Research Review

Evidence-Based Programs and Practices: What Does It All Mean?

Children's Services Council's Mission and Vision and Evidence-Based Programs

In order for CSC to truly achieve its mission and vision, we as an organization must expect programs and services we fund to be able to demonstrate through data that they are achieving positive results and "doing no harm" to the recipients. CSC must be accountable to the children, families and taxpayers of Palm Beach County. In order to do that, the best possible programs and services must be in place. This means that we are either funding programs that are already evidence-based, are on a continuum of becoming evidence-based, or are providing services that enable children to enter programs that are evidence-based.

Children and families will be able to reach their full potential if we as an organization and our providers and partners offer the best possible programs and services. We must remember that we are only at the beginning of this journey and are all in it together.

In order to assist in this process, CSC has organized an evidence-based programs committee consisting of a cross section of divisions and outside consultants. Its primary purpose is two-fold, (1) to gather research on nationally rated, evidence-based programs and (2) to construct an assessment tool comprised of specific criteria to rate our currently funded programs. This tool will enable us to see where programs/services fall on a continuum of effectiveness so that we can better understand program needs and also assist programs in their journey towards becoming more effective. This effort will also help agencies see where they are on the continuum and help them improve their programs. More specifically, the more information CSC and providers have, the better equipped we are in regards to either implementing a nationally rated program or helping to refine current programs in order to demonstrate their effectiveness.

In January, at a Senior Executive Policy Institute, providers engaged in an activity aimed at helping the committee examine what criteria should be included in the assessment tool. There was also an inquiry regarding what CSC could do to help providers move towards becoming evidence-based. Information like this is quite helpful to committee work and provider input and participation will continue to be needed along the way.

The Science of Investing in Evidence-Based Programs: Advocacy and Impact

It should be the charge of any social service organization to not only affect the children it serves, but to improve the lives of those who it will never serve directly. This goal is often done though advocacy (Pizzigati, Stuck, and Ness, 2002). To that end, it is critical that CSC advocate for other service providers in the county to begin or continue to research and implement evidence-based programs. Why? The answer is simple, because they are a good return on investment and, more importantly, research shows they work for the children and families they serve.

Tana Ebbole
CEO

While the phrase "evidence-based programs and practices" has been a common one within the medical field, it is becoming much more widespread in other disciplines, including early childhood education, academia, and juvenile justice. However, what has become quite evident is that not all disciplines are using it in the same way; in other words, there is no consensus regarding the definition and what criterion makes a program evidence-based. This of course leads us to the problem of miscommunication.

It is critically important that each area have a common understanding of what the term means and of issues that are relevant to this work. Children's Services Council Palm Beach County believes it is imperative that we help to inform and educate the community regarding what evidence-based programs mean to our organization and to the programs and services we fund.

What is an Evidence-Based Program?

No universal definition exists for the term "evidence-based program." (See Appendix A) Evidence-based is often used synonymously with research-based and 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 and each is**

Children's Services Council
PALM BEACH COUNTY
Expect Success

Table of Contents

defined differently by the various organizations defining them. There are at least 23 organizations that have created criteria to rate program effectiveness (See Appendix B).

For example, the Substance Abuse and Mental Health Services Administration (SAMHSA) uses the term "science-based programs" and defines them as "programs which have been reviewed by experts in the field according to accepted standards of empirical research. Science-based programs are conceptually sound and internally consistent, have sound research methodology, and can prove the effects are clearly linked to the program itself and not extraneous events" (Kyler, Bumbarger, and Greenberg, 2005, p. 2). The No Child Left Behind Act uses the term "scientifically based research" program, which is defined as having "reliable evidence that the program or practice works" (U.S. Department of Education, 2003).

Each organization may have different criteria for determining whether a program is evidence-based. Reviewers for SAMHSA's National Registry of Evidence-based Programs and Practices use criteria to measure the quality of research and readiness for dissemination. They look at the following quality of research criteria: reliability, validity, fidelity, attrition and missing data, potential confounding variables, and appropriateness of analysis. The readiness for dissemination criteria is as follows: Availability of implementation materials, training and support resources, and quality assurance procedures.

"Effective Programs do not try to do everything for everyone. Their design and operation reflect clear priorities and goals in terms of the type of youth they target, what they seek to accomplish, and the kinds of services, supports, and activities they offer" (Promising & Effective Practices Network, 2001).

The Department of Education uses specific criteria noted in the No Child Left Behind Act, which includes

- "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; and
- experimental studies are presented in sufficient detail and clarity to allow for replication, or, offer the opportunity to build systematically on their findings; and 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, 2001, pp. 1964-65).

Because organizations use different criteria, it is extremely important that a funder or provider of evidence-based programs understands who is assigning a rating to the programs and how they are defining it.

Evidence-based practice stands in contrast to approaches that are based on tradition, convention, belief, or anecdotal evidence (National Registry of Evidence-based Programs and Practices).

Common Elements

Although this may be confusing, most definitions of evidence-based do include common elements such as: a strong theoretical foundation; intended for a developmentally appropriate population; quality data collection and procedures; and evidence of effectiveness. For a program to show effectiveness, generally there must be strong evidence that the program results are the direct result of the activities of the program. This means that no other factor or factors were major contributors to the outcomes or that the changes did not happen by chance. For example, while we would expect early education programs to produce favorable effects on children, a scientifically sound evaluation is absolutely required in order to know whether they fulfill their promise (Karoly, Kilburn and Cannon, 2005). To truly say that a program is effective, there must be a strong research design testing the outcomes. This means using an experimental/Randomized Control Trial (RTC) or quasi-experimental design. The experimental design is often referred to as the "gold standard" in research.

While an in-depth discussion of research designs and methodology is outside the scope of this brief, it is important to note that there are specific types of studies needed in order to say that a program is working and achieving specific child-level outcomes. Without an evaluation that compares a group that received the program or intervention with another group that did not, it would be difficult to determine whether or not the program/

intervention caused the differences between the two groups of children. Also, if you just measure children before and after they receive treatment then you can not say that the gains they made would not have occurred despite the intervention. As Karoly and colleagues note (2005), we do not want to attribute a positive effect to a program without a comparison with what would have happened in the absence of the program, holding all other factors constant. Ultimately, we want to answer the question "Compared to what?" to determine whether a program is "effective" (p. 27).

