 22 weeks which are delivered by counselors or other staff. Both components seek to strengthen children's social and emotional competencies such as appropriate classroom behavior, effective problem-solving strategies, understanding and communicating feelings, managing anger, and practicing friendship and conversational skills at home and at school.

A teacher-training component encourages teachers to use praise and encouragement, proactive teaching strategies, and offers ways to manage inappropriate classroom behavior and build positive relationships with students.

Evaluation Methods and Program Outcomes
More than a dozen experimentally designed studies of IY have shown statistically significant outcomes in the areas of positive and nurturing parenting including decreasing harsh, coercive and negative parenting. Results also show decreases in child behavior problems and increases in prosocial behavior, social competence, and school readiness skills. Outcomes also indicate increases in parent bonding and involvement
with teacher and school and improvements in teacher classroom management skills. Evaluation methods include parent reports, teacher reports, independent observations, and tests of children's social problem solving. Study methods have included comparisons of intervention groups with non-treatment control groups and/or other interventions and have included various ethnic groups.

The following outcomes were found in six experimental studies published between 1994 and 2007 by the program developer (Reid, Webster-Stratton & Baydar, 2004; Reid, Webster-Stratton & Hammond, 2003 and 2007; Webster-Stratton, 1994; Webster-Stratton & Hammond, 1997; Webster-Stratton, Reid & Hammond, 2004 and 2001).

Treatment groups for the parenting program received the parent training in combination with child and/or teacher training, while the control groups received only child and/or teacher training. Treatment groups for the children's programs received parent, child, and/or teaching training; while the control group did not receive the IY intervention (i.e. children were either on a waiting list or received Head Start only).

- Parents in treatment groups showed a significant increase in positive and nurturing parenting relative to parents in comparison groups.
- Parents in treatment groups showed a significant reduction in harsh, coercive, and negative parenting relative to parents in comparison groups.
- Children in treatment groups showed a significant reduction in behavior problems at home or at school compared with controls.
- Children in treatment groups showed a significant increase in positive behaviors at home or at school compared with controls.
- The combination of parent and teacher training resulted in significantly higher levels of parent-child bonding and parent involvement with the teacher and school compared with child training only and control conditions.
- Four of the treatment conditions studied—(1) child training alone, (2) parent and teacher training, (3) child and teacher training, and (4) parent, child, and teacher training—resulted in significantly better teacher management in the classroom compared with control conditions.

**Program Contact Information**

Incredible Years  
1411 8th Avenue West, Seattle, WA 98119  
Phone: (888) 506-3562 or (206) 285-7565  
Fax: (888) 506-3562
Multidimensional Treatment Foster Care (MTFC)

Program Overview
Multidimensional Treatment Foster Care (MTFC) offers an alternative to group residential treatment, incarceration, or hospitalization for adolescents engaging in chronic and severe criminal behavior or suffering from emotional disturbances. MTFC's two stated objectives are “to create opportunities so that youth are able to successfully live in families rather than in group or institutional settings, and to simultaneously prepare their parents, relatives, or other aftercare resources to provide youth with effective parenting so that the positive changes made while the youth are placed in MTFC can be sustained over the long-run” (Multidimensional Treatment Foster Care, 2009).

Adolescents reside with a community foster family that has received intensive training in behavior management and is in daily contact with MTFC staff. Program duration is typically six to nine months long. During that time, the foster family tracks and regulates the youth's behavior, while the youth is mentored and encouraged to develop academic and positive living skills. The setting provides daily structure with clear expectations and limits, and the youth is rewarded with increased freedom for positive behavior, and is sanctioned with increased restrictions for non-compliant behavior. Association with deviant peers is deterred for the duration of the program.

Evaluation Methods and Program Outcomes
MTFC has been the subject of eight randomized trials evaluating feasibility and effectiveness. It is currently undergoing three additional randomized trials.

A study published in 1997 evaluated the impact of MTFC on a sample of 79 adolescent males, ages 12 to 17 that averaged 13 prior arrests and 4.6 prior felonies (Chamberlain & Reid, 1997). The youth were randomly assigned to either MTFC (n=37) or group care (GC) (n=42) placements after being mandated by juvenile courts to receive out-of-home care. Boys were evaluated at baseline, three months into treatment, and then every six months for two years following treatment. Measures included delinquent and criminal activity, supervision, discipline, positive relationships with adult caregivers, and association with deviant peers and were based on self-report, caregiver-report and official arrest data.

At the three-month follow-up, MTFC youth did better on measures of being supervised, receiving more consistent discipline, better relationships with adults and less association with deviant peers than their GC counterparts. MTFC boys reported fewer daily problem behaviors than GC boys; however, this was not consistent with caretaker-reports which rated the boys similarly.
At one year post-treatment, MTFC boys had significantly fewer arrests and were incarcerated significantly less often than GC boys. MTFC participants were also significantly more likely to report no further arrests post-treatment. Finally, boys receiving MTFC reported significantly less criminal and delinquent behaviors.

A second randomized controlled trial of 85 chronic and serious juvenile male offenders showed several outcomes (Eddy, Whaley & Chamberlain, 2004). The sample averaged 15 years of age, and had an average of 14 lifetime criminal offenses and more than four felonies. The overall attrition rate at the two-year follow-up was 7% for official criminal data and 21% for survey data. Data were collected for all MTFC participants, regardless of whether or not they completed the entire program.

At the two-year follow-up, MTFC participants had a much lower percentage of one or more official criminal referrals for violent offenses (21% of the MTFC group vs. 38% of the control group). MTFC participants had a much lower percentage with two or more official criminal referrals for violent offenses – 5% vs. 24%. The percentage of youth with self-reported violent offenses declined by 62% in the MTFC group over the two years compared to 28% for the control group.

**Program Contact Information**
TFC Consultants, Inc.
Gerard Bouwman, President
1163 Olive St.
Eugene, Oregon 97401
Phone: 541-343-2388 ext. 204
Cell phone: 541-954-7431
Fax: 541-343-2764
The Journey to Evidence-Based Programming

Teen Outreach Program (TOP)

Program Overview
The Wyman Teen Outreach Program (TOP) is a life skills curriculum for 12- to 17-year-olds that aims to prevent negative youth behaviors, particularly school failure and early pregnancy. It is based on the premise that a heightened awareness of life options and an increased understanding of how to effectively pursue these options will lead to a variety of positive outcomes. These outcomes include success in school, as measured by reduced course failure and suspension rates, and a decrease in pregnancy rates.

TOP includes three essential program components: classroom/group instruction, community service, and service learning. Urban high school students are the primary recipients, although the program has been administered to middle school students as well. It is currently being used in 16 states, the Virgin Islands, and the United Kingdom.

Classroom instruction can take place during school time for credit or not, or may be offered as an afterschool program. The curriculum is flexible and can be altered to fit a community's particular needs. Trained facilitators deliver the curriculum in weekly classes throughout the school year. Fundamental elements of the classroom instruction include learning life skills, understanding social and emotional issues important to teens, and discussing feelings and attitudes about a variety of subjects. The topics are part of the Changing Scenes curriculum, originally created in 1988 and revised in 1996. While pregnancy prevention is a goal of TOP, sex education comprises roughly 15% of the curriculum, and may be skipped if it overlaps with other school programs.

