Evaluating the Impact of Evidence-Based Practice and Policy in Public Health

A Case Study on Parent-Child Interaction Therapy

Sol Bee Jung
2011-12-09
# Table of Contents

**ABSTRACT** .................................................................................................................. 4

**Introduction** ................................................................................................................. 5

The Role of Evidence in Public Health ........................................................................... 5

The Evidence-based Practice Movement and Its Strict Criteria for Evidence ................. 5

**Implementation of Evidence-Based Practice** ............................................................... 7

System change .................................................................................................................. 9
Practitioner selection and training .................................................................................... 9
Policy-practice knowledge flow ....................................................................................... 11

**Barriers to Implementation of EBPs** ........................................................................... 11

**Barriers to Achieving the Intended Impact of EBPs** .................................................... 13

**Exchange Club’s Family Center in Durham** ................................................................. 14

PCIT ................................................................................................................................... 15

Home-based (In-Home) PCIT ......................................................................................... 18
Family Center’s Implementation of PCIT ......................................................................... 20

**Methods** ....................................................................................................................... 22

Description of Data .......................................................................................................... 22

Subjects ............................................................................................................................ 22

Practitioners ...................................................................................................................... 23

**Measures** ..................................................................................................................... 24

AAPI .................................................................................................................................. 24
Satisfaction Survey .......................................................................................................... 27
Referral Source ................................................................................................................ 27
Closure Code .................................................................................................................... 28
Mandatory Referral .......................................................................................................... 29

**Data Analysis** ............................................................................................................... 29

Covariates ......................................................................................................................... 30

**Results** .......................................................................................................................... 31

General Linear Model Analysis of the AAPI Results ....................................................... 31

Inappropriate Expectations .............................................................................................. 33
Lack of Empathy ............................................................................................................... 35
Physical Punishment ......................................................................................................... 36
Role Reversal .................................................................................................................... 37
General Linear Model Analysis of the Satisfaction Survey Results ........39
Analysis of Provider Effect on Outcomes ........................................... 41

Discussion ........................................................................................................ 41

Interpretation of AAPI Outcomes ................................................................. 42
  Inappropriate Expectations ...................................................................... 43
  Lack of Empathy ......................................................................................... 43
  Physical Punishment ..................................................................................... 44
  Role Reversal ................................................................................................. 45
  Limitations .................................................................................................... 47

Interpretation of Satisfaction Survey Results .................................................. 48

Conclusion ........................................................................................................... 48

References ....................................................................................................... 53
ABSTRACT

The purpose of this study was to explore the evidence-based practice movement and its implications for the realm of public health. The paper first described the evidence-based practice movement and examined its pros and cons, including some of the barriers to implementation. A case study on Parent-Child Interaction Therapy (PCIT), an evidence-based practice (EBP), was then conducted to examine whether an EBP resulted in better outcomes than treatment as usual for clients from one agency in Durham. The findings from the data analyses showed that PCIT did not necessarily result in better outcomes than treatment as usual based on the AAPI results. PCIT also did not serve all racial groups equally well. Latino families achieved greater improvement on most subscales from treatment as usual than from PCIT, whereas African American and Caucasian clients achieved more improvement from PCIT than from the comparison treatment. Analysis of satisfaction surveys showed that PCIT clients expressed greater satisfaction than comparison treatment clients in two aspects. PCIT clients tended to believe that their providers were more knowledgeable and that their providers respected the family’s way of doing things. No significant provider effects could be observed from the data analysis.
Introduction

The Role of Evidence in Public Health

“Evidence” holds significance in public health. Researchers studying existing medical interventions and probing into innovative treatment methods are devoted to finding substantial and useful evidence from their studies. Some politicians and decision makers in the government rely on evidence to strengthen their arguments for political debates. Insurance companies and other funders of clinical treatments are interested in the evidence provided by researchers in order to understand their investment options in pursuit of highest returns. Service providers depend on evidence to make decisions and resolve any dilemmas they come across during clinical sessions. Evidence provides them with the assurance and confidence that their services are helping their clients because it certifies that their treatments are beneficial rather than harmful or ineffective. Additionally, clients or patients who are the end consumers of these interventions may desire to be informed better about the treatments they are receiving.

The Evidence-based Practice Movement and Its Strict Criteria for Evidence

While it is difficult to deny the importance of evidence, the research-to-practice gap has been a chronic problem in the mental health field. A need to close this gap brought about a movement to promote the use of quality evidence to support clinical services, called the evidence-based practice (EBP) movement.

From a few decades ago, with the advent of technological advances contributing to the rapid dissemination of information, the public increasingly began to realize that policies grounded in erroneous evidence can harm consumers. There was increased public awareness about the prevalence of “quackery,” or non-scientific, fake clinical treatments harming patients rather than helping them. Reliance on scientific research data and good evidence was
believed to be a warrant for any action taken in clinical settings (Goodman, 2003). Many also recognized that at times, overly optimistic evaluations of policies, bolstered by suspect evidence, may preclude research efforts and development in fields that need further investigation (Gambrill, 2006). In addition, practitioners were also expected to fully understand the services they provided. It became more and more important that they knew what quality evidence was so that they could utilize this knowledge to correctly interpret and apply research findings to their treatments. It was recommended they make independent assessment of evidence and combine it with their own clinical judgments for maximum effectiveness. (Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996). These rising expectations partially gave rise to the recent EBP movement.

The rigorous standards of this movement require that a practice, in order to be labeled as evidence-based, have the best scientific support for its effectiveness. Here, “best scientific support” refers to critically assessed evidence obtained from well-designed, impartial studies and reviews, which are free of research biases, such as the confirmation, allegiance, or sampling bias.

One method of acquiring such evidence is the randomized controlled trial (RCT). RCTs are used in order to minimize testing bias and control for confounding factors that may alter conclusions. The rigid design of such studies and close monitoring of external variables allow researchers to minimize the possibility that findings are due to chance. For this reason, RCTs can be replicated under routine conditions by any researcher, as long as the conditions of the experiments are strictly adhered to (Victora, Habicht, & Bryce, 2004).

A systematic review, such as meta-analysis, is the second method of acquiring sound evidence. In conducting systematic reviews, researchers search for all relevant literature, in all languages, from both published and unpublished sources, incorporating diverse methods including hand searches of journals. Systematic reviews of past literature and studies ensure
that all available data are critically analyzed and incorporated into the comprehensive evaluation of a treatment. No studies are neglected or left out for reasons such as convenience or personal agenda. Researchers acknowledge that sometimes studies are left unpublished because they lack substantial results and do not offer new insights into a topic. This is coined the “file-drawer problem.” Researchers realize that despite their extensive search methods, some unpublished studies may still be left out from their collection of past literature. In conducting systematic reviews, they use scientific and statistical tools to account for this possibility and measure the hypothetical impact that omitted studies may have on their conclusions (Gambrill, 2006).