Other characteristics of a rigorous research design are an adequate sample size (meaning there were a sufficient number of research subjects who received the intervention); a measurement of sustainability; replication; and a measure of participants' gains or changes in knowledge, attitudes, and behaviors.

Evidence-Based Practice (Best Practices) Versus Program: Is There A Difference?

While many use the terms "programs" and "practices" interchangeably, more and more researchers and practitioners are beginning to differentiate between these terms. A practice is defined as a habitual or customary performance or operation action or something that a professional does in order to achieve a positive outcome. More specifically, according to Fixsen et al. (2005), evidence-based practices are skills, techniques, and strategies that can be used when a practitioner is interacting

Commonly Used Terms

Attrition

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

Experimental Design (classical experiment)

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).

Fidelity

Extent to which delivery of an intervention adheres to the protocol or program model originally developed.

Quasi-Experiment

This research design is very similar to and almost meets criteria for an experimental design, but is unable to control 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. Intervention adheres to the protocol or program model originally developed.

directly with a customer (p. 26). They are sometimes called core intervention components when used in a broader program context.

Evidence-based programs may be defined as "organized, multi-faceted interventions that are designed to serve consumers with complex problems. Such programs, for example, may seek to integrate social skills training, family counseling, and educational assistance, where needed, in a comprehensive yet individualized manner, based on a clearly articulated theory of change, identification of the active agents of change, and the specification of necessary organizational supports" (Fixsen et al., 2005, p. 82). Programs integrate various practices within specific settings and with targeted customers. At CSC, we differentiate between these two terms as follows:

- **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. EBPs may incorporate a number of evidence-based practices in the delivery of services.
- **Evidence-Based Practice** – An approach, framework, collection of ideas or concepts, adopted principles and strategies supported by research.

History of Evidence-Based Programs

The idea of evidence-based programs is quite new overall, and it is even more recent for the social service arena. The premise of evidence-based originated in the medical field. One landmark in the movement towards evidence-based programs was the establishment of the Food and Drug Administration,

which is responsible for testing the safety of medical treatments (Leff, 2002). Another landmark was in the use of randomized control studies. It was only in 1948 that the first such study took place – researching the efficacy of streptomycin in treating tuberculosis. By the 1960s the number of randomized control trials reached into the hundreds, and today there are tens of thousands occurring every day (Dodge, 2006).

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. It really 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 was to assist in disseminating research and effective programs/services regarding problem behaviors. In 1999, the American Psychological Association established a task force for the main 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 current best 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 in that they believed that "it infringes on their autonomy and dehumanizes clients" (Dodge, 2006, p. 477). There was concern that clients vary too much in regards to disorders, co-morbidity, personality, race, ethnicity, and culture to use a one size fits all "cookie-cutter" approach (Levant, 2005).

In education, the No Child Left Behind Act of 2001 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 (See page 3 for definitions and criteria). Prevention science has been the last discipline to welcome evidence-based programs. In 1996, the Center for the Study and Prevention of Violence in Colorado began examining various youth programs to determine which ones worked to reduce crime and violence. As can be seen from the previous examples, the areas of substance use, mental health, and juvenile justice have been working towards using evidence-based programs for the past 15 years, but a systematic review of programs in the primary prevention and early intervention areas, such as early care and education is just now taking off.

Legislation and Evidence-Based Programs

In March 2007, Senators Salazar (D-Colorado) and Specter (R-Pennsylvania) introduced the Healthy Children and Families Act, a bill to expand the evidence-based program Nurse-Family Partnership to all 50 states. This bill would allow states to draw down federal dollars in support of their State's Children's Health Program (S-CHIP). If passed, as many as 570,000 mothers and children could gain access to the program each year.

Nurse-Family Partnership is a home visiting program for first-time, low-income mothers. It has been researched using randomized control trials on three different occasions with the findings published in 1978, 1990, and in 1994. Outcomes achieved range from positive birth outcomes

As of 2004, Louisiana was the only state that used Medicaid to fully fund the Nurse-Family Partnership through Targeted Case Management. Since inception, premature births have decreased by 52% and low birthweight deliveries decreased by 22% for participating mothers (O'Connor, 2004, p. 7.)

to a reduction in maternal antisocial/ criminal behavior and child abuse. This program has been researched by both the Washington State Institute for Public Policy and the RAND Corporation and has shown a positive return on the dollar ($2.88 and $4 for every dollar invested, respectively). Economic benefits include a reduction in emergency room visits, school dropout, arrests and incarceration, and an increase in employment (Yeager, 2007).

Examples of Evidence-Based Programs

1. **Nurse-Family Partnership (David Olds)** – A home visiting program for first-time, low-income, at-risk mothers promoting improved maternal, prenatal, and early childhood health. The outcomes achieved include the following:

   - **Improved Birth Outcomes**: low birthweight, preterm delivery, neurodevelopmental impairment
   - **Improved Outcomes for At-risk Mothers**: reduced rates of subsequent pregnancy, reduction in maternal behavioral problems due to substance use, reduction in school dropout rates, reduction in unemployment, reduced use of welfare and food stamps, and fewer arrests
   - **Improved Child Outcomes**: reduced rates of childhood injury, abuse, and neglect. Long-term follow-up also

shows children have fewer sexual partners, reduced cigarette smoking and alcohol use, and fewer arrests and convictions 15 years later.

2. **High/Scope Perry Preschool Program (David Weikart)** - A universal preschool program that utilizes an active learning environment to encourage independence, self-esteem, confidence, problem-solving skills, social cooperation, and promotes school bonding. The outcomes achieved include the following:

   - **Improved Child Outcomes**: reduction in need for special education classes, increased academic success (high school graduation), increased adult financial stability (employment, home ownership, monthly income, lowered incidence of use of welfare and other social services), and reduction in arrests

3. **Incredible Years (Carolyn Webster-Stratton)** – A program for children ages two to eight living in poverty with conduct problems that teaches children to manage anger and frustration in a healthy manner. It provides parents with effective parenting skills to work with their child's problem behaviors, and provides teachers with appropriate classroom management skills to address and reduce problem behaviors. The outcomes achieved include the following:

   - **Improved Child Outcomes:** positive peer association and interaction, positive behavior interactions at home and school, emotional and social competence, increase in problem-solving and anger management skills, school readiness, academic success, and prevention and reduction of aggressive and problem behaviors
   - **Improved Parental And School Outcomes:** parents and teachers use appropriate and effective discipline practices and praise and increased parental involvement in school and positive relationships between parents and teachers

What Organizations Promote the Use of Evidence-Based Programs?