The community service component involves some volunteer experience in the community, with a goal of providing participants with the opportunity to develop feelings of competence and autonomy. These activities may include volunteering at food kitchens, graffiti removal, fund raising, petition drives, or other community improvement activities. This experience is a minimum of 20 hours per school year, but may vary from program to program. Finally, the service learning component helps students process their community service and tie it in to the classroom learning with the goal of maximizing learning from the service experience.

Evaluation Methods and Program Outcomes
Measures of outcomes in several program evaluations demonstrate positive results from the curriculum. Wyman (2009) maintains that studies demonstrate a 60% lower rate of dropping out, a 33% lower rate of pregnancy, a 14% lower rate of school suspension, and an 11% lower rate of school course failure.
In a 1997 study by Allen and colleagues, TOP participants’ school suspension rate decreased by 24% over the course of the study while the control group experienced a 21% increase in suspension rate; TOP group’s course failure rate decreased by 12% after the study while the control group experienced a 24% increase in failure rate; TOP group’s pregnancy rate among female participants was significantly affected, with the TOP pregnancy rate decreasing 31% and the control group’s pregnancy rate decreasing only 2%. There was no significant relationship between program dosage and pregnancy rates or program dosage and suspension (Allen, Philliber, Herrling & Kuperminc, 1997).

Program Contact Information

Wyman
600 Kiwanis Drive
Eureka, Missouri 63025
Phone: 636-938-5245
Fax: 636-938-5289
The following table provides a brief snapshot of the reviewed evidence-based programs, which have been rigorously researched, demonstrate sizable effects, have been replicated, and/or show sustainable outcomes according to CSC rating and other national organizations. For additional information on the individual rating systems and criteria used by organizations found below, see Appendix D.

### Ratings of Top-Rated programs

<table>
<thead>
<tr>
<th>Program</th>
<th>Blueprints for Violence Prevention</th>
<th>Center for Mental Health Services</th>
<th>Promising Practices</th>
<th>OJJDP Model Programs Guide</th>
<th>Strengthening America's Families</th>
<th>Office of the Surgeon General</th>
<th>Coalition for Evidence-based Policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multi-systemic Therapy</td>
<td>Model</td>
<td>Proven</td>
<td>Exemplary</td>
<td>Exemplary 1</td>
<td>Model 1</td>
<td>Effective</td>
<td></td>
</tr>
<tr>
<td>LifeSkills Training</td>
<td>Model</td>
<td>Proven</td>
<td>Exemplary</td>
<td></td>
<td>Model 2</td>
<td>Effective</td>
<td></td>
</tr>
<tr>
<td>Olweus Bullying Prevention Program</td>
<td>Model</td>
<td></td>
<td>Effective</td>
<td></td>
<td>Promising 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nurse-Family Partnership</td>
<td>Model</td>
<td>Proven</td>
<td>Exemplary</td>
<td>Exemplary 2</td>
<td>Model 1</td>
<td>Effective</td>
<td></td>
</tr>
<tr>
<td>Triple P</td>
<td>Promising</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Promoting Alternative Thinking Strategies</td>
<td>Model</td>
<td>Effective</td>
<td>Exemplary</td>
<td></td>
<td>Promising 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Big Brothers Big Sisters of America</td>
<td>Model</td>
<td>Effective</td>
<td>Proven/promising</td>
<td>Exemplary</td>
<td></td>
<td>Effective</td>
<td></td>
</tr>
<tr>
<td>Check and Connect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Effective</td>
</tr>
<tr>
<td>Functional Family Therapy</td>
<td>Model</td>
<td></td>
<td>Exemplary</td>
<td>Exemplary 1</td>
<td>Model 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incredible Years</td>
<td>Model</td>
<td>Proven</td>
<td>Exemplary</td>
<td>Exemplary 1</td>
<td>Promising 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multi-dimensional Treatment Foster Care</td>
<td>Model</td>
<td></td>
<td>Exemplary</td>
<td>Exemplary 1</td>
<td>Model 1</td>
<td>Effective</td>
<td></td>
</tr>
<tr>
<td>Teen Outreach Program</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Effective</td>
</tr>
</tbody>
</table>
References


Conclusion

With shrinking budgets, accountability and return on investment have risen to the forefront of everyone's minds. Funders, grant makers, and community organizations want to ensure that the limited funding is spent on the most effective programming. For this reason, evidence-based programming has become a hot topic in the human services field. This publication aimed to make the case for investing in evidence-based programs and provide some strategies for moving forward with the process.

While we have come a long way in terms of research and evaluation and demonstrating that there are prevention and early intervention programs that are effective and show long-term outcomes, the human services field's knowledge and information gathering in this area are still in the infancy stage. There are thousands of programs being implemented across the country, the majority of which have never been rigorously evaluated. Thus, we have no idea if they work. There are simply not enough evidence-based programs to meet all of a community's needs. For this reason, new ones are developed and in many cases, the organization or individual who creates it does not think about evaluating it from the start or is not considering their program's viability as a national evidence-based model. There may in fact be many promising programs, but we will never know unless they are put to the test.

To begin changing social services across the country, we must following a strategy for improving outcomes that includes implementing evidence-based programs when they exist and continue thinking about how to move locally developed programs towards becoming evidence-based. This is no easy task. It can only be accomplished with judicial use of resources to create infrastructures for data collection, research and evaluation, as well as professional development. Providing resources for these infrastructures will afford funders the ability to hold service providers accountable - not just for process outcomes (i.e. how many people served), but did the lives of those served change for the better. If the human services field does not begin thinking in this way, we will never reach our full potential at achieving the ultimate goal: helping the children and families in local communities and across the country.
Appendix A: Glossary of Evidence-Based Terms

**Attrition** – A gradual, natural reduction in client participation or of personnel within a program.

**Client-level outcome** – The actual impact, benefit, or change for an individual or group of participants as a direct correlate or effect of the program.

**Comparison group** – A group in quasi-experimental research that is similar to the experimental group, but who do not receive the experimental intervention (e.g. treatment, therapy, or curriculum). Comparing these groups allows the researcher to identify relationships associated with the intervention.

**Control group** – A group in experimental research that is similar to the experimental groups, but who do not receive the experimental intervention (e.g. treatment, therapy, or curriculum). Comparing these groups allows the researcher to identify effect of the intervention. This group is similar to the comparison group in quasi-experimental research, but is randomly assigned (Maxfield & Babbie, 2005, p. 435).

**Cost-benefit analysis** – An assessment of whether the cost of the intervention or program is worth the benefit by measuring both in the same unit of analysis (Centre for Evidence-Based Medicine, 2004).

**Data** – Information collected in a systematic manner in order to help measure performance. This collection of observations or recorded factual material will support research and evaluation efforts.