Quality evidence obtained mainly through these two methods—RCTs and systematic reviews—possess three main characteristics: rigor, replicability, and independence. Increasingly, researcher, practitioners, policymakers, and the general public have come to understand that evidence can have varying degrees of strength, reliability, and rigor. They also realized that it was not enough for clinical interventions to be developed upon scientific grounds, but that these grounds must also be tested. Such understanding led to rising demands for strong evidence in policy and service-related decision-making and has created a stronger push for the EBP movement.

**Implementation of Evidence-Based Practice**

Despite the persistent push for the movement, for years, EBP had made little progress in actual practice and field settings because dissemination of information and application of research findings in actual treatment was considered to be a passive process. In this process, researchers had the role of conducting rigorous studies for innovation and advancement of interventions and publishing findings of wide-scale impact. Application of this research was left to the agency administrators and practitioners in charge of implementing the interventions.
Clients would be affected by these modifications without information on how they are actually impacting them and their well-being (Fixsen, Blase, Naoom, & Wallace, 2009).

Only in the past few decades did implementation—of which factors like practitioner training and treatment fidelity are critical components—begin to be considered more seriously as an important part of the EBP. The implementation process involves researchers and agencies and practitioners, and influences clients; the aim is to achieve high fidelity utilization of research findings and sound scientific evidence, in order to ensure greatest benefits for the clients or consumers (NIRN, 2008b).

These trends are reflected in the policy sector as well. Policy decision makers are realizing that the human services, interventions serving the population in diverse ways to ensure individual well-being and comfort, are often conducted in an inconsistent manner, produce only limited impact and sometimes cause harm rather than benefit to clients. As a result, there is a call for system transformations, led by a reliance on the evidence-based approach to bring about innovations and progress (NIRN, 2008a).

There are several components of implementation. As general awareness about the importance of implementation began to increase, the research and service community began to take greater interest in determining the specific steps and skills required for efficient implementation of treatments and interventions.

The central components of implementation have been identified through research. They are staff selection, pre-service and in-service training, ongoing coaching and consultation, staff evaluation, decision support data systems, facilitative administrative support, and systems interventions. These crucial components can largely be broken down into system change, practitioner selection and training, and policy-practice knowledge flow.
System change

System change refers to the change of stakeholder mindsets and increase in external support. Possibly the most important piece of this component is the agency administrators’ willingness to support the shift to EBPs (Fixsen, Blasé, Naoom, & Wallace, 2009). These people possess the greatest influence over the practitioners under their supervision. EBP is a collective effort. Unless agency heads can be persuaded to change common practice in their organizations, EBP cannot become the normal standard of practice.

Practitioner selection and training

Practitioner selection and training are also crucial and difficult elements of implementation. Effective preparation preceding the actual implementation of practice can be broken down into two parts: training and coaching. Training is where the dissemination of updated information from the research community takes place. Coaching assures that the training is leading to changes in practice. This makes the implementation process more active and eventually reduces the gap between science and service. However, there are some skills that cannot be acquired even with the most intensive preparation programs. They are the innate characteristics of practitioners and have little relation to the amount of training or schooling acquired. These qualities encompass common sense, sense of social justice, ethics, willingness to learn, empathy, and the ability to relate to clients (Fixsen, Blasé, Naoom, & Wallace, 2009).

Training and coaching are two necessary parts to the practitioners’ acquisition of full mastery in applying interventions and carrying out their core components. From two decades ago, the perception of training objective has undergone big shifts. Joyce and Showers, two authors who have studied in depth about the practice of implementation, claim that contemporary training designs distinguish more effectively between raising practitioner
awareness of innovative and efficient practice and inducing concrete behavior change as objectives of training. To accomplish the latter goal, coaching plays a crucial role (Joyce & Showers, 1988).

Research on training effectiveness found that traditional ways of training practitioners—first teaching theories behind treatments, followed by discussions and demonstrations, and providing numerous practice sessions to try the acquired skills—resulted in only 5-10% of behavior modification in actual service. On the other hand, coaching on the job led to behavior changes, or adherence to learned methods, in 95% of the trained practitioners (Fixsen, Blasé, Naooom, & Wallace, 2009). Coaching is only possible with the full support and participation of agency heads and the willingness of the practitioners to follow instructions and modify their treatments. Joyce and Showers also introduce the concept of peer coaching.

As a part of practitioner development, evaluations also play a big role. Evaluations are important because they assess how well trained staff are utilizing acquired skills and implementing the treatment as informed and taught. Furthermore, evaluations ensure that staff—practitioners, supervisors, agency leaders—are keeping with the core factors of EBPs identified through research, RCTs, systematic reviews and other methods of independent, rigorous trials. These critical elements of EBPs must be strictly adhered to. In many cases, the effects of such services are nullified as a result of idiosyncratic adaptation, which deprives the clients of their right to receive healthcare service of the highest quality that is available (Fixsen, Blasé, Naooom, & Wallace, 2009).

These assessments of practice include measures of fidelity. High fidelity implementation produces the expected results of treatments which have been rigorously tested and proven to be beneficial.
Policy-practice knowledge flow

Third, the policy-practice knowledge flow refers to continuous evaluation process, which provides both sectors with better, more updated information about interventions. EBP is policy-led practice enabled by practice-informed policies. The information is flowing in mutual directions, and advancements are ceaselessly sought by both the policy community—informed and assisted by the researchers—and the service providers. Old versions of interventions are continuously updated and norms of practice are not taken at face value. Implementation is only complete when practice is evolving with the incoming flow of knowledge and when practitioners and agency administrators are receptive to such updated information and are fully open to incorporating them into their practices (Aos, Lieb, Mayfield, Miller, & Pennucci, 2004).

Barriers to Implementation of EBPs

Some barriers lie on the road to the broad implementation of EBPs. There are feasibility issues due to all the conditions that EBPs must meet, such as the required use of RCTs and systematic reviews. Moreover, political interests and incumbency issues arise, further inhibiting the use of EBPs as standard practice.

It is indeed not to be assumed that there will be universal responses in every patient to clinical interventions. There is too much variability in real-world practice to expect there to be one-size-fits-all solutions with cookbook-style implementation manuals to any public health or social issue. External influences can vary from fidelity of implementation and training of practitioners to characteristics and values of clients and available resources. Because practitioners have already had firsthand experience coping with unpredictable situations that take place on the field, they reject the idea of relying purely on science to tell them exactly what to do (Victora, Habicht, & Bryce, 2004). They also hold doubts about the
validity and usability of the empirical data presented by researchers through controlled studies. Therefore, they believe evidence can play only a limited role and that the gaps need to be filled in through clinical experience and expert judgment (Clarence, 2002).

The practitioners’ concerns about the feasibility and actual impact of EBPs indeed are valid points. It is not enough to create and test EBPs; they must be implemented correctly in order to have the anticipated impact. Attention must be given to application barriers in agencies, and there must be sufficient support for these efforts. Better evaluations are necessary to assess the returns for the budget spent in conducting or switching to EBPs.