There are many organizations that now promote the use of evidence-based programs. It is important to remember that definitions of evidence-based programs and rating standards vary greatly between organizations. Some of the most well-known organizations include (See Appendix B for links):

- Office of Juvenile Justice and Delinquency Prevention (OJJDP)
- SAMHSA's National Registry of Effective Programs and Practices (NREPP)
- Blueprints for Violence Prevention
- Promising Practices

- What Works Clearinghouse
- Strengthening America's Families
- Center for Mental Health Services (2000)
- Center for Substance Abuse Prevention (CSAP)
- Office of the Surgeon General
- Child Welfare League of America

Why Implement an Evidence-Based Program?

"To date, few of the programs identified as model or exemplary programs have been successfully implemented on a wide scale" (Elliott and Mihalic, 2004).

Despite decades of research on the causes and treatments of various problems within the social service arena, children and families still find themselves in crisis. Most approaches aimed at helping these families have shown only modest effect (August, et al., 2004, p. 2018). Thus, service providers have begun searching for programs with scientifically proven results. For example, SAMHSA reviewed more than 600 programs and only 11 programs were found to be effective. The National Registry of Evidence-based Programs and Practices, which examined programs in the substance use and mental health disciplines reviewed more than 1,100 programs and only found 150 that were viewed as model, effective, or promising programs. This is very telling and shows that most programs in the prevention field have either not been sufficiently researched to draw conclusions or have been and do not show positive effects.

Return on Investment

According to researchers, implementing evidence-based programs helps ensure that a program is based on a proven or tested theory of change. The results or client outcomes are directly related to the services received from the program. Second, evidence-based programming helps to ensure that agencies are spending resources on a proven program that works. We must be accountable to the families we serve, as well as to community stakeholders, funders, and taxpayers (Hyde, Falls, Morris, and Schoenwald, 2003). Third, funders want to invest in programs that have demonstrated outcomes, meaning a good return on investment. "In an era of increasingly tight fiscal budgets, public sector policymakers need more objective and impartial means of reviewing publicly funded programs to determine if the greatest value is being provided for the taxpayer's dollars. No longer can these policymakers assume that programs are effective simply because the program's supporters assert that they are effective" (Brown, 2005).

The 2003-05 Washington state operating budget required that the Washington State Institute for Public Policy (WSIPP) conduct research examining the benefits and costs of prevention and early intervention programs for children and youth. There were three main focus areas:

1. Identifying which programs produce a positive return on the dollar.
2. Developing criteria to ensure fidelity and quality of program implementation.
3. Developing recommendations for state legislation encouraging local governments to invest in evidence-based programming and providing these governments reimbursements for implementing such programs (Pizzigati, Stuck, and Ness, 2002).

What they found was that in fact, there are some programs that do produce positive effects and also generate more benefits than costs. Conversely, they also found that some programs were not good investments and were therefore an inefficient use of taxpayer money. According to Washington State Institute for Public Policy researchers, "the market for rigorously researched prevention and early intervention programs is young, but is evolving quickly. Most high-quality evaluations have been completed only in the last two decades, and many new rigorous studies will become available

Highlight

As of December 2006 in Florida, 405 youth had completed the Redirection Program, a program that utilizes the evidence-based programs multi-systemic therapy and functional family therapy. This program achieves the same outcomes as residential delinquency programs and in fact, when examined closely, youth that graduated from the Broward and Escambia counties' sites achieved significantly better outcomes, including lower re-arrest rates. Also noteworthy, the program has cost $3.1 million dollars so far in comparison to the $8.9 million it would have cost if these 405 youth had been placed in residential care (Rhen, 2007).

Research Review - Evidence-Based Programs and Practices: What Does It All Mean?

7

"As Peter Greenwood states in his book, "Changing Lives," a culture of accountability is the only way to "keep everyone honest" because it "allows for fair comparisons between competing approaches and programs." In the absence of such a culture, Greenwood predicts the industry will be unable to avoid the "flavor of the month phenomenon, which conjures up new, supposedly promising programs on the strength of their novelty, rather than proven track record" (Rhen, Evidence-based Associates, p. 1).

in the years ahead" (Pizzigati, Stuck, and Ness, 2002).

Local communities are being asked more than ever to invest in and implement proven programs. "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, and Greenberg, 2005). This means either implementing programs that have already been labeled evidence-based through a national process or proving that the programs they are running would be considered evidence-based if rated. These communities must prove that their programs work (i.e. that they are effective for the children and families they serve).

One of the major problems is that communities and small local agencies do not have the resources necessary to prove that their programs are effective because the type of studies that need to be conducted, namely randomized control 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.

Accountability

Implementing evidence-based programs assists agencies and organizations in moving towards accountability. Why does everyone need to be accountable? The answer is simple: too often programs continue to run without ever showing that what they do works for the children and families they serve. 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. Strangely enough, the government does not always support programs that have been shown to work. In fact, there is evidence that programs that do not work are being supported and funded, such as DARE, boot camps, and Scared Straight.

Case Study: Scared Straight

The Scared Straight Program, while implemented across the nation, has actually been shown to cause a small increase in subsequent criminal activity by participating youth. However, the Governor of Illinois recently signed legislation that required schools in Chicago to implement this program even though it is known to have harmful effects (Dodge, 2006).

The fact remains that we spend billions of dollars on social programs that may have absolutely no effect on the problems they are trying to eradicate, and in some cases may be harmful to participants.

HOT TOPIC #1

Concerns Regarding Evidence-Based Programs

Although there appears to be widespread movement towards evidence-based programs and practices, there is some skepticism. As with any movement, there is sometimes opposition, which is critical for the success of the change movement and journey. Opposition and thoughtful critique help leaders think about the goals and objectives of social change in a thorough and responsive way. In addition, there are many concerns that are currently being addressed or hopefully will be considered in the future. Here are a few examples:

1. Evidence-based programs do not take into account professional experience of practitioners.
2. Evidence-based programs and practices do not exist for all identified needs or for all target populations.
3. Researching programs in order to define them as evidence-based is very expensive.
4. Implementing evidence-based programs can be very expensive.
5. Providers may not have the capacity to implement an evidence-based program.
6. Providers may believe that adaptation is needed for program success.