**Effect Size** – a measure of the strength or size of the relationship between two variables.

**Essential elements** – The crucial components of an evidence-based program. These are the components that create the benefits or outcomes for participants. Other research may refer to as core components.

**Evaluation** – “The systematic collection of information about activities, characteristics, and outcomes of programs to reduce uncertainties, improve effectiveness, and make decisions with regard to what those programs are doing and effecting.” (Patton, 1982).

**Evaluation research** – An evaluation of the effects of a program in regard to its stated outcomes or goals (Maxfield & Babbie, 2005, p. 436).

**Evidence-based practice** – An approach, framework, collection of ideas or concepts, adopted principles and strategies supported by research.
Evidence-based program (EBP) – Programs comprised of a set of coordinated services/activities that demonstrate effectiveness based on research. Criteria for rating as such depend upon organization or agency doing the rankings, but typically includes researched using a rigorous design and demonstrated positive, sustainable and replicable effects. EBPs may incorporate a number of evidence-based practices in the delivery of services.

Experimental design (classical experiment) – A research design where participants are randomly assigned to either an experimental group (treatment) or a control group (no treatment/placebo). This allows the researchers to examine whether the intervention/treatment caused the outcomes or effect to take place (causal inference).

Experimental group – A group in experimental research that is similar to the control group, but who receives the experimental intervention (e.g. treatment, therapy, and curriculum). Comparing these groups allows the researcher to identify effect of the intervention (Maxfield & Babbie, 2005, p. 436).

Fidelity – Extent to which delivery of an intervention adheres to the protocol or program model originally developed (Mowbray, Holter, Teague & Bybee, 2003).

Level of significance – The degree of probability that the finding could be attributed to sampling error or that if we took another sample we might find no effect ($p \leq .05 = $ if there is 5% or less possibility that a relationship is due to chance or sampling error, we conclude the relationship is real).

Logic model – A diagram that shows the interrelationships between activities and their outcomes, using arrows to indicate which sets of activities are believed to contribute to specific outcomes.

Measurement – Assessing changes in characteristic(s) or attributes of subjects as a result of participation in a program or receipt of a treatment.

Outcome – Benefit for participants during or after their involvement with a program. Outcomes may be related to knowledge, skills, attitudes, values, behavior, condition or status. There can be “levels” of outcomes, with initial outcomes being the first change that can be expected, leading to intermediate and longer-term outcomes that can be sustained over time.

Process evaluation – An evaluation of whether a program is implemented as planned or as intended (Maxfield & Babbie, 2005, p. 438).
**Program** – A collection of services, activities, or projects intended to meet a public (or social) need and identified goals (e.g. Nurse-Family Partnership).

**Qualitative research** – Research involving detailed descriptions of characteristics, cases, and settings. This research technique derives data from observation, interviewing, and document review and focuses on the meanings and interpretations of the participants.

**Quantitative research** – Research that examines phenomenon through the numerical representation of observations and statistical analysis. The systematic scientific collection and measurement of data that are expressed as a specific unit/number that define, measure, and report on the relationships between various variables, characteristics or concepts.

**Quasi-experimental** – This research design is very similar to and almost meets criteria for an experimental design, but is unable to control all potential factors and does not include random assignment of participants.

**Replication** – Process of repeating services and/or a program model undertaken by someone else using the same methodology. Commonly the location and participants will be different. Replication results either support earlier findings or question the accuracy of earlier results.

**Target population** – The sample of participants that a program is designed to help.

**Theoretical framework** – A theoretical framework is a collection of interrelated concepts that guide one's research, determining what things a person will measure, and what statistical relationships will be identified.

**Theory of Change** – Guided by the theoretical framework, a detailed narrative that describes a process of planned social change from the assumptions that guide its design to the long-term goals it seeks to achieve.

**Variable** – A variable is anything that takes on different values. It is a measurable factor, characteristic, or attribute that varies over time.

- **Independent variable** – A variable which is actively controlled/manipulated to see if there is a change in the dependent variable and used to measure the causal construct.
- **Dependent variable** – A variable used to assess the affected construct. Rather, the dependent variable is the value that changes as a result of the manipulation of the independent variable.
References


## Appendix B: Example of a Pathways Document

### Palm Beach County’s Pathway to Early Childhood Development

<table>
<thead>
<tr>
<th>OUTCOMES</th>
<th>MEASURABLE CONDITIONS</th>
<th>INTERVENTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>#1 HEALTHY BIRTHS</strong>&lt;br&gt;measured by reduced infant mortality fewer low-birthweight births</td>
<td>• mothers receiving early and regular prenatal care&lt;br&gt;• mothers receiving appropriate preconception/interconception health care&lt;br&gt;• births to teens/ subsequent teen births&lt;br&gt;• women depressed</td>
<td>High-quality, affordable, accessible parental care. Access to specific services and supports appropriate to mothers needs. High-quality, affordable, accessible family planning services and other preconception/interconception health care. Alternatives for teens to risky sexual behaviors. Screening and follow up to assure early detection and intervention of developmental obstacles, signs of toxic childhood stress and abuse. Guidance for coping with stress, social emotional developmental, family support services, and respite care for children. Child protection agencies work in partnership with families and community organizations. Coordination to link families with services to respond to problems of substance abuse, domestic violence, child abuse or unmeet basic needs. High-quality, affordable, accessible child health care.; medical home; childhood nutrition. Screening and follow-up to assure early detection of developmental or behavioral obstacles. High-quality, affordable, accessible early care &amp; education. Parenting supports for families, including safety and nutrition. Individuals, programs, and systems where families are seen routinely provide links to specialized help and supports. Connections among schools, early childhood providers, and families focus on the transition to school and the alignment of curriculum and expectations. Schools and school districts support excellent teaching and learning in the early grades and assure most effective staff are placed in high-risk communities. High quality activities during out-of-school time compliment curricula and promote student success. Schools, community organizations, providers of services, and parents work to establish and maintain trusting relationships.</td>
</tr>
<tr>
<td><strong>#2 FEWER CHILDREN BIRTH TO FIVE ABUSED, NEGLECTED</strong>&lt;br&gt;measured by index based on number of reports of abuse and neglect (new, substantiated, or repeat)</td>
<td>• early care and education programs, sources of health care, other entry points with capacity to connect high-risk families to needed services and supports&lt;br&gt;• children identified as needing Early Intervention, social emotional wellness and disability services and are receiving these services&lt;br&gt;• parents knowledgeable about child development</td>
<td>Screening and follow up to assure early detection and intervention of developmental obstacles, signs of toxic childhood stress and abuse. Guidance for coping with stress, social emotional developmental, family support services, and respite care for children. Child protection agencies work in partnership with families and community organizations. Coordination to link families with services to respond to problems of substance abuse, domestic violence, child abuse or unmeet basic needs. High-quality, affordable, accessible child health care.; medical home; childhood nutrition. Screening and follow-up to assure early detection of developmental or behavioral obstacles. High-quality, affordable, accessible early care &amp; education. Parenting supports for families, including safety and nutrition. Individuals, programs, and systems where families are seen routinely provide links to specialized help and supports. Connections among schools, early childhood providers, and families focus on the transition to school and the alignment of curriculum and expectations. Schools and school districts support excellent teaching and learning in the early grades and assure most effective staff are placed in high-risk communities. High quality activities during out-of-school time compliment curricula and promote student success. Schools, community organizations, providers of services, and parents work to establish and maintain trusting relationships.</td>
</tr>
<tr>
<td><strong>#3 EAGER AND READY TO LEARN</strong>&lt;br&gt;(Aspirational)&lt;br&gt;measured by more children with higher scores on FLKRS</td>
<td>• children in high-quality early care and education programs&lt;br&gt;• parents reading to children&lt;br&gt;• children interact positively with peers and adults&lt;br&gt;• children with health insurance&lt;br&gt;• children with “medical home”&lt;br&gt;• children in good health&lt;br&gt;• children without undetected developmental/behavioral problems&lt;br&gt;• parental and provider depression</td>
<td>High-quality, affordable, accessible early care &amp; education. Parenting supports for families, including safety and nutrition. Individuals, programs, and systems where families are seen routinely provide links to specialized help and supports. Connections among schools, early childhood providers, and families focus on the transition to school and the alignment of curriculum and expectations. Schools and school districts support excellent teaching and learning in the early grades and assure most effective staff are placed in high-risk communities. High quality activities during out-of-school time compliment curricula and promote student success. Schools, community organizations, providers of services, and parents work to establish and maintain trusting relationships.</td>
</tr>
<tr>
<td><strong>#4 THIRD GRADE SCHOOL SUCCESS</strong>&lt;br&gt;(Aspirational)&lt;br&gt;measured by more children scoring on grade level in 3rd Grade Reading (FCAT)</td>
<td>• children in schools that connect with early care and education programs and families&lt;br&gt;• children in classes with highly skilled teachers with high expectations&lt;br&gt;• children in schools that address health, family, developmental, and attendance issues&lt;br&gt;• children in high-quality out-of-school time activities</td>
<td>High-quality, affordable, accessible child health care.; medical home; childhood nutrition. Screening and follow-up to assure early detection of developmental or behavioral obstacles. High-quality, affordable, accessible early care &amp; education. Parenting supports for families, including safety and nutrition. Individuals, programs, and systems where families are seen routinely provide links to specialized help and supports. Connections among schools, early childhood providers, and families focus on the transition to school and the alignment of curriculum and expectations. Schools and school districts support excellent teaching and learning in the early grades and assure most effective staff are placed in high-risk communities. High quality activities during out-of-school time compliment curricula and promote student success. Schools, community organizations, providers of services, and parents work to establish and maintain trusting relationships.</td>
</tr>
</tbody>
</table>
The Journey to Evidence-Based Programming
Appendix C: Example of a Program Logic Model