Many have attempted to formulate means by which to stably establish EBP as standard practice. They have contemplated the realistic issues, including problems of cost, which arise during the transition of EBP. A well-known advocate of evidence-based practice and policymaking, Steve Aos, describes the portfolio approach as an effective way to allocate scarce resources to acquire maximum efficiency and impact for the cost spent on public interventions. The portfolio approach divides total resources into several investments. Some options would have been proven through rigorous studies, such as RCTs and systematic reviews, to work and others would be treatments that have not yet been explored in depth and need to be further investigated. This allows for continuous growth and progress in both research and practice, eventually leading to greater benefits for the stakeholders, especially the clients, or consumers, of intervention. It is more realistic in the sense that the approach takes into account the possibility of unexpected events, where something goes wrong with approved treatments or the conditions change affecting the nature or scale of their impact (Aos, Lieb, Mayfield, Miller, & Pennucci, 2004). Also, it allows for development of services for sub-populations and minorities, for whom a mainstream intervention may not work.

For the EBP movement to be accepted and fully implemented, knowledge and skills in this area must be enhanced. The goals of EBP and the means by which these goals are
achieved—through the collaborative efforts of all the stakeholders—need to be explained better and communicated to those who possess doubts (Melnyk, Fineout-Overholt, Stone, & Ackerman, 2000).

**Barriers to Achieving the Intended Impact of EBPs**

There are also additional barriers to the movement due to the difficulties and overall costs of developing EBPs and the challenges to implementing them in the real world exactly as intended and evaluating their impact on the society, which are often large-scale and hard to measure.

There are also doubts about the actual impact that EBPs would have. It is possible that regardless of trainings and coaching sessions, practitioners will draw on their own experiences and accumulated knowledge anyway. Therefore, there may be a limit to preventing this artificially by providing strict restrictions on the kinds of decisions practitioners can make during clinical sessions.

In addition, EBPs would make a difference only if they were implemented strictly as provided by research and were updated constantly as researchers acquire additional information on the ways to further improve them. This is the central idea of EBPs—that treatments are continuously developed through the flow of knowledge within the research, practice, and policy communities, further supported by client contributions. However, with most public health interventions, there is great variability in delivery, training, population characteristics and responses to treatments, and the availability of resources according to locality (Victora, Habicht, & Bryce, 2004). Therefore, it may be likely that the dosage of interventions will be watered down in real-world settings. In this case, EBPs may not attain the level of effectiveness and social impact for the additional costs spent on pilot studies, rigorous and systematic research, or practitioner training and coaching.
Therefore, it is dire that all steps in the process--from development to implementation to evaluation--are successfully undertaken. One missing link in this chain may stifle other advances (Gambrill, 2006). For this reason, evaluations are also critical in that they provide detailed information on both delivery and compliance, which can later be used to improve services.

**Exchange Club’s Family Center in Durham**

Established in 1992 by the members of Exchange Clubs in Durham and surrounding areas, the Exchange Club’s Family Center in Durham is a provider of services focused on preventing child maltreatment and neglect. It is committed to assisting struggling families through programs aimed largely at support and education. The Family Center’s practices are grounded in the belief that parents have the greatest influence on their children and that parenting plays a significant role in determining child outcomes.

For over 10 years after it was established, the Family Center helped families by teaching parents effective and positive parenting skills to eventually develop more nurturing relationships between parents and their children. Their practice did not follow a specific model developed by researchers, but they had incorporated the findings from their own research and knowledge acquired through professional experience in order to devise a helpful and effective program. First, the service providers visited clients’ homes and treated the entire family together in every session. They focused on building positive parent-child relationships. Second, they incorporated coaching into their treatments to follow up on the parenting skills education. They asked parents to demonstrate and practice the learned skills. This gave the clinicians a chance to point out immediately what mistakes the parents were making before any wrong habits could develop. Third, they evaluated their program in order to assess if they were providing effective and satisfactory services. When families completed their treatment,
the parents were given two self-report assessments. One was the Adult-Adolescent Parenting Inventory (AAPI), a measure of parenting attitudes, proven to be a good predictor of child maltreatment. The other form was a satisfaction survey.

Over the years that they served families using this program, the Family Center received positive feedback on the satisfaction surveys and the AAPI outcomes indicated that the Family Center’s treatment was reducing parent attitudes predictive of child maltreatment. They have data showing that they made a statistically significant difference in reducing the risk of child maltreatment for the families who completed their program, which is over 65% of the families who were served by the Family Center.

However, when evidence-based practices began to gain widespread popularity in the area, the Family Center decided to switch to more structured models with scientific evidence bases. Parent-Child Interaction Therapy (PCIT) is the model they selected for families with children of ages 2-12. The following subsection describes PCIT in greater depth and explains its different components.

**PCIT**

PCIT was originally developed to treat children with oppositional behavioral problems by focusing on the parent-child relationship. As more studies were conducted, PCIT was found to have significant influence on parental attitudes toward children and their styles of discipline and attachment. This reduced negative parental influences on the children and taught and reinforced positive parenting methods, thereby ultimately improving children’s behavioral problems. PCIT teaches parents with conduct-disordered children coping skills, punishment tactics and strategies, and ways to reduce disruptive behaviors in their children through positive reinforcement rather than power assertion. It provides the parents with the language for developing more positive relationships with their children, such as praises and
criticisms (Thomas & Zimmer-Gembeck, 2011). There are more than 100 randomized controlled trials that support the effectiveness of PCIT. These evaluations have also shown that the effects of PCIT and the positive gains from treatment are maintained for up to six years (Pearl, Thieken, Olafson, Boat, Connelly, Barnes, & Putnam, 2011).

Many studies have also examined PCIT as a program for the prevention of child maltreatment. For example, Thomas and Zimmer-Gembeck (2011) studied the effects of PCIT through a randomized controlled trial on mothers with a history of child maltreatment. When mothers who had received 12 weeks of PCIT were compared with the waitlisted mothers, it was found that PCIT improved parent-child interactions, reduced child abuse potential, and increased maternal sensitivity. Overall, PCIT developed a more nurturing response to child needs in mothers who had high risks of abuse. In addition, those who completed the treatment—where completion was defined by successful acquisition of skills and readiness to terminate treatment—were less likely to be reported to child welfare for suspected maltreatment. Another study, which looked into the effectiveness of PCIT with families of preschool-age children with oppositional defiant disorder, found that parents who received the treatment reported decreased parenting stress and more control. This change in parent attitudes also led to statistically significant improvements in children's behavior (Schuhmann, Foote, Eyberg, Boggs, & Algina, 1998).

There are three critical implementation components identified through the research on PCIT (Herschell, Calzada, Eyberg, & McNeil, 2002). First, it is crucial that children and parents are treated together. PCIT is divided into two modules. The child-directed interaction focuses on changing the relationship between parents and children. Parents are taught nondirective play skills. The parent-directed interaction focuses on teaching parents clear, age-appropriate instructions. For example, they learn praises and ways to utilize time-out as a punishment method (Schuhmann et al., 1998). Second, coaching parents to use the acquired
skills and attain perfect mastery significantly increases effectiveness. PCIT utilizes live coaching and bug-in-the-ear wireless communication for immediate clinician feedback within the session. Usually in the beginning, clinicians praise and provide positive feedback in order to encourage parents. Later, they provide more noncritical corrections and coaching becomes more directive. Finally, assessment is a critical element, which requires the use of various instruments and methods. Clinical interviews and in-session observations provide some information about the progress that the families have made. Satisfaction surveys offer information about the quality of services. Evaluation is important because it ensures that PCIT is implemented to achieve maximum effectiveness (Herschell et al., 2002).