What About Professional Expertise?

There has been growing concern from those working in the field that definitions of "evidence-based" do not take into account the personal and professional experience of those providing services to clients. At the same time, there is really no argument that not all programs work for all individuals or families. Because there has been concern regarding implementing programs as is, without taking into consideration the providers and their knowledge and expertise, many organizations have begun adopting definitions that emphasize a balance between research and practice.

For example, Buysse and Wesley from the FPG Child Development Institute at the University of North Carolina Chapel Hill define evidence-based practice as "a decision-making process that integrates the best available research evidence with family and professional wisdom and values." The American Psychological Association defines evidence-based practice as "the integration of the best available research with clinical expertise in the context of patient characteristics, culture, and preferences" (p. 5). The Institute of Medicine defines it as "the integration of the best research evidence with clinical expertise and patient values" (Levant, 2005, p. 5).

The issue that some researchers have with possibly changing a program/practice because of professional expertise and client need is that if there is a change in the client's behavior and the practitioner believes it was due to the program, the claim can not be substantiated without being scientifically studied. It could have been the result of maturation or because of additional assistance from family and friends (Leff, 2002). The concept is that after a program/practice is altered, it must be researched again for effectiveness.

Needs and Target Population Concerns

The research on what works in prevention and early intervention is in its infancy. As can be gathered from a historical perspective, this area of study has only been in existence for the past 60 years and only recently has the social service prevention field begun to scientifically study programs and discuss the possibility of dissemination. In addition, this was done first for the juvenile justice, substance abuse and mental health areas. There has been very little done in the areas of primary prevention and early intervention, such as early care and education. Thus, some skeptics note that because there are few evidence-based programs to choose from it is unethical to refuse access to programs that are in effect. However, as asserted by Elliot (2007), would it not be unethical to provide a program

to children that either does not work or may in fact be harmful? Thus, we have to be patient and remind ourselves that we are at the forefront of this movement and it is important to realize that what we are doing right now may in fact be effective. The only things to do now is research it and find out if we are really as effective as we think we are.

> The positive effects resulting from these programs may not be visible for years, and the small size of many programs makes it unlikely that they alone could affect city- or county-wide risk factors . . .
> (Moore, 2007)

Research is Costly

There is not always the opportunity or resources needed to conduct a strong research design. Many studies of this magnitude are extremely expensive and time consuming due to the need to test for sustainability. Most evidence-based programs show that their outcomes are sustainable for at least one year after leaving the program. A program can show positive effects, but if there is not a permanent change to the recipient's attitudes, knowledge or behaviors after program participation has ended, then the effects are not sustainable and are, therefore, inadequate.

In medical care, "an average of about 17 years is required for new knowledge generated by randomized controlled trials to be incorporated into practice…" (Leff, Conley, and Hennessy, 2006). This is a perfect example of just how long it can take to prove that a program or service works and have it ready for dissemination and use by others in the field. Another example is the Nurse-Family Partnership home visiting model. Program developer David Olds researched his program for over 21 years before allowing it to be replicated and disseminated to the general public. He has also established a national site that assists with the implementation of his program so that effectiveness is ensured.

Because demonstrating effectiveness is such a lengthy process, there are some advocates favoring implementation of programs based on having a strong theoretical foundation, background research supporting the program model and its activities, and finally, clinical experience. There is a debate whether this type of support is enough to move forward with implementing or continuing to support a program based on these factors. It is believed that while some programs may have credible evidence for their support, in the end they still may not be effective when outcomes are measured using a strong research design. In practice, this decision must be made by the funding organization. It is also critically important to understand that changes and effects may not be seen for years, and while very expensive, in the long run the costs of researching programs and finding what works and what does not will pay for itself in the positive effects it produces for the children and family it serves. In fact, many evidence-based programs have demonstrated effects decades later for the participants and have even shown positive impacts for the children of those that participated.

Implementation: A Key to Success

> "Although we have taken giant strides forward in determining "what works" and promoting the use of science-based programs, we have lagged behind in building the internal capacity of designers to deliver their programs. To move forward with a national prevention initiative, this gap must be addressed by funders and policymakers" (Elliott and Mihalic, 2004).

There has been some discussion and concern that the primary focus of evidence-based programs has been on researching what types of programs are researched-based and much less attention given to whether or not there is capacity to implement them. Implementation occurs in stages and there can be problems at any one of them. For example, according to Fixsen et al. (2005), there is what is termed "paper implementation," which is when a program completes the recorded theory of change for the new program, the second stage is "process implementation," also called the expressed or active theory of change and involves such components as training.

The last phase is called "performance implementation" and is known as the integrated theory of change and involves actually carrying out the program leading to outcomes for clients (p. 6).

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. They are simply not ready for the complexity of implementing such a program. Actually, it can take upwards of six to nine months to get a site ready for implementation (Elliott and Mihalic, 2004). One reason is that when an agency decides to implement an evidence-based program, there is almost always the need for some organizational change (Fixsen et al., 2005, p. 64). Fixsen and colleagues (2005, pp. 64-65) report that there are specific factors that are critical to such organizational change, including

- commitment of ongoing resources and support for providing time and scheduling for coaching, participatory planning, exercise of leadership, evolution of teamwork;
- commitment of leadership to the implementation process;
- involvement of stakeholders in planning and selection of programs to implement;
- creation of an implementation taskforce made up of consumers, stakeholders;
- suggestions for "unfreezing" current organizational practices;
- resources for extra costs, effort, equipment, manuals, materials, recruiting, access to expertise, re-training for new organizational roles;
- alignment of organizational structures to integrate staff selection, training, performance evaluation, and on-going training;
- alignment of organizational structures to achieve horizontal and vertical integration; and

**"To be effective, any design process must intentionally be,
from the beginning, a redesign process"** (Felner, et al., 2001, p. 189).

While disseminating information about evidence-based programs is useful, if there is no capacity to put the program into practice, then the likelihood of achieving positive outcomes becomes quite limited. As Chinman and colleagues (2005) argue, there are many points where the prevention process can falter, each increasing the possibility of poor outcomes. According to them, some of the most critical factors and steps that need to be addressed are as follows:

- Complexity of prevention programming - Conducting needs assessments; setting goals and objectives; choosing appropriate programming that fits the local context given current resources; planning, implementing, evaluating, and sustaining programs.
- System-level factors – Differences in theoretical orientations of researchers and practitioners; differences in training; lack of coordination between agencies and systems of care; lack of community readiness either to adopt programs or implement them with fidelity.
- Resources – Lack of needed resources to implement or sustain programming including both financial and technical.
- Adaptation – Issues concerning adapting programs to fit community characteristics – developers may not take into consideration dissemination issues and implementers may not consider issues such as generalizability and fidelity concerns.