Good Behavior Game – Sheppard Kellam

Theory | Strategies/Activities | Immediate Outcomes | Intermediate Outcomes | Long Term Outcomes
--- | --- | --- | --- | ---
Classroom-based 1st and 2nd grade classrooms | A structured and organized classroom environment reduces aggressive behavior in young children and positively impacts their developmental path into adulthood | **Classroom Management** (Good Behavior Game)
Team-based classroom behavior management strategy based on group reinforcement and mutual self-interest for an academic year
- Class broken into teams, each with equal number of aggressive/disruptive students
- Behavioral expectations clearly defined
- Teams receive check marks for each violation of the behavioral expectations
- Winning team(s) receive tangible rewards in the beginning, with rewards becoming less tangible and less predictable as year progresses | Teachers learn an effective strategy to manage aggressive behavior exhibited in the classroom
Children learn and understand behavioral expectations | For aggressive students: Decreased frequency of aggressive/disruptive behavior during the Good Behavior Game, progressively generalizing across multiple settings
For shy/withdrawn students: Decrease in shy behavior during Good Behavior Game, progressively generalizing across multiple settings | Childhood
Decreased incidence of aggressive behaviors and incidence of ATOD use
Adolescence
Reduction in antisocial behavior and drug abuse
Adulthood
Reduction in antisocial behavior and drug abuse

\[14\text{John Jay College of Criminal Justice Research & Evaluation Center}\]
## Appendix D: National Rating Systems

<table>
<thead>
<tr>
<th>ORGANIZATION</th>
<th>FOCUS</th>
<th>RATINGS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blueprints for Violence Prevention</td>
<td>Violence Prevention</td>
<td>• Model</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Promising</td>
</tr>
<tr>
<td><a href="http://www.colorado.edu/cspv/blueprints">www.colorado.edu/cspv/blueprints</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prevention Research Center for the Promotion of Human Development</td>
<td>Reducing the risk/effects of Psychopathology</td>
<td>Effective Promising</td>
</tr>
<tr>
<td>- Penn State University (2000 Report)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><a href="http://prevention.psu.edu/pubs/documents/MentalDisorders">http://prevention.psu.edu/pubs/documents/MentalDisorders</a> fullreport.pdf</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Department of Education</td>
<td>Reducing substance use, violence, and other conduct problems</td>
<td>• Exemplary</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Promising</td>
</tr>
<tr>
<td><a href="http://www.ed.gov">www.ed.gov</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td><a href="http://www.ed.gov/admins/lead/safety/exemplary01/exemplary01.pdf">http://www.ed.gov/admins/lead/safety/exemplary01/exemplary01.pdf</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mihalic and Aultman-Bettridge (2004)</td>
<td>Reducing school disciplinary problems, suspensions, truancy, dropout, and improving academic achievement</td>
<td>• Exemplary/Model</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Favorable</td>
</tr>
</tbody>
</table>
### RATINGS DEFINED

<table>
<thead>
<tr>
<th>Model:</th>
<th>Programs that show evidence of a deterrent effect using either an experimental or quasi-experimental design, show sustained effects for at least one year post-treatment, and include replication at more than one site with demonstrated effects.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Promising:</td>
<td>Programs that show a deterrent effect using either an experimental or quasi-experimental design.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Effective:</th>
<th>Programs that are evaluated using comparison groups with either a randomized or quasi-experimental design using a control group, must have pre- and post-test data and preferably follow-up data, a written implementation manual, and must demonstrate positive outcomes.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Promising:</td>
<td>Programs that appear promising, but are not proven, meaning they lack a controlled design, contain very small samples, or have findings that are indirectly related to mental health outcomes.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exemplary:</th>
<th>The program is based on empirical data and demonstrates evidence of effectiveness in improving student achievement.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Promising:</td>
<td>The program provided sufficient evidence to demonstrate promise for improving student achievement.</td>
</tr>
</tbody>
</table>

Programs are rated according to evidence of efficacy, quality, educational significance, and usefulness to others.