Researchers have attempted to study the impact of client characteristics, such as race, ethnicity, and age, on the effectiveness of PCIT, which is used with diverse populations. Cultural adaptations for Mexican Americans and Native American families have been developed in order to better meet the needs of the client families. A study on this variation of PCIT has surprisingly found that the translated version of PCIT achieves treatment effects similar to the culturally adapted version called the GANA program (McCabe, Yeh, Garland, Lau, & Chavez, 2005). A study was conducted to find the effectiveness of PCIT for 75 foster parents and their foster children compared with 98 non-abusive biological parent-child dyads treated because of the children’s behavior problems (Timmer, Urquiza, & Zebell, 2005). It showed decreases in child behavioral problems and caregiver distress from pre- to post-treatment for both foster and biological parent-children dyads. The findings suggested that PCIT is a promising therapy for children living with foster parents.

Researchers also have looked into the impact of modifying certain elements of PCIT. One example is a randomized controlled study of Enhanced Parent-Child Interaction Therapy (EPCIT), which incorporates more individualized services with additional intervention components based on needs (Chaffin, Silovsky, Funderburk, Valle, Brestan, & Balachova,
In this study, clinicians implemented the treatment in the homes of clients. They paid special attention to parental depression, substance abuse, domestic violence, and other problems, factors that have been shown to increase the likelihood of child maltreatment and eventually cause or worsen behavioral problems in children. The original PCIT model did not treat these internal problems and focused on observable behavioral changes.

The result revealed that EPCIT actually led to less behavior changes in clients than the traditional model. The study did note that the providers of PCIT were well-trained to carry out the treatment than were the staff treating client families with EPCIT. The researchers concluded that despite the added components, the modified version of PCIT does not necessarily have better outcomes. One possible explanation is that PCIT typically focuses on behavioral modification for immediate solutions to the problem of child oppositional behavior, which can improve parent-child relationships and thus reduce maltreatment. By attempting to fix the parents’ psychological issues or deeper personal struggles during the treatment, ECPIT may have opened up more long-term treatment needs and consequently diluted the immediate behavioral changes.

This study also suggests that modifications to the original PCIT may have unexpected influence on its effectiveness. The Family Center administers PCIT in clients’ homes rather than in clinical settings. The next subsection discusses the possible impact of this alteration.

**Home-based (In-Home) PCIT**

PCIT was originally designed to be administered in clinics or controlled settings using a one-way mirror with a bug-in-the-ear for parent coaching. However, in reality, this is not feasible for all clinics, nor is it practical for all client populations. Researchers have investigated the impact of in-home PCIT, or home-based PCIT. This treatment follows
standard practice of PCIT, but all the sessions are conducted in clients’ homes. Some modifications are necessary due to the lack of resources and the therapist’s unfamiliarity with the environment. These include eliminating all disturbances, such as TVs or computers, from the room selected for the treatment session, finding effective ways to provide transitional instructions to parents from inside the room, and developing lists of toys that the children favor for every household.

One study conducted in 2008 (Masse & McNeil, 2008) found that there are some advantages to this modified version of the treatment, as well as some downsides that may critically impact treatment. First, with the home-based PCIT, clinicians can more precisely individualize PCIT based on a client's living situation and be mindful of how the acquired parenting skills will be applied within the house. They can pay closer attention to logistical details like placing of the timeout chair. Second, in-home PCIT is less likely to cause anxiety in children because of their familiarity with the environment. Third, high attrition rates are a problem that providers of PCIT often face. The in-home treatment minimizes issues of client no-shows or irregular visits.

However, despite these positive aspects of home-based PCIT, practitioners have to be careful that any inevitable variations made to PCIT do not water down the impact of treatment. First, practitioners administering PCIT in clients’ homes are unfamiliar with the environment and therefore have little control over it. For example, they do not know of all the possible escape paths for children during their timeout sessions. There may also be other distractions they are not aware of, such as computer or television, which would diminish the effectiveness of the treatment. Second, clients’ homes are not equipped with one-way mirrors or bug-in-ear devices, and therefore practitioners must coach parents in the same room as the parent and the child. Consequently, the presence of the practitioner in the room induces children to behave and follow directions.
However, coaching in the house does have its downsides. Because children can hear the clinicians' instructions for their parents, they are more likely to do as they are told and behave. Some parents reported that children who behaved and followed directions well during clinical sessions did not follow the same instructions when the therapist left the house. Third, unanticipated disruptions such as phone calls or neighbor's visits may occur. Fourth, PCIT is a coached play model and utilizes toys and other activities into the treatment. The amount of resources that clinicians can carry to individual households is limited. It may require a few initial trials to determine which play tools to bring to each session.

The effectiveness of this home-based, or in-home, PCIT model was explored through a case study. It was found that parents can effectively acquire the skills, such as labeled praises, reflections, and behavior descriptions through this modified version of the treatment. However, data showed that caregivers' reported parenting stress did not decrease significantly (Ware et al., 2008). It should be noted that case studies often do not provide generalizable findings, since they are not as rigorous or replicable as randomized controlled tests. However, more systematic study into this PCIT model would greatly benefit future clients and practitioners.

**Family Center’s Implementation of PCIT**

There are several noteworthy aspects about the Family Center’s practice of PCIT. First, the Family Center serves its families at their homes and rarely have the clients come into the Center for a session. However, unlike the in-home PCIT described above, the Family Center does use walkie-talkies instead of the bug-in-ear devices. Before the sessions begin, providers also warn children that they will not respond to any questions or comments directed at them. They also request beforehand that families do not allow for any distractions or interruptions during the clinical sessions. Second, the Family Center has
frequently had trained interns treat its clients. Although these temporary workers are trained before they begin to serve families and are observed and coached for several rounds of sessions, they lack years of experience. This may compromise the impact or the quality of service because experience allows practitioners to make the right clinical decisions and implement treatments effectively. The Family Center has made an effort to assess their practitioners and give feedback on their performance. The practitioners are regularly observed by more experienced members, who fill out fidelity checklists to help them recognize whether they are properly implementing the treatment. These checklists are provided to the staff after each observed session for their reference.

The Family Center is committed to serving its clients with approved methods and finding what works best. The providers are concerned about the quality of their treatment and would like to know if the switch to PCIT, an EBP, has made a difference in serving their clients. They are especially interested in finding out if their services are having varying degrees of impact on families of different races.