Going to Scale and Fidelity

There continues to be questions about whether many of these programs can be brought to scale, meaning replicated with fidelity, given real-life circumstances. Moving a program to a community setting from a research setting is not just a change in location. For example, according to August et al. (2004) there are (1) Client factors, (2) Practitioner factors, (3) Intervention Structure factors, and (4) Organizational culture/climate factors that can impact implementation success.

- **Client Factors** - In real life settings, clients/patients can not be chosen in the same rigorous manner as is typical in research studies where these programs were tested. Other client factors include things like cost of the program and logistic issues, such as transportation. These are not usually problems when researching these programs in a controlled setting.
- **Practitioner Factors** - There will probably be a high degree of variation in the education, practice orientations, and qualifications of those individuals delivering the program. These practitioners will have backgrounds that are much different than those that provided the program to recipients in the research study. Most practitioners do not have experience delivering an evidence-based program. Individuals that are part of research projects are also very committed to the program model, implementing it with fidelity, and often have high job satisfaction.
- **Intervention Structure factors** - In a research study, scientists have complete control over the program implementation. It is implemented according to a scripted manual and there is strict supervision. These structures are quite different, and there will probably be less support, in real-life settings.
- **Organizational Culture/climate Factors** - Once a program moves out of the controlled setting, the organization or agency that decides to implement the program will have its own leadership with its own attitudes and management style, issues with financial and human resources, and organizational stress.

Possible Solutions to Implementation Concerns

August (2004) reports that there are in fact some important things an agency can do to help with implementing a program successfully. These points include making sure that there are collaborative relationships with program developers and other agencies and stakeholders. Each partner must feel ownership of the program to ensure accountability. The host organization and its staff must have a high degree of readiness and motivation for implementation. To measure this there are readiness tools, such as the Organizational Readiness to Change Scale and the Evidence-based Practice Attitude Scale for staff. There must also be open communication among all parties. Implementers must have staffs with sufficient education, who are given appropriate training and supervision. Personality of staff and theoretical orientation must also be examined. There must be cultural awareness and how it can influence outcomes. Particular attention must be given to recruitment and retention of participants. The implementing organization must also consider potential problems and begin exploring ways to solve future crises.

Elliott and Mihalic (2004) have various training recommendations from past research on replication of evidence-based programs, which include the following:

- Use interactive training methods (e.g. videos, role playing).
- Be firm in explaining the formal eligibility requirements for program staff (e.g. required skills, formal training, and education).
- Hire entire staff before training.
- Introduce the program to the staff before beginning training.
- Encourage administration to attend training.
- Expect staff turnover and begin planning and budgeting for it.
- Be ready to begin implementation right after training ends.

HOT TOPIC #2

Fidelity versus Adaptation

"…the critical question may not be will this program fit in this local context, but how does this context have to change for us to successfully implement this program here?" (Elliott and Mihalic (2004) quoting Lisbeth Schorr)

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). Adaptation or changing the program design is usually done because it makes the program more acceptable to the local environment. 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 Life Skills 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. Other typical adaptations include eliminating training components and changing dosage (Elliott and Mihalic, 2004).

Implementers will sometimes jeopardize fidelity for sustainability of the program. The problem is that the effects may be sustainable, but no longer effective. Many assumptions are made regarding implementing evidence-based programs, namely that in order to get local buy-in, the program must be changed, such as decreasing the intensity. Elliott and Mihalic (2004) report that there is really little research that supports the need to adapt programs, but acknowledge that language and cultural adaptations may be the exceptions (p. 51). However, they also state that every program does not need separate treatments for different sexes or racial/ethnic groups, especially when the program is geared towards children and adolescent populations. There is a question though about whether a program that works well for rural teen mothers would work with inner-city teens. The program may or may not be able to be generalized to this different population and would need to be further evaluated.

Does changing or adapting the original design of an evidence-based program mean that it will not be as effective? Not necessarily, the adaptation may be as effective, more effective, or not effective at all. The problem with adaptation is that, in many cases, we just do not know if it will be as effective as the original program because it has not been experimentally evaluated.

Some adaptations have been evaluated. For example, there was an adaptation to the Nurse-Family Partnership program using paraprofessionals to complete the home visits instead of nurses. What researchers found was that the program was not as effective in reaching its outcomes. Further analysis revealed that nurses completed more visits than paraprofessionals and spent more time focusing on personal health during pregnancy and on parenting an infant. Paraprofessionals visited for longer time periods than nurses and spent more time on environmental concerns (safety of the environment, including living conditions and domestic violence issues and the ability to provide adequate food, clothing, and shelter). Paraprofessionals experienced greater staff turnover. These differences in implementation caused the program adaptation to fail to achieve positive outcomes for participants (Korfmacher, O'Brien, Hiatt, and Olds, 1999; Hiatt, Sampson, and Baird, 1997).

If a program is implemented as designed with the intended population, the need for an outcome evaluation is eliminated and the evaluation becomes focused on process and program-design adherence. According to many advocates of evidence-based programs, strict fidelity is essential to program effectiveness. On the other hand, there is some research that shows that "sensitivity and flexibility in administering therapeutic interventions produces better outcomes than rigid application of manuals or principals (Levant, 2005, p. 14).

Evidence-based program developers do not necessarily disagree with this belief or this research. That is why there has been some discussion regarding moving towards researching what the core components or critical ele-

ments are to any intervention. The federal government is supporting current research in this area. Core components are defined as the "essential and indispensable" elements of a program or practice needed in order to reach outcomes (Fixsen, 2005, p. 24). The goal is to highlight these and then be able to do some adaptation without decreasing effectiveness. Replication of programs will need to occur before these core components can be established. Furthermore, there is also some agreement that if at all possible, a program should first be implemented with fidelity before adaptation begins and research shows that when this occurs, adaptations are more successful (Fixsen, 2005).