<table>
<thead>
<tr>
<th>Exemplary/Model:</th>
<th>Programs that show evidence of a deterrent effect using either an experimental or quasi-experimental design, show sustained effects for at least one year post-treatment, and include replication at more than one site with demonstrated effects. (Based on Blueprints)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Promising:</td>
<td>Programs that show a deterrent effect using either an experimental or quasi-experimental design. (Based on Blueprints)</td>
</tr>
</tbody>
</table>

<p>| Favorable: | Programs have experimental or matched control group designs, show evidence that behavioral effects are due to the intervention and not other factors, but may have weaker research designs than the standard held for Blueprints. |</p>
<table>
<thead>
<tr>
<th>ORGANIZATION</th>
<th>FOCUS</th>
<th>RATINGS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>• Promising</td>
</tr>
<tr>
<td>Full report to Congress (1997) <a href="http://www.ncjrs.org/works/wholedoc.htm">www.ncjrs.org/works/wholedoc.htm</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>The Gilford Center</strong></td>
<td>Adult Services; Substance Abuse Prevention and Treatment</td>
<td>• Best Practice</td>
</tr>
<tr>
<td><a href="http://www.guilfordcenter.com/provider/practices/default.htm">http://www.guilfordcenter.com/provider/practices/default.htm</a></td>
<td>Child Services, Mental Health and Systems of Care; Developmental</td>
<td>• Emerging Best Practice</td>
</tr>
<tr>
<td></td>
<td>disabilities</td>
<td>• Evidence-based</td>
</tr>
<tr>
<td><strong>Promising Practices Network</strong></td>
<td>Children and Families</td>
<td>• Proven</td>
</tr>
<tr>
<td><a href="http://www.promising">http://www.promising</a> practices.net/programs.asp</td>
<td></td>
<td>• Promising</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Proven/Promising</td>
</tr>
</tbody>
</table>
### RATINGS DEFINED

<table>
<thead>
<tr>
<th><strong>Working/Effective:</strong></th>
<th>Programs that have at least two level 3 evaluations with statistical significance tests and the preponderance of all available evidence showing effectiveness of crime prevention or in reducing risk factors for crime, and findings can be generalizable.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Promising:</strong></td>
<td>Programs that have at least one level 3 evaluation and the preponderance of the remaining evidence showing effectiveness, but have a low level of certainty to support generalizability.Mui Programs are rated according to research design and internal validity using the Maryland Scale of Scientific Methods. Level 3 - A comparison between two or more comparable units of analysis, one with and one without the program (Research Design), causal direction and history are not threats to validity (internal validity).</td>
</tr>
<tr>
<td><strong>Best Practice:</strong></td>
<td>Generally accepted as a successful intervention currently believed to improve consumer outcomes. Evidence based practices are a type of best practice that has been established and supported by scientific evidence. The terms “best practice” and “evidence-based practice” are often used interchangeably.</td>
</tr>
<tr>
<td><strong>Emerging best practice:</strong></td>
<td>Interventions or services that have shown benefit to consumers, but have not yet been established as evidence-based practices through rigorous scientific research.</td>
</tr>
<tr>
<td><strong>Evidence-based:</strong></td>
<td>Practice is an intervention for which there is consistent scientific evidence showing that it improves client outcomes.</td>
</tr>
<tr>
<td><strong>Proven:</strong></td>
<td>Programs have at least one credible, scientifically rigorous study that demonstrates improvement on at least one indicator. To be rated as proven, all of the following must be met: (1) must improve an indicator related to children and family outcomes; (2) at least one outcome is changed by 20%, 0.25 standard deviations, or more; (3) at least one outcome with a substantial effect size is statistically significant at the 5% level; (4) study design uses a convincing comparison group to identify program impacts, including randomized controlled trial (experimental design) or some quasi-experimental designs; (5) sample size of evaluation exceeds 30 in both the treatment and comparison groups; (6) program evaluation documentation is publicly available.</td>
</tr>
<tr>
<td><strong>Promising:</strong></td>
<td>Programs have at least some evidence that the program improves outcomes for children and families. To be rated as promising all of the following must be met: (1) may affect intermediary variables rather than direct outcomes; (2) change in outcome is more than 1%; (3) outcome change is significant at the 10% level (marginally significant); (4) study has a comparison group, but it may exhibit some weaknesses, e.g., the groups lack comparability on pre-existing variables or the analysis does not employ appropriate statistical controls; (5) sample size of evaluation exceeds 10 in both the treatment and comparison groups; (6) program evaluation documentation is publicly available.</td>
</tr>
<tr>
<td><strong>Proven/Promising:</strong></td>
<td>Program affects more than one indicator, and the level of evidence differs across indicators. Additional considerations play a role on a case-by-case basis. These may include attrition, quality of outcome measures, and others.</td>
</tr>
<tr>
<td>ORGANIZATION</td>
<td>FOCUS</td>
</tr>
<tr>
<td>--------------</td>
<td>-------</td>
</tr>
</tbody>
</table>
| **Juvenile Justice Evaluation Center** | Youth Violence | • Model programs  
• Promising approaches  
• Innovative approaches |
| **Substance Abuse and Mental Health Services Administration Center for Substance Abuse Prevention** | Substance abuse prevention Criteria applied to High Risk Youth programs (HRY) and Pregnant and Postpartum Women and their Infants programs (PPWI) | • Type 1  
• Type 2  
• Type 3  
• Type 4  
• Type 5 |
RATINGS DEFINED

**Model programs**: Model programs are those that have demonstrated definitive success in multiple evaluations. These are sometimes referred to as exemplary programs.

**Promising approaches**: Those for which evaluation evidence is suggestive of success, but not definitive.

**Innovative approaches**: Those for which no evidence exists, but may be based on prior research or evaluation.

<table>
<thead>
<tr>
<th>Type 1:</th>
<th>Not scientifically defensible. The program/principle has been defined or recognized publicly, and has received awards, honors, or mentions.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 2:</td>
<td>Not scientifically defensible. The program/principle has appeared in a non-refereed professional publication or journal. It is important to distinguish between citations found in professional publications and those found in journals.</td>
</tr>
<tr>
<td>Type 3:</td>
<td>Expert/peer consensus process - scientifically defensible. The program's source documents have undergone thorough scrutiny in an expert/peer consensus process for the quality of implementation and evaluation methods, or a paper has appeared in a peer-reviewed journal. All dosage information and data collection processes are detailed, all analysis are presented for review. Reviewers trained as evaluators, code the implementation variables and activities, as well as the findings.</td>
</tr>
<tr>
<td>Type 4:</td>
<td>Qualitative or quantitative meta-analysis - scientifically defensible. The program/principles have undergone either a quantitative meta-analysis or and expert/peer consensus process in the form of a qualitative meta-analysis.</td>
</tr>
<tr>
<td>Type 5:</td>
<td>Replications of programs/principles - scientifically defensible. Replications of program/principle have appeared in several refereed professional journals. Evidence of a program's effectiveness is that it can be replicated across venues and populations, demonstrating credibility, utility, and generalizability.</td>
</tr>
</tbody>
</table>

⚠️ Matrix applied to establish Scientific Credibility of a program with overall program ratings on a scale of 1-5 by Integrity and Utility.