In this study, I compared families who received PCIT to the families who were in the Nurturing Program for treatment impact on parents, based on two scales used by the Family Center to measure the therapy outcome, the Adult-Adolescent Parenting Inventory (AAPI) and satisfaction surveys. These measures will be discussed in a later section. The purpose of this study is to find out whether the use of PCIT can eventually make a tangible difference. I hypothesized the following:

1. Compared to parents who received the comparison treatment, parents who received PCIT will show greater improvement on the 4 AAPI subscales.
2. PCIT serves clients of different races equally well, and improvements on AAPI scores pre- and post-treatment do not vary significantly from group to group. On the other hand, the comparison intervention results in better outcomes for some race
groups than others.

3. PCIT has better or equally good levels of satisfaction from clients as the comparison treatment.

4. Different practitioners with differing levels of familiarity with PCIT and experience will affect clients’ AAPI outcomes and levels of satisfaction.

Methods

Description of Data

The Family Center provided a set of data compiled from the year 2000 to September 2011. For recent clients, only those who had closure codes (whose cases had been closed prior to September 2011) were included in the analysis. These data included the AAPI subscale scores and the satisfaction surveys results, along with the demographic information of clients, or parent-child dyads.

Subjects

There were a total of 326 client families who had children between the ages of 2 and 12. Client families were divided into the PCIT group and the comparison group based on the type of treatment they received. The comparison group had 117 client families, and the PCIT group had 211 client families. From 2000 to December 2006, the Family Center only offered the comparison treatment for families referred to its Family Support program for children's behavioral problems or other concerns. The comparison treatment involved in-home services but did not involve the use of standardized treatment manuals. Services were tailored to each family and lacked a consistent structure. After January 2007, the Family Support program began to offer PCIT. For a short transition period, some practitioners administered PCIT and some provided service as usual. For families who have received multiple rounds of treatments,
only their first round was considered in these analyses to limit the effect of previous treatments at the Family Center on these clients.

Families were excluded from the subject pool based on three criteria. First, a family was considered to be an inappropriate case if their child was too old for treatments provided in the Family Support Program. These families were then referred to the Parenting of Adolescents program at the Family Center. Second, in some instances, if the parent had substance abuse problems or other issues that inhibited them from participating and fully engaging in the therapy sessions, or the household blatantly had domestic violence problems, these cases were considered inappropriate. Third, parents with no direct contact with their children (e.g., cases where children had been removed from the home) were considered inappropriate for Family Center services. The race composition of the subjects was 55.4% African American, 33.0% Latino, and 11.6% Caucasian.

Practitioners

The total number of practitioners who provided the comparison treatment was 11 and the total number of practitioners who provided the PCIT was 10. There were 17 practitioners in total, which means that 4 practitioners served both treatments and counted twice. However, it should be noted that only three providers served a significantly proportion of the comparison group. One experienced practitioner served 65%, and the other two providers each served 14% and 8% of the comparison group. On the other hand, PCIT was served by more providers and the proportion of the group served by each provider was relatively equal. Four providers each served 33%, 24%, 17%, and 7% of the PCIT group.
Measures

AAPI

The AAPI is a self-response survey intended for adult and adolescent parents (Bavolek & Keene, 1999). There are 40 questions in total. The Family Center has consistently used only the four subscales (35 items) in the original version of the AAPI. The items each belong to a different parenting construct, or subscale. Each of the five responses per question is assigned a numerical value of 1 to 5, which is the raw score to be added up. The raw scores are added up separately for the four (or five) subscales. Then, according to the person assessed, different norm tables are used to get the standard scores, ranging from 1 to 10 (sten scores). There are different norm tables for Form A and Form B, for male and female survey takers, and for adult parents, adolescent parents, or adolescent non-parents. If there are any missing responses, the full score can be estimated by obtaining the total score for the completed questions in the subscale, multiplying this score by the total number of questions in the subscale, and dividing by the number of items that were completed.

The AAPI asks about parenting attitudes and beliefs about children and their roles. This instrument is convenient to use in that it is shorter than other instruments which measure similar factors. It takes only about 20 minutes to administer. The AAPI is intended for adults and adolescents with reading levels of fifth grade or higher. If survey takers have a lower reading level, the survey can also be read aloud.

The scores on the original 4 constructs of the AAPI—(1) Inappropriate expectations of children, (2) Parental lack of empathy towards children’s needs, (3) Strong belief in the use of corporal punishment as a means of discipline, and (4) Reversing parent-child role responsibilities—are good predictors of child maltreatment and of children’s behavioral problems (Bavolek & Keene, 1999). Each construct is described below:

1. Inappropriate expectations of children (Inappropriate Expectations)
Ten items assess accuracy of parents’ perceptions of children’s skill and abilities. These items evaluate unrealistic expectations parents may possess stemming from lack of knowledge about children’s developmental or growth stages.

2. Parental lack of empathy towards children’s needs (Lack of Empathy)

Seven items assess parents’ level of empathy towards and understanding of their children’s expression of emotions and needs.

3. Strong belief in the use of corporal punishment (Physical Punishment)

Eleven items assess the value parents place on physical punishment as a means of discipline and their understanding of its impact on children.

4. Reversing parent-child role responsibilities (Role Reversal)

Seven items assess parents’ belief that children should adopt behaviors traditionally associated with parents, such as providing care and comfort.

A fifth subscale—oppressing children’s power and independence—has been added to AAPI-2, but the Exchange Club’s Family Center in Durham has not used this scale for consistency of measurement. Moreover, often times, their clients give too much power and independence to their children because they possess unrealistic expectations. Exchange Club staff believed that this subscale therefore would provide little information about the impact of PCIT on their clients.

Normative data for the AAPI come from fifty-three agencies from twenty-three different states, comprised of almost 1,000 adults and adolescents (Bavolek & Keene, R., 1999). The AAPI consists of two forms, Form A and Form B. They are meant to be used pre-intervention and post-intervention to measure the difference or change in scores. The two forms were tested for correlation to examine whether they may be used interchangeably. The correlation between the different subscales of Forms A and B ranged from 0.80 to 0.92. The
high correlation suggested that the order in which the forms were used would not influence the results.

A factor analysis on the AAPI concluded that the instrument measures at least two constructs that it intends to measure and therefore the total score would provide information of value (Conners et al., 2006). However, individual scores on the subscales must be interpreted with caution because according to the factor analysis, every subscale is measuring more than one construct. Since the Family Center utilizes the AAPI as an instrument to assess the pre- and post-treatment differences in the parental attitudes toward parenting and does not use it for diagnostic purposes, it was not considered problematic that each AAPI subscale may lack unidimensionality.

According to over 20 years of research, AAPI is a valid and reliable inventory that assesses parenting attitudes predictive of abusive behaviors (Bavolek & Keene, 1999). It significantly discriminates between adult and adolescent parents and non-parent adolescents with abusive attitudes and those with non-abusive attitudes. The AAPI is often used to determine the impact of parenting education or therapies aimed at changing the attitudes of parents toward their children. The scores are used also to predict parents' risk of child maltreatment or abuse. Higher scores on the survey indicate more sensitive and nurturing parenting attitude.