While it is in its infancy, some research is examining generalizability and transportability of evidence-based programs into community settings. It is important that if you are contemplating adapting a program from its original format that (1) you contact the program developer and ask about the core components. The developer may have an understanding of how important an omission may be to the outcomes, and (2) you understand the theoretical foundation that the program is premised on so that you can preserve it when making changes (Chinman, Imm, Wanderman, 2004, p. 47).

References

August, G.J, Winters, K.C., Realmuto, G.M., Tarter, R., Perry, C., and Hektner, J.M. (2004). Moving Evidence-Based Drug Abuse Prevention Programs From Basic Science to Practice: "Bridging the Efficacy-Effectiveness Interface." *Substance Use and Misuse.* 39(10-12): pp. 2017-2053.

Brown, A. (July 2005). Determining What Works (and What Doesn't) in the Public Sector. *Public Management.* Civic Renewal and Community Building.

Buysse, V. and Wesley, P. W. (September 2006). *Evidence-Based Practice Empowers Early Childhood Professionals and Families.* FPG Child Development Institute at The University of North Carolina at Chapel Hill. FPG Snapshot #33.

Centre for Evidence-based Medicine. (2004). Glossary of EBM Terms. University Health Network. <www.cebm .utoronto.ca/glossary/index.htm>.

Chinman, M., Hannah, G., Wandersman, A., Ebener, P., Hunter, S.B., Imm, P. and Sheldon, J. (June 2005). Developing a Community Science Research Agenda for Building Community Capacity for Effective Preventive Interventions. *American Journal of Community Psychology.* 35(3/4): pp. 143-157.

Chinman, M., Imm, P., and Wanderman, A. (2004). *Getting To Outcomes 2004: Promoting Accountability Through Methods and Tools for Planning, Implementation, and Evaluation.* Chapter Three- Question #3: Which Evidence-Based Programs Can Be Used to Reach Your Goal? (Best Practice). Rand Corporation. <http://www.rand.org/pubs /technical_reports/TR101/>

Crane, J. and Barg, M. (April 2003). D*o Early Childhood Programs Really Work?* Coalition for Evidence-Based Policy.

Dodge, K. (2006). Professionalizing the Practice of Public Policy in Prevention of Violence. *Journal of Abnormal Child Psychology.* 34: pp. 475-479.

Elliott, D.S. and Mihalic, S. (March 2004). Issues in Dissemination and Replicating Effective Prevention Programs. *Prevention Science.* 5(1): pp. 47-53.

Elliott, D.S. (2007). Personal Communication via the Senior Executive Policy Institute.

Felner, R. D., Favazza, A., Shim, M., Brand, S., Gu, K., & Noonan, N. (2001). Whole school improvement and restructuring as prevention and promotion – Lessons from STEP and the project on high performance learning communities. *Journal of School Psychology*, 39(2), pp. 177-202.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M. & Wallace, F. (2005). *Implementation Research: A Synthesis of the Literature.* Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication #231).

Hiatt, S.W., Sampson, D., and Baird, D. (1997). Paraprofessional Home Visitation: Conceptual and Pragmatic Considerations. *Journal of Community Psychology.* 25(1): pp. 77-93.

Hyde, P.S., Falls, K., Morris, J.A., and Schoenwald, S.K. (Fall 2003). *Turning Knowledge into Practice: A Manual for Behavioral Health Administrators and Practitioners about Understanding and Implementing Evidence-Based Practices.* The American College of Mental Health Administration.

Karoly, L.A., Kilburn, M.R., and Cannon, J.S. (2005). *Early Childhood Interventions: Proven Results, Future Promise.* Rand Corporation. <www.rand.org>.

Frequently Asked Questions

(written by EBP committee)

What is an evidence-based program?

An evidence-based program is 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. EBPs may incorporate a number of evidence-based practices in the delivery of services.

Why are evidence-based programs important?

Implementing evidence-based programs is important to ensure that resources are spent on programs that have a high probability of achieving desired, long-term outcomes and that incorporate principles of effective programming that bring about positive results. The advantage to both funders and providers is that EBPs eliminate the costly and time-consuming efforts of exploring and experimenting with new methods, strategies and technologies. They provide the best return on investment.

Who makes the distinction whether something is evidence-based?

There are many organizations that now have procedures for rating programs in order to designate them as evidence-based. CSC will also be creating criteria in which to determine which programs are evidence-based and what is needed to move a program towards becoming evidence-based.

Why is CSC interested in EBP? Why now?

CSC's top priority has always been to see that the children and families served by its funded programs achieve the best possible outcomes. We have also, since our inception demanded the highest level of accountability – of ourselves and those we fund – to provide the highest return on investment for our county's taxpayers and stakeholders. As this report details, the development of evidence-based programming in social services is a relatively new one, but one that vastly reduces hit or miss outcomes and results. This continuous improvement mindset is part of CSC's leadership philosophy, core values, and behaviors. Now that CSC is at a place where it can provide a supportive infrastructure (e.g. training) for EBP development, we will move forward to assist programs/agencies with this advanced level of work.

How do we know the movement towards EBP is not just another fad?

As grants change, tax dollars dwindle, and boards require higher standards for the accountability of provided resources, social service programs much adapt by better measuring program success and evidence-based outcomes. This expectation to provide evidence-based outcomes is happening nationally across many states. As a result, EBP is not being viewed as another "fad," but rather it is an understandable expectation by funders and boards to ensure accountability of provided resources and to produce outcomes/results that truly make a difference for the clients being served.

I thought we were doing that already, what is different now?

Historically, CSC has funded a few, select programs that are considered evidence-based (e.g. HIPPY). While CSC has required agencies and programs to provide data in the past, and will continue to do so in the future, this does not necessarily mean that these programs are evidence-based. The term "evidence-based" refers to programs that have theoretical underpinnings, have met specific criteria, and have been proven to be effective. It is CSC's goal to help its currently funded agencies and programs move toward becoming evidence-based. This can be a long and laborious process; however, prior work (i.e., data collection, development of logic models and theories of change) has actually helped

move programs forward in this direction. Such prior work also allows programs to work toward continuous quality improvement, ensuring accountability and provision of the best services possible for our families.

How can we learn about evidence-based programs? And how is CSC going to help us become evidence-based?