⚠️ Must score 3 or greater
<table>
<thead>
<tr>
<th>ORGANIZATION</th>
<th>FOCUS</th>
<th>RATINGS</th>
</tr>
</thead>
</table>
| What Works Clearinghouse (WWC) | Educational interventions (programs, products, practice, and policies) | • Meets Evidence Standards  
• Meets Evidence Standards with reservations  
• Does not meet Evidence Screens                                         |
| www.whatworks.ed.gov/        |                                                                      |                                                                        |
| Helping America's Youth      | Prevent and reduce delinquency or other youthful (up to age 20) problem behaviors (e.g. drug and alcohol use). | • Level 1  
• Level 2  
• Level 3                                                             |
<p>| <a href="http://guide.helpingamericas">http://guide.helpingamericas</a> youth.gov/default.htm |                                                                      |                                                                        |</p>
<table>
<thead>
<tr>
<th>RATINGS DEFINED</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Meets Evidence Standards:</strong> Randomized controlled trials (RCTs) that do not have problems with randomization, attrition, or disruption, and regression discontinuity designs that do not have problems with attrition or disruption.</td>
</tr>
<tr>
<td><strong>Meets Evidence Standards with Reservations:</strong> Strong quasi-experimental studies that have comparison groups and meet other WWC Evidence Standards, as well as randomized trials with randomization, attrition, or disruption problems and regression discontinuity designs with attrition or disruption problems.</td>
</tr>
<tr>
<td><strong>Does not meet Evidence Screens:</strong> Studies that provide insufficient evidence or causal validity or are not relevant to the topic being reviewed</td>
</tr>
</tbody>
</table>

In addition, the standards rate other important characteristics of study design, such as intervention fidelity, outcome measures, and generalizability.

<p>| Level 1: Programs have been scientifically demonstrated to prevent delinquency or reduce/enhance risk/protective factors for delinquency and other child and youthful problems using a research design of the highest quality (i.e. an experimental design and random assignment of subjects). |
| Level 2: Programs have been scientifically demonstrated to prevent delinquency or reduce/enhance risk/protection for delinquency and other child and youthful problems using either an experimental or quasi-experimental research design, with a comparison group, and the evidence suggest program effectiveness, but the evidence is not as strong as the Level 1 programs. |
| Level 3: Programs display a strong theoretical base and have been demonstrated to prevent delinquency and other child and youthful problems or reduce/enhance risk/protective factors for them using limited research methods (with at least single group pre- and post –treatment measurements). The evidence associated with these programs appears promising but requires confirmation using more rigorous scientific techniques. |</p>
<table>
<thead>
<tr>
<th>ORGANIZATION</th>
<th>FOCUS</th>
<th>RATINGS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communities That Care, Developmental Research and Programs</td>
<td>Substance abuse, delinquency, teen pregnancy, school dropout, violence, and child and youth development</td>
<td>• Effective</td>
</tr>
<tr>
<td>Office of the Surgeon General</td>
<td>Youth Violence</td>
<td>• Model</td>
</tr>
<tr>
<td><a href="http://www.surgeon">http://www.surgeon</a> general.gov/library/youthviolence/chapter5/sec2.html#Scientific Standards</td>
<td></td>
<td>• Promising</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Does Not Work</td>
</tr>
<tr>
<td>OJJDP Model Programs Guide</td>
<td>Entire continuum of youth services from prevention through sanctions to reentry.</td>
<td>• Exemplary</td>
</tr>
<tr>
<td><a href="http://www.dsgonline.com/mpg2.5/mpg_index.htm">http://www.dsgonline.com/mpg2.5/mpg_index.htm</a></td>
<td></td>
<td>• Effective</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Promising</td>
</tr>
</tbody>
</table>
### RATINGS DEFINED

**Effective:** (1) Programs address research based risk factors for substance abuse, delinquency, teen pregnancy, school dropout and violence (2) Increase protective factors; (3) intervene at developmentally appropriate age; and (4) show significant effects on risk and protective factors in controlled studies or community trials.

**Model:** Rigorous experimental design (experimental or quasi-experimental); Significant deterrent effects on: Violence or serious delinquency (Level 1) or any risk factor for violence with a large effect (.30 or greater) (Level 2); Replication with demonstrated effects; and Sustainability of effects.

**Promising:** Rigorous experimental design (experimental or quasi-experimental); Significant deterrent effects on: Violence or serious delinquency (Level 1) or any risk factor for violence with an effect size of .10 or greater (Level 2); Either replication or sustainability of effects.

**Does Not Work:** Rigorous experimental design (experimental or quasi-experimental); Significant evidence of null or negative effects on violence or known risk factors for violence; Replication, with the preponderance of evidence suggesting that the program is ineffective or harmful.

**Exemplary:** In general, when implemented with a high degree of fidelity these programs demonstrate robust empirical findings using a reputable conceptual framework and an evaluation design of the highest quality (experimental).

**Effective:** In general, when implemented with sufficient fidelity these programs demonstrate adequate empirical findings using a sound conceptual framework and an evaluation design of the high quality (quasi-experimental).

**Promising:** In general, when implemented with minimal fidelity these programs demonstrate promising (perhaps inconsistent) empirical findings using a reasonable conceptual framework and a limited evaluation design (single group pre- post-test) that requires causal confirmation using more appropriate experimental techniques.

The Model Programs Guide (MPG) evidence ratings are based on the evaluation literature of specific prevention and intervention programs. The overall rating is derived from four summary dimensions of program effectiveness: (1) the conceptual framework of the program; (2) the program fidelity; (3) the evaluation design; and (4) the empirical evidence demonstrating the prevention or reduction of problem behavior; the reduction of risk factors related to problem behavior; or the enhancement of protective factors related to problem behavior.
<table>
<thead>
<tr>
<th>ORGANIZATION</th>
<th>FOCUS</th>
<th>RATINGS</th>
</tr>
</thead>
</table>
| The National Centers for Career and Technical Education | Secondary & postsecondary career and technical education | • Exemplary  
• Promising |

This program is a joint effort of The Ohio State University (OSU) and the University of Illinois at Urbana-Champaign (UIUC). *Funding for Rating process ended in 2002
RATINGS DEFINED

Exemplary or Promising:

A. Program Quality
1. Learning goals and objectives are clear, challenging, and appropriate for the intended population. 2. The program content is aligned with learning goals, is accurate, current, and appropriate for the intended learner population. 3. The program reflects the vision promoted in recognized academic and occupational standards and by state and national legislation, as appropriate. 4. Collaborations with internal and external organizations and stakeholders are maintained to strengthen the quality and effectiveness of the program.

B. Educational Significance
1. The program addresses important individual and societal needs. 2. The program contributes to educational excellence for all learners and leads to other positive results or outcomes. 3. The program design is innovative, reflects current research, and is worthy of replication.

C. Evidence of Effectiveness and Success
1. The program makes a measurable difference in learning for all participants. 2. The program meets or exceeds identified performance goals. 3. A systematic evaluation process is used to continuously improve the program.

D. Replicability/Usefulness to Others
1. The whole program or elements of the program can be successfully implemented, adopted, or adapted in other educational settings.