The Family Center gave the pre-treatment AAPI to clients within the first 2~3 sessions. The post-treatment AAPI was given during the last 2~3 sessions of a full cycle of treatment, which consists of 10 to 12 treatments. Practitioners tried to collect AAPI results for all their clients, whether they had completed a full cycle or had stopped participating for various reasons. However, in cases where families relocated without notice or refused for any reason to fill out the AAPI, scores could not be obtained.
Satisfaction Survey

The satisfaction survey—called Family Support Program Exit Questionnaire—asked four questions in total, three of which were of interest in this study. This survey was developed because the practitioners in this program felt it was important to know whether their clients believe that the services they receive are helpful and whether clients have respect and trust for their therapist and the work that he/she does with them. Out of the entire subject pool, a total of 228—150 for the PCIT group and 78 for the comparison group—had data for this survey.

The first of the three questions considered in this study asked clients what factors were most helpful to them. The responses available were: program helped with understanding of child's behavior, program helped with setting goals and making a plan to meet them, program help learn about local resources, program provided assistance working out family problems, and there was someone to talk to. The second question asked what the clients liked about working with the Family Support Specialist. The responses were: practitioner was knowledgeable, practitioner was respectful of the family's way of doing things, and meeting times were good. The third question asked about the reasons for discontinuing the program. The responses were: family moved, program was unhelpful, specialist was not knowledgeable, specialist was not respectful of family's way of doing things, client completed stated goals, and client was too busy and had other priorities. Each response for each question could be answered yes or no and was considered a categorical dependent variable. Clients’ answers were coded 0 for unchecked or 1 for checked response.

Referral Source

Clients in the Family Support Program were referred to the Family Center through different ways and by different entities. These referral sources were coded into 9 categories:
<table>
<thead>
<tr>
<th>Code</th>
<th>Referral Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Department of Social Services (DSS)</td>
</tr>
<tr>
<td>2</td>
<td>School</td>
</tr>
<tr>
<td>3</td>
<td>Self</td>
</tr>
<tr>
<td>4</td>
<td>Relatives</td>
</tr>
<tr>
<td>5</td>
<td>Mental Health Clinics or Centers</td>
</tr>
<tr>
<td>6</td>
<td>Other</td>
</tr>
<tr>
<td>7</td>
<td>Court Ordered</td>
</tr>
<tr>
<td>8</td>
<td>Children's Developmental Services Agency</td>
</tr>
<tr>
<td>9</td>
<td>Durham County Health Department</td>
</tr>
</tbody>
</table>

**Closure Code**

For this study, only families whose cases had been closed were included into the subject pool. Therefore, these families were given one of the four closure codes upon the termination of their treatment. The codes are: 1 – Completed goals, 2 – Client stopped participating, 3 – Client refused to participate, and 4 – Family Relocated.

Every client in the Family Support Program upon the start of their treatment had to set one concrete, 30-day goal for themselves through a conversation with their practitioner. Some examples of these goals are getting the child to say “no” to the parent less or reducing the frequency of a bad behavior, such as hitting or throwing toys. For families that attended at least 10 sessions and reached their goals prior to discontinuation of services, the first code, “1- Completed Goals,” was given.

The next two, “2 - Client Stopped Participating” and “3 – Client Refused to Participate,” can be differentiated by the level of communication involved in the process of stopping treatment. In cases where families told the practitioners that they refuse to have any more sessions or they did not come to the door when the practitioners visited their homes for a session, even though they could be seen through the window, the clients were given the “3 – Client Refused to Participate” code. Code “2 - Client Stopped Participating” was used
primarily when families disappeared and could not be contacted or tracked down, but providers had no knowledge of relocation. When families moved out of the Family Center’s jurisdiction, their cases were dropped and given the “4 – Family Relocated” code.

**Mandatory Referral**

For the analysis, it was assumed that all families registered with the DSS were to some degree mandated or strongly recommended to receive treatment. Therefore, they were classified as “mandated clients.” The other group of mandated clients was the court-ordered families. Combined, the mandated clients were given the code “1.” Clients that were not mandated were given the code “0.”

**Data Analysis**

A general linear model analysis was conducted to test the effects of treatment type on post-treatment AAPI scores, controlling for baseline AAPI scores, race, interaction of treatment type by race, number of sessions, and mandatory referral. Baseline AAPI scores were included in order to remove any bias resulting from different baseline scores since they were likely to be correlated with post-treatment scores. This analysis tested the first hypothesis—that subjects who received PCIT showed greater improvement than comparison groups—when the effect of other covariates are taken out.

There was no significant difference ($\chi^2 = 2.00, p = 0.16$) in the completion rates of the AAPI surveys: 63.48% of the comparison group and 71.09% of the PCIT group completed them. This meant that any difference in AAPI scores due to treatment type was not confounded by the difference in AAPI completion rates.
Covariates

Four covariates were included: race, the interaction of race by treatment type, number of sessions, and mandatory referral. T-tests and chi squares were tested for many variables to see which factors had significant differences in the two treatment groups, PCIT and comparison. The variables considered were: race, completion of stated goals, referral sources, parent gender, parent age, mandatory referral, number of children in the family, and number of sessions. Variables with significant differences between the two treatment groups were included in analyses as covariates.

Race

A significant difference existed in the racial make-up in the two treatment groups ($\chi^2 = 13.76, \ p = 0.001$). Within the PCIT group, the breakdown by race was 48.3% African American, 12.0% Caucasian, and 39.7% Latino. In the comparison group, the race composition was 68.8% African American, 11.0% Caucasian, and 20.18% Latino. Over time, the Family Center served more Latino families, consistent with the change in Durham’s population. Because of this shift in demographics, greater percent of Latino families were provided PCIT compared to the other two races. The interaction of treatment type by race was included as an additional covariate in order to see whether the benefit of one treatment over the other differs by race.

Number of Sessions

The t-test analysis showed that number of sessions differed significantly by treatment type ($p = 0.008$). The PCIT group received a mean of 14.28 sessions, and the comparison group received a mean of 11.77 sessions. Therefore, number of sessions was included as a covariate.
**Mandatory Referral**

The percent of mandated clients in the comparison group (53.04%) was significantly greater (p = 0.0025) than the percent of mandated clients in the PCIT group (35.71%). Therefore, mandatory referral was included as a covariate in the analyses of treatment differences on pre- and post-treatment AAPI subscale scores.

**Satisfaction Survey Results**

Two types of data analysis were conducted on the satisfaction survey results. First, chi-square tests of independence were performed to compare satisfaction survey results of the PCIT and the comparison group. Second, a general linear model procedure was conducted on each response option of three questions on the exit questionnaire. Treatment type was the independent variable; race, mandatory referral, number of sessions, and the interaction of treatment type by race were used as covariates. This analysis tested whether the Family Center served clients better (from the parent perspective) with PCIT or the comparison treatment.