There are a number of resources available to educate individuals about evidence-based programs, including the various organizations that define and rate programs using specific criteria. Links to these organizations can be found in Appendix B of this document. The reference section of this document also provides many sources of information. CSC will be hosting training on evidence-based programs beginning in September 2007 as part of a mini-series dedicated to the topic of effectiveness. There is also a CSC Professional Development Committee working on preparing professional development opportunities to prepare staff to build their knowledge, skills and abilities necessary to work within an evidence-based program environment. CSC will make available resources for professional development, monitoring and technical support.

Are we "ready" for evidence-based practices?

We won't know unless we try. We must follow logical steps during the planning and design stages – do a substantial amount of research about our own population and the programs and services deemed evidence-based that may lead to the outcomes we are seeking and determine if we can match the two.

How long does it take to become evidence based?

While there is no specific time table in becoming evidence-based, key elements must be in place before evaluating success. This process can take extend over several years. Every area and program is different and some take longer to determine whether they are successful. Much also depends on the results sought. Lastly, considering that a program must have completed randomized control trials and have shown positive outcomes which have lasting client effects, the process is long and can be costly.

What measures can we take to move ourselves towards evidence-based programming?

It starts with a solid foundation. To begin with, you must ensure that your agency is clear on its mission and goals and that the program staff is committed to the goal(s), outcomes and procedures required by the funding agency. Programs can begin by developing a sound theory of change stating what they believe will effect change for the clients; develop a logic model (a road map as to how they plan to achieve their outcomes); provide data that is submitted on time and is clean and easily manipulated for evaluation; and implement their existing program with fidelity. Programs should be fully aware of what is in their PIE (Program Implementation and Evaluation) and assuring that they are adhering to it. A program that wishes to move towards becoming evidence-based should be collecting data for their program and be able to understand how to implement change based on findings from the data.

All members of the agency's staff involved in the program must take advantage of ongoing professional development and periodic performance assessment. Finally, evidence-based programs not only assess themselves regularly for continuous quality improvement, but seek ongoing feedback from clients regarding satisfaction.

Will we be de-funded if we are not evidence-based?

As we will continue to emphasize, we are all in the business of improving outcomes for our children and families. CSC has and will continue to make data-driven decisions – on services and programs needed, outcomes sought and who is most capable of providing those services and achieving those outcomes. The road toward becoming evidence based only heightens the importance of clean, accurate, complete data for you, as well as for CSC to be able to see what's working and what's not and make mid-course corrections.

Korfmacher, J., O'Brien, R., Hiatt, S. and Olds, D. (1999). Differences in Program Implemenattion Between Nurses and Paraprofessionals Providing Home Visits During Pregnancy and Infancy: A Randomized Trial. *American Journal of Public Health*. 89(12): pp. 1847-1851.

Kyler, S.J., Bumbarger, B.K., and Greenberg, M.T. (April 2005). *Technical Assistance Fact Sheets: Evidence-Based Programs*. Pennsylvania State University Prevention Research Center for the Promotion of Human Development.

Leff, H. S. (2002). Section V: Insurance for Mental Health Care: Chapter 17. *A Brief History of Evidence-Based Practice and a Vision for the Future*. SAMHSA's National Mental Health Information Center. <http://mentalhealth .samhsa.gov/publications/allpubs/SMA04-3938/Chapter17.asp>.

Leff, S., Conley, J., Hennessy, K. (Summer, 2006). Marketing behavioral Health: Implications of Evidence Assessments for Behavioral Health Care in the United States. *International Journal of Mental Health*. 35(2): pp. 6-20.

Levant, R.F. (2005, July 1). Report of the 2005 Presidential Task Force on Evidence-Based Practice. *American Psychological Association*.

U.S. Department of Education (2003). *Proven Methods: Questions and Answers on No Child Left Behind* – Doing What Works. <www.ed.gov/nclb/methods/what works/doing.html>.

Yeager, C. (Summer 2007). *Nurse-Family Partnership: Helping First Time Parents Succeed*. In Why Evidence-Based Associates. The EB-Advocate. Evidence-Based Associates.

Appendix A: CSC's 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 groups, 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 and 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.

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 regards to its stated outcomes or goals (Maxfield and 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. 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 the 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 and Babbie, 2005, p. 436).

Fidelity – Extent to which delivery of an intervention adheres to the protocol or program model originally developed (Mowbray, Holter, Teague, and 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 and 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 and Brief Strategic Family Therapy).

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-experiment – This research design is very similar to and almost meets criteria for an experimental design, but is unable to control 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 our research, determining what things you 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.

Appendix B: National Ratings

Organization	Focus	Rating	Ratings Defined
Blueprints for Violence Prevention <www.colorado.edu/cspv/blueprints>	Violence Prevention	• Model • Promising	**Model:** 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. **Promising:** Programs that show a deterrent effect using either an experimental or quasi-experimental design.
Center for Substance Abuse Prevention (CSAP) <www.modelprograms.samhsa.gov>	Substance Abuse Prevention	• Model • Effective • Promising	**Model**: Programs that are evidence-based (conceptually sound, internally consistent, have sound methodology, credible, can be generalized). Programs have utility, are well-implemented, well-evaluated, produce a consistently positive pattern of results to the majority of intended recipients Developers must show that the programs is available for dissemination and provide TA to others wishing to implement the program (Must score ≥ 4.0). **Effective**: Same as above, however, not currently available for wide dissemination to the general public (Must score ≥ 4.0). **Promising**: Programs that demonstrate some positive outcomes, but require additional evidence and research showing consistent positive results. (Must score ≥ 3.33).
Center for Mental Health Services (1999 Report) <http://www.prevention.psu.edu/pubs/docs/CMHS.pdf>	Reducing the Risk/Effects of Psychopathology	• Effective • Promising	**Effective**: 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. **Promising**: 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.
Department of Education <www.ed.gov> <http://www.ed.gov/admins/lead/safety/exemplary01/exemplary01.pdf >	Reducing Substance Use, Violence and Other Conduct Problems	• Exemplary • Promising	**Exemplary**: The program is based on empirical data and demonstrates evidence of effectiveness in improving student achievement. **Promising**: The program provided sufficient evidence to demonstrate promise for improving student achievement. Programs are rated according to evidence of efficacy, quality, educational significance, and usefulness to others.