Practices have been rated in 15 different categories (all related to technical and career education): (1) Access and inclusiveness; (2) Alignment with standards; (3) Certification and credentialing; (4) Curriculum reform; (5) Evaluation and continuous improvement; (6) Placement and retention; (7) Partnerships; (8) Professional development; (9) Program and instructional delivery; (10) Program and institutional leadership; (11) Technology enhancements; (12) Transitions options; (13) Student development and leadership; (14) Sustainability and finances; (15) Systemic and whole school reform.
<table>
<thead>
<tr>
<th>ORGANIZATION</th>
<th>FOCUS</th>
<th>RATINGS</th>
</tr>
</thead>
</table>
| Exemplary and Promising Safe, Disciplined and Drug-Free Schools Programs 2001 | Safe, Disciplined and Drug Free Schools | • Exemplary  
• Promising |


| Federal Government/Office of Management and Budget | Dept. of Energy to Homeland Security to the Interior, etc.  
Health and Human Services - There was a range of topics including, but not limited to: health, childcare, adoption, family planning, developmental disabilities, maternal child health, substance abuse, mental illness, homelessness, universal newborn screenings, TANF and immigration | • Effective  
• Moderately Effective  
• Adequate |

http://www.whitehouse.gov/omb/expectmore/perform.html
### RATINGS DEFINED

<table>
<thead>
<tr>
<th>Exemplary:</th>
<th>Based on empirical data a program was effective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Promising:</td>
<td>There is sufficient evidence to demonstrate that the program showed promise for improving student achievement</td>
</tr>
</tbody>
</table>

Ratings use the following criteria:

**A. Evidence of Efficacy**
- **Criterion 1**: The program reports relevant evidence of efficacy/effectiveness based on a methodologically sound evaluation.

**B. Quality of Program**
- **Criterion 2 (Goals)**: The program’s goals with respect to changing behavior and/or risk and protective factors are clear and appropriate for the intended population and setting.
- **Criterion 3 (Rationale)**: The rationale underlying the program is clearly stated, and the program’s content and processes are aligned with its goals.
- **Criterion 4 (Content Appropriateness)**: The program’s content takes into consideration the characteristics of the intended population and setting (e.g., developmental stage, motivational status, language, disabilities, culture) and the needs implied by these characteristics.
- **Criterion 5 (Implementation Methods)**: The program implementation process effectively engages the intended population.

**C. Educational Significance**
- **Criterion 6**: The application describes how the program is integrated into schools’ educational missions.

**D. Usefulness to Others**
- **Criterion 7 (Ability to be replicated)**: The program provides necessary information and guidance for replication in other appropriate settings.

<table>
<thead>
<tr>
<th>Effective:</th>
<th>This is the highest rating a program can achieve. Programs rated Effective set ambitious goals, achieve results, are well-managed and improve efficiency.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moderately Effective:</td>
<td>In general, a program rated Moderately Effective has set ambitious goals and is well-managed. Moderately Effective programs likely need to improve their efficiency or address other problems in the programs’ design or management in order to achieve better results.</td>
</tr>
<tr>
<td>Adequate:</td>
<td>This rating describes a program that needs to set more ambitious goals, achieve better results, improve accountability or strengthen its management practices.</td>
</tr>
<tr>
<td>ORGANIZATION</td>
<td>FOCUS</td>
</tr>
<tr>
<td>------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Strengthening America's Families</td>
<td>Behavioral Parent and Family Skills Training or Behavioral Family therapy, Family therapy, Family In-home Support, Comprehensive Approaches, incorporates universal, selected (at risk) and indicated (crisis) prevention efforts</td>
</tr>
<tr>
<td>Ohio State CLEX</td>
<td>Youth Behavior Mental Health Alt. Educ</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### RATINGS DEFINED

**Exemplary:** Programs that are well-implemented, are rigorously evaluated, and have consistent positive findings (integrity ratings of “A4” or “A5”).

**Model:** Programs that have consistent integrity ratings of “A3” and “A4”

**Promising:** Programs that have mixed integrity ratings but demonstrate high integrity ratings in at least 3 - 4 categories.

- Programs are rated across 14 dimensions receiving rating from A1 for “very low quality,” to A5 for “very high quality.” Dimensions include: (1) Theory; (2) Fidelity of Interventions; (3) Sampling Strategy & Implementation; (4) Attrition; (5) Measures; (6) Missing Data; (7) Data Collection; (8) Analysis; (9) Other plausible threats to validity; (10) Replications; (11) Dissemination Capability; (12) Cultural & Age Appropriateness; (13) Integrity; and (14) Utility

<table>
<thead>
<tr>
<th><strong>Evidence checklist:</strong></th>
<th>Implementable, based on effective principles, customer satisfaction, change reports, comparison group, random assignment to control group, longitudinal impact, multiple site replication, dosage analysis, meta-analysis, expert review &amp; consensus.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-2 checks: <strong>Unproven approach:</strong></td>
<td>no documentation approach has either ever been used or has been implemented successfully w/no evaluation.</td>
</tr>
<tr>
<td>3-5 checks: <strong>Promising Approach:</strong></td>
<td>implemented and significant impact evaluations have been conducted. Data is promising; its scientific rigor is insufficient to suggest causality. Multiple factors contribute to the success of participants.</td>
</tr>
<tr>
<td>6-10 points: <strong>Evidence Based:</strong></td>
<td>compelling evidence of effectiveness. Attribute participant success to the program itself, and have evidence that the approach will work for others in different environments.</td>
</tr>
</tbody>
</table>

**Model:** meets the satisfactory standards of specific criteria as an effective program.
<table>
<thead>
<tr>
<th>ORGANIZATION</th>
<th>FOCUS</th>
<th>RATINGS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Child Welfare League of America</strong></td>
<td>Child Welfare</td>
<td>• Exemplary Practice</td>
</tr>
<tr>
<td><a href="http://www.cwla.org/programs/r2p/levels.htm">http://www.cwla.org/programs/r2p/levels.htm</a></td>
<td></td>
<td>• Commendable Practice</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Emerging Practice</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Innovative Practice</td>
</tr>
<tr>
<td><strong>Child Trends</strong></td>
<td>Life Course Models, teen programs, school readiness, and afterschool</td>
<td>• What works</td>
</tr>
<tr>
<td><a href="http://www.childtrends.org/what_works/clarkwww/clarkwww_intro.asp">http://www.childtrends.org/what_works/clarkwww/clarkwww_intro.asp</a></td>
<td></td>
<td>• What doesn’t work</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Mixed Reviews</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Best Bets</td>
</tr>
</tbody>
</table>
### RATINGS DEFINED

**Levels of Research Rigor** - Each program or practice included in the Research to Practice (R2P) initiative has been identified as effective with successes supported by a research component. R2P has developed the following categories to describe the level of empirical support available. All programs and practices exist within an organizational context with many factors that may influence outcomes.

*Exemplary Practice:* Must have: Randomized study, Control group, Posttests or pre- and posttest, Effects sustained for at least one year, Multiple replications.

*Commendable Practice:* Must have a majority of the following characteristics: Randomized or quasi-experimental study, Control or comparison group, Posttests or pre- and posttests, Follow up, Replication.

*Emerging Practice:* Must have a majority of the following characteristics: quasi-experimental study, correlational or ex post facto study, Posttest only, Single group pre- and posttest, Comparison group.

*Innovative Practice:* Must have a majority of the following characteristics: Case study, Descriptive statistics only, Treatment group only.

**What works** – Programs with specific evidence from experimental studies that show a significant positive impact on a particular developmental outcome.

**What doesn't work** – Programs with experimental evidence that, to date, an outcome has not been positively affected by a particular program. These findings should not be construed to mean that the program can never positively affect outcomes or that it cannot be modified to affect outcomes positively.

**Mixed reviews** – Programs with experimental evidence that a program has been shown to be effective in some, but not all, studies or that it has been found to be effective for some, but not all, groups of young people.

**Best bets** – Programs with promising approaches or practices that have not been tested through experimental research but that may be important from a theoretical standpoint. These include results from quasi-experimental studies, multivariate analyses, analyses of longitudinal and survey studies, non-experimental analyses of experimental data, and wisdom from practitioners working in the field. The term “best bets” is not intended to highlight these as the recommended practices for programs, but as promising approaches worthy of consideration by program designers or policymakers.
<table>
<thead>
<tr>
<th>ORGANIZATION</th>
<th>FOCUS</th>
<th>RATINGS</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Community Corrections Institute</td>
<td>Substance Abuse Offenders</td>
<td>• No specifics</td>
</tr>
<tr>
<td><a href="http://www.accilifeskills.com/evidence-based/research.htm">http://www.accilifeskills.com/evidence-based/research.htm</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Substance Abuse &amp; Mental Health Services Administration - National Registry of Evidence-based Programs and Practices</td>
<td>Behavioral Health – Substance Use and Mental Health</td>
<td>• No overall ratings</td>
</tr>
<tr>
<td><a href="http://www.nrepp.samhsa.gov/">http://www.nrepp.samhsa.gov/</a></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### RATINGS DEFINED

**Fifteen Point Rating Criteria For Evidence-Based Programs**

1. **Theory**: the degree at which programs reflect clear principles about substance abuse behavior and how it can be changed.
2. **Intervention Fidelity**: how the program ensures consistent delivery.
3. **Process Evaluation**: whether the program implementation was measured
4. **Sampling Strategy and Implementation**: how well the program selected its participants and how well they received it.
5. **Attrition**: whether the program retained participants during evaluation.
6. **Outcome Measures**: the relevance and quality of evaluation measures.
7. **Missing Data**: how developer addressed incomplete measurements.
8. **Data Collection**: the manner in which data were gathered.
9. **Analysis**: the appropriateness and technical adequacy of data analyses.
10. **Other Plausible Threats to Validity**: the degree to which the evaluation considers other explanations for program effects.
11. **Replications**: number of times the program has been used in the field.
12. **Dissemination Capability**: whether program materials are ready for implementation by others in the field.
13. **Cultural Age Appropriateness**: the degree to which the program addresses different ethnic, racial and age groups.
14. **Integrity**: overall level of confidence of the scientific rigor of the evaluation.
15. **Utility**: overall pattern of program findings to form prevention theory and practice.

**Quality of Research** – (1) Reliability of measures; (2) Validity of measures; (3) Intervention fidelity; (4) Missing data and attrition; (5) Potential confounding variables; and (6) Appropriateness of analysis

**Readiness for Dissemination** - (1) Availability of implementation materials; (2) Availability of training and support resources; and (3) Availability of quality assurance procedures

Scales of 0.0 to 4.0, with 4.0 being the highest rating given for each component

---

*Note: Format and select information from Mihalic, S.F. Agency and Practitioner Rating Categories and Criteria for Evidenced Based Programs. Center for the Study and Prevention of Violence, Blueprints Initiative.*
### Appendix E: Checklist for Evidence-Based Programs

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Are outcomes aligned with CSC’s Logic Model?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Are outcome sustainable for at least one year?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Research:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Experimental Design</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Quasi Design</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Peer Review</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Has the program been replicated with more than one site?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Has the program been replicated with more than one population?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Does this program meet one of the sentinel outcomes? Please check all that apply.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Healthy Births</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Eager to Learn</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Child Abuse Prevention</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**

________________________________________________________________________________________
________________________________________________________________________________________
________________________________________________________________________________________
________________________________________________________________________________________
________________________________________________________________________________________
________________________________________________________________________________________
________________________________________________________________________________________
Appendix F: Sample Questions from Assessment Tool

Strong Theoretical Foundation

Question 1: What theory/model underlies the framework of your program (e.g. social-learning, cognitive-behavioral)?

Question 2: What research supports your program design?

Question 3: Who is your target/intended population?

Question 4: How are the program activities developmentally/age appropriate for the intended population?

Question 5: How are the program activities culturally appropriate?

Logic Model and/or Theory of Change

Question 1: Do you have a Theory of Change? If yes, is the theory of change research-based?

Question 2: Does your theory of change address the following components:
   ♦ Problem statement?
   ♦ Desired results?
   ♦ Influential factors?
   ♦ Strategies?
   ♦ Assumptions?

Question 3: Do you have a logic model for your program?

Question 4: If yes, does your logic model address the following components:
   ♦ Targeted Population?
   ♦ Activities, Dosage, Frequency?
   ♦ Initial Outcomes?
   ♦ Intermediate Outcomes?
   ♦ Long-term Outcomes?

Question 5: Does the program address risk and protective factors of target population?
   If yes, which ones?
Outcomes

Question 1: What outcomes does your program address?
   ♦ Process?
   ♦ Client Level?

Training to the Model

Question 1: Does your program require specific training to a model?

Question 2: Does your program have an accountability system to ensure fidelity to the model?

Quality Data Collection

Question 1: What is your data collection capacity?

Question 2: Do you collect client-level data, process data, or both?

Question 3: How are you measuring accuracy?

Monitor Client Attrition

Question 1: Do you monitor attrition?

Question 2: How does program staff ensure successful and positive working relationships with program participants?

Question 3: Are the participants effectively engaged?
Monitor for Fidelity

Question 1: Are program activities being implemented with fidelity as per program design/model?
   ♦ Clients are selected as per the program model?
   ♦ Clients are assessed as per the program model?
   ♦ Clients are receiving services/activities/treatment as per the program model?
   ♦ Service/Treatment dosage adheres to the program model?
   ♦ Client to staff ratio adheres to program model?

Question 2: How are you ensuring fidelity to the model?

Question 3: What data are you collecting to measure fidelity?

Process Evaluation

Question 1: Have you conducted a process evaluation?

Outcome Evaluation

Question 1: Have you conducted an outcome evaluation?

Question 2: What type of research design does the evaluation utilize?

Question 3: Which outcomes are statistically significant?

Question 4: What are the effect sizes of the outcomes?

Question 5: Have you measured sustainability? If yes, how long? And which outcomes are sustainable?

Question 6: Has the program been replicated? If yes, are the outcomes consistent across sites or across populations?