**Results**

**General Linear Model Analysis of the AAPI Results**

The results of the general linear model analysis of the AAPI post-treatment scores, including the F-statistic, degrees of freedom, and p-values, are reported in Table 1. Predictors include the main independent variable, treatment type, and five covariates. The linear relationship of each variable to AAPI post-treatment scores was examined, controlling for all other variables.
Table 1: Results from a general linear model analysis of AAPI post-treatment scores

<table>
<thead>
<tr>
<th>Variable</th>
<th>F-statistic</th>
<th>DF</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Inappropriate Expectations (IE) post-treatment score</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full Model *</td>
<td>6.27</td>
<td>8, 188</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Predictors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IE baseline score *</td>
<td>13.75</td>
<td>1, 188</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Treatment type</td>
<td>0.38</td>
<td>1, 188</td>
<td>0.73</td>
</tr>
<tr>
<td>Race *</td>
<td>13.44</td>
<td>2, 188</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Treatment type by Race *</td>
<td>3.68</td>
<td>2, 188</td>
<td>0.03</td>
</tr>
<tr>
<td>Number of sessions</td>
<td>0</td>
<td>1, 188</td>
<td>0.93</td>
</tr>
<tr>
<td>Mandatory referrals</td>
<td>5.65</td>
<td>1, 188</td>
<td>0.52</td>
</tr>
<tr>
<td><strong>Lack of Empathy (LE) post-treatment score</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full Model *</td>
<td>13.5</td>
<td>8, 188</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Predictors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LE baseline score *</td>
<td>61.64</td>
<td>1, 188</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Treatment type</td>
<td>5.17</td>
<td>1, 188</td>
<td>0.02</td>
</tr>
<tr>
<td>Race *</td>
<td>3.79</td>
<td>2, 188</td>
<td>0.02</td>
</tr>
<tr>
<td>Treatment type by Race *</td>
<td>4.25</td>
<td>2, 188</td>
<td>0.02</td>
</tr>
<tr>
<td>Number of sessions</td>
<td>1.42</td>
<td>1, 188</td>
<td>0.23</td>
</tr>
<tr>
<td>Mandatory referrals</td>
<td>0.09</td>
<td>1, 188</td>
<td>0.77</td>
</tr>
<tr>
<td><strong>Physical Punishment (PP) post-treatment score</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full Model *</td>
<td>16.2</td>
<td>8, 188</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Predictors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP baseline score *</td>
<td>102.98</td>
<td>1, 188</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Treatment type</td>
<td>0.61</td>
<td>1, 188</td>
<td>0.44</td>
</tr>
<tr>
<td>Race *</td>
<td>12.16</td>
<td>2, 188</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Treatment type by Race</td>
<td>0.56</td>
<td>2, 188</td>
<td>0.57</td>
</tr>
<tr>
<td>Number of sessions</td>
<td>0.39</td>
<td>1, 188</td>
<td>0.53</td>
</tr>
<tr>
<td>Mandatory referrals</td>
<td>0.17</td>
<td>1, 188</td>
<td>0.68</td>
</tr>
<tr>
<td><strong>Role Reversal (RR) post-treatment score</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full Model *</td>
<td>21.28</td>
<td>8, 187</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Predictors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RR baseline score *</td>
<td>85.17</td>
<td>1, 188</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Treatment type</td>
<td>1.11</td>
<td>1, 188</td>
<td>0.29</td>
</tr>
<tr>
<td>Race *</td>
<td>10.41</td>
<td>2, 188</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Treatment type by Race</td>
<td>1.09</td>
<td>2, 188</td>
<td>0.34</td>
</tr>
<tr>
<td>Number of sessions</td>
<td>2.25</td>
<td>1, 188</td>
<td>0.14</td>
</tr>
<tr>
<td>Mandatory referrals *</td>
<td>4.98</td>
<td>1, 188</td>
<td>0.03</td>
</tr>
</tbody>
</table>

* p < 0.05
Inappropriate Expectations

For the Inappropriate Expectations subscale, the overall model was significant. Controlling for all the covariates, treatment type did not significantly predict post-treatment Inappropriate Expectations; however, race (p < 0.0001) and the interaction of treatment type by race (p = .03) were significant predictors. As shown in Table 2, the mean post-treatment score was highest for Latino families, meaning that these families had the most appropriate expectations for their children.

Table 2: Mean Baseline and Post-treatment Subscale Scores by Race

<table>
<thead>
<tr>
<th></th>
<th>Inappropriate Expectations mean</th>
<th>Lack of Empathy mean</th>
<th>Physical Punishment mean</th>
<th>Role Reversal mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Post</td>
<td>Pre</td>
<td>Post</td>
</tr>
<tr>
<td>African American (n = 104)</td>
<td>4.41</td>
<td>4.63</td>
<td>4.06</td>
<td>4.88</td>
</tr>
<tr>
<td>Caucasian (n = 31)</td>
<td>4.58</td>
<td>5.13</td>
<td>4.97</td>
<td>5.94</td>
</tr>
<tr>
<td>Latino (n = 78)</td>
<td>5.15</td>
<td>5.90</td>
<td>5.78</td>
<td>6.00</td>
</tr>
</tbody>
</table>
African American clients (+0.89) and Caucasian (+0.57) clients showed larger improvements with PCIT than with the comparison treatment, as depicted in Graph 1. On the other hand, Latino families improved 1.01 points less with PCIT than with the comparison intervention.
Lack of Empathy

For the Lack of Empathy subscale, the overall model was significant ($F (6, 190) = 16.03, p < 0.0001$). For the Lack of Empathy subscale post-treatment score, baseline score ($p < 0.0001$), race ($p = 0.02$), treatment type ($p = 0.02$) and the interaction of treatment type by race ($p = 0.02$) were significant. As shown in Table 2, the post-treatment scores for this subscale were highest for Latino subjects and lowest for African American subjects. Also, the scores were higher for the comparison group. Comparison of the average score change from baseline to post-treatment between the comparison group and the PCIT groups revealed Caucasian (-1.32) and Latino (-0.70) clients showed far less improvement on this subscale with PCIT than they did with comparison treatment (see Graph 2). In contrast, empathy changes for African American clients were similar for the two treatment types.

<table>
<thead>
<tr>
<th></th>
<th>African American</th>
<th>Caucasian</th>
<th>Latino</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparison</td>
<td>0.76</td>
<td>1.82</td>
<td>0.79</td>
</tr>
<tr>
<td>PCIT</td>
<td>0.86</td>
<td>0.5</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Graph 2: AAPI Lack of Empathy subscale: Comparison group vs PCIT group by Race
Physical Punishment

For the Physical Punishment subscale, the overall model was significant (F (6, 190) = 21.51, p < 0.0001). Controlling for covariates, the baseline score for the Physical Punishment subscale (p < 0.0001) and race (p < 0.0001) showed significant linear relation with the post-treatment AAPI score. Post-treatment scores were higher for participants with higher baseline scores. As shown in Table 2, Latino clients had higher mean post-treatment score than Caucasian clients, Caucasian clients higher than African American clients. Controlling for all covariates, treatment type did not significantly predict post-treatment Physical Punishment AAPI scores, either as a main effect or in interaction with race. Graph 3 depicts the score increases or decreases (pre-intervention to post-intervention) in the Physical Punishment subscale for the two treatment types for each race group. It should be noted that the interaction of treatment type by race was not a significant a covariate for this subscale.
Role Reversal

For the Role Reversal subscale, the overall model was significant ($F (6, 189) = 27.98$, $p < .0001$). Baseline score ($p < 0.0001$), mandatory referral ($p = 0.03$) and race ($p < .0001$) showed significant linear relation with the post-treatment score. Post-treatment scores were higher for participants with higher baseline scores and for clients who were mandated. The Role Reversal subscale scores for clients who were not mandated improved 0.74 on average, whereas clients who were mandated improved by 1.10, as shown in Table 3. Controlling for all covariates, treatment type did not significantly predict post-treatment Role Reversal scores.

Table 3: Role reversal pre/post-treatment average scores by mandatory referral

<table>
<thead>
<tr>
<th></th>
<th>Role reversal baseline mean</th>
<th>Role reversal post-treatment mean</th>
<th>Role reversal average score change (post-pre)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not mandated = 0</td>
<td>5.12</td>
<td>5.86</td>
<td>0.74</td>
</tr>
<tr>
<td>Mandated =1</td>
<td>4.39</td>
<td>5.45</td>
<td>1.10</td>
</tr>
</tbody>
</table>
Graph 4 depicts the score increases or decreases (pre-intervention to post-intervention) in the Role Reversal subscale for the two treatment types for each race group. It should be noted that the interaction of treatment type by race was not a significant covariate for this subscale.

<table>
<thead>
<tr>
<th></th>
<th>African American</th>
<th>Caucasian</th>
<th>Latino</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparison</td>
<td>0.67</td>
<td>-0.18</td>
<td>1.36</td>
</tr>
<tr>
<td>PCIT</td>
<td>0.59</td>
<td>0.68</td>
<td>1.48</td>
</tr>
</tbody>
</table>
General Linear Model Analysis of the Satisfaction Survey Results

The first question asked clients what factors were most helpful to them. For 4 out of 5 responses, the overall model was significant: understanding of child's behavior \( (F(7, 201) = 2.88, p=0.01) \), setting goals and making a plan to meet them \( (F(7, 201) = 2.80, p = 0.01) \), learning about local resources \( (F(7, 201) = 3.57, p=.001) \), and assistance working out family problems \( (F(7, 200) = 2.83, p = 0.01) \). As shown in Table 4, treatment type was a significant factor for the response “understanding of child's behavior” \( (p = 0.02) \) and the response “learning about local resources” \( (p = 0.04) \) when adjusted for all other covariates. The PCIT group had better satisfaction scores for “understanding of child's behavior” than the comparison group. The comparison group on the other hand had better satisfaction scores for “learning about local resources.”

<table>
<thead>
<tr>
<th>Variable</th>
<th>F-statistic</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment helped understanding of child's behavior</td>
<td>( F(1, 201) = 5.59 * )</td>
<td></td>
</tr>
<tr>
<td>Comparison</td>
<td>74.36</td>
<td></td>
</tr>
<tr>
<td>PCIT</td>
<td>93.38</td>
<td></td>
</tr>
<tr>
<td>Treatment helped learn about local resources</td>
<td>( F(1, 201) = 4.09 * )</td>
<td></td>
</tr>
<tr>
<td>Comparison</td>
<td>65.38</td>
<td></td>
</tr>
<tr>
<td>PCIT</td>
<td>53.64</td>
<td></td>
</tr>
<tr>
<td>Provider was knowledgeable</td>
<td>( F(1, 201) = 4.78 * )</td>
<td></td>
</tr>
<tr>
<td>Comparison</td>
<td>83.33</td>
<td></td>
</tr>
<tr>
<td>PCIT</td>
<td>94.70</td>
<td></td>
</tr>
<tr>
<td>Provider respected the family’s way of doing things</td>
<td>( F(1, 201) = 6.08 * )</td>
<td></td>
</tr>
<tr>
<td>Comparison</td>
<td>73.08</td>
<td></td>
</tr>
<tr>
<td>PCIT</td>
<td>90.73</td>
<td></td>
</tr>
</tbody>
</table>

* \( p < 0.05 \)

The second question asked what the clients liked about working with the Family Support Specialist. For all three responses, the overall linear model was significant:

39
knowledgeable (F (7, 201) = 2.67, p=0.01), respectful of the family's way of doing things (F (7, 201) = 2.17, p=0.04), and good meeting times (F (7, 201) = 3.10, p=0.004). As shown in Table 3, treatment type was a significant variable for the response "knowledgeable” (p=.03) and the response "respectful of family's way of doing things” (p = 0.01), when controlling for all covariates. The PCIT group had better satisfaction scores than the comparison group on both responses.

For the third question that asked about the reasons for discontinuing the program, the overall model was significant for the responses "completed goals" (F (7, 201) = 7.24, p<0.0001) and "too busy/other priorities" (F (7, 201) = 4.32, p = 0.0002). Treatment type was not a significant variable for any of the responses when controlling for all other covariates. One additional thing to note about covariates other than treatment type is that mandatory referral (p=0.0002) was shown to be significant for the response “too busy/other priorities.”
Analysis of Provider Effect on Outcomes

The AAPI outcomes did not vary significantly by provider for either treatment type, perhaps given the large variability among providers for each treatment. Separate analyses were run for one experienced provider who served the largest number of clients and provided both types of interventions. This provider served 65% of the comparison group and 7% of the PCIT group, which meant that she/he had less experience serving PCIT in terms of the number of clients. The average score change from pre-treatment to post-treatment on each AAPI subscale was greater for the PCIT clients than for the comparison group (see Table 5). Controlling for covariates, this difference approached significance for role reversal ($F(7,42) = 3.04, p = .09$), but was not significant for any of the other AAPI outcomes.

Table 5: Average Score Change on the AAPI Subscales

<table>
<thead>
<tr>
<th></th>
<th>Inappropriate Expectations average score changes (post-pre)</th>
<th>Lack of Empathy average score changes (post-pre)</th>
<th>Physical Punishment average score changes (post-pre)</th>
<th>Role Reversal average score changes (post-pre)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCIT</td>
<td>2.4</td>
<td>2.6</td>
<td>1.6</td>
<td>1.6</td>
</tr>
<tr>
<td>Comparison</td>
<td>0.5</td>
<td>1.1</td>
<td>0.6</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Discussion

This study’s primary goal was to compare PCIT with treatments as usual by comparing AAPI post-treatment scores of clients who had received either type of treatment at the Exchange Club’s Family Center in Durham over the last 11 years. The study further examined the effect of race on the levels of improvement in each AAPI subscale to observe any differences by race in the PCIT outcomes and racial differences in the outcomes for treatment as usual, or the comparison treatment. The study also analyzed satisfaction survey
responses to find differences in the way people felt about the services they received based on treatment type. Lastly, the study looked at provider influences on the outcomes of intervention. The follow discuss the findings about these four areas of inquiry.