Organization	Focus	Rating	Ratings Defined
Mihalic and Aultman-Bettridge (2004) In William L. Tulk (Ed.), *Policing and school crime*. Englewood Cliffs, NJ: Prentice Hall Publishers.	Reducing school disciplinary problems, suspensions, truancy, dropout, and improving academic achievement	• Exemplary/ Model • Promising • Favorable	**Exemplary/Model**: 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) **Promising**: Programs that show a deterrent effect using either an experimental or quasi-experimental design. (Based on Blueprints) **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.
National Institute of Justice (NIJ) –Report by Sherman, et al (1998) Research in Brief (1998) <http://www.ncjrs.gov /pdffiles/171676.pdf> Full report to Congress (1997) <www.ncjrs.org /works/wholedoc.htm>	Crime and drug abuse prevention	• Working/ Effective • Promising	**Working/Effective**: 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. **Promising**: 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. ▶ 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).
The Gilford Center <http://www.guilfordcenter .com/provider/practices /default.htm>	Adult services; substance abuse prevention and treatment child services, mental health and systems of care; developmental disabilities	• Best Practice • Emerging Best Practice • Evidence-Based Practice	**Best Practice**: 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. **Emerging Best Practice**: Interventions or services that have shown benefit to consumers, but have not yet been established as evidence-based practices through rigourous scientific research. **Evidence-Based Practice**: Intervention for which there is consistent scientific eveidence showing that it improves client outcomes.

Organization	Focus	Rating	Ratings Defined
Promising Practices Network <http://www.promisingpractices.net/programs.asp>	Children and Families	• Proven • Promising • Proven/Promising	**Proven**: 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-control 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. **Promising**: 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. **Proven/Promising**: 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.
Juvenile Justice Evaluation Center <http://www.jrsa.org/jjec/resources/evidence-based.html>	Youth Violence	• Model Programs • Promising Approaches • Innovative Approaches	**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.

Organization	Focus	Rating	Ratings Defined
Substance Abuse and Mental Health Services Administration Center for Substance Abuse Prevention <http://www.whitehousedrugpolicy.gov/prevent/pdf/science.pdf>	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	**Type 1**:Not scientifically defensible. The Program/principle has been defined or recognized publicly, and has received awards, honors, or mentions. **Type 2**: 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 **Type 3**: Expert/peer consensus process - scientifically defensible. The program's source documents have undergone thorough scrutiny in a 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. **Type 4**: 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. **Type 5**: 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. ▶ 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
What Works Clearinghouse <www.whatworks.ed.gov/>	Educational Interventions (programs, products, practice, and policies)	• Meets Evidence Standards • Meets Evidence Standards With Reservations • Does Not Meet Evidence Screens	**Meets Evidence Standards**: 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. **Meets Evidence Standards with Reservations**: 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. **Does Not Meet Evidence Screens**: Studies that provide insufficient evidence or causal validity or are not relevant to the topic being reviewed ▶ In addition, the standards rate other important characteristics of study design, such as intervention fidelity, outcome measures, and generalizability.

Organization	Focus	Rating	Ratings Defined
Helping America's Youth <http://guide.helpingamericasyouth.gov/default.htm>	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	**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, an 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.
Communities That Care, Developmental Research and Programs Posey, R., Wong, S., Catalano, R., Hawkins, D., Dusenbury, L., & Chappell, P. (2000). *Communities That Care Prevention Strategies: A Research Guide to What Works*. Seattle, WA: Developmental Research and Programs, Inc. <www.preventionscience.com>	Substance abuse, delinquency, teen pregnancy, school dropout, violence, and child and youth development.	• Effective	**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.
Office of the Surgeon General <http://www.surgeongeneral.gov/library/youthviolence/chapter5/sec2.html#ScientificStandards>	Youth Violence	• Model • Promising • Does Not Work	**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.

Organization	Focus	Rating	Ratings Defined
OJJDP **Model Programs Guide** <http://www.dsgonline.com /mpg2.5/mpg_index.htm>	Entire continuum of youth services from prevention through sanctions to reentry.	• Exemplary • Effective • Promising	**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.

Organization	Focus	Rating	Ratings Defined

Exemplary and Promising Safe, Disciplined and Drug-Free Schools Programs 2001 <http://www.ed.gov/admins/lead/safety/exemplary01/report_pg3.html>	Safe, Disciplined and Drug Free Schools	• Exemplary
		 • Promising
			Exemplary: Based on empirical data a program was effective.
			 Promising: There is sufficient evidence to demonstrate that the program showed promise for improving student achievement.
			
			 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 (Replicability): The program provides necessary information and guidance for replication in other appropriate settings.

Federal Government/ Office of Management and Budget <http://www.whitehouse.gov/omb/expectmore/perform.html>	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	**Effective**: This is the highest rating a program can achieve. Programs rated Effective set ambitious goals, achieve results, are well-managed and improve efficiency. **Moderately Effective**: 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. **Adequate**: This rating describes a program that needs to set more ambitious goals, achieve better results, improve accountability or strengthen its management practices.

Organization	Focus	Rating	Ratings Defined
Strengthening America's Families <http://www.strengtheningfamilies.org/html/programs_1999/Review_Criteria.html>	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	• Exemplary • Model • Promising	**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
Ohio State CLEX <http://www.alted-mh.org/ebpd/criteria.htm>	Youth Behavior Mental Health Alternative Education	• Evidence • Model • Promising	**Evidence Checklist**: 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 and consensus. • 0-2 checks: ***Unproven approach***: No documentation approach has either ever been used or has been implemented successfully w/no evaluation. • 3-5 checks: ***Promising Approach***: Implemented & 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. • 6-10 points: ***Evidence Based***: Compelling evidence of effectiveness. Attribute participant success to the program itself, & have evidence that the approach will work for others in different environments. **Model**: Meets the satisfactory standards of specific criteria as an effective program.

Organization	Focus	Rating	Ratings Defined
Child Welfare League of America <http://www.cwla.org /programs/r2p/levels.htm>	Child Welfare	• Exemplary Practice • Commend-able Practice • Emerging Practice • Innovative Practice	**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 1 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.
Child Trends <http://www.childtrends .org/what_works/clarkwww /clarkwww_intro.asp>	Life Course Models, teen programs, school readiness, and afterschool	• What Works • What Doesn't Work • Mixed Reviews • Best Bets	**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, nonexperimental 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.

Organization	Focus	Rating	Ratings Defined
American Community Corrections Institute <http://www.accilifeskills .com/evidence-based /research.htm>	Substance Abuse Offenders	• No specifics	**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.

Children's Services Council
PALM BEACH COUNTY

2300 High Ridge Road
Boynton Beach, FL 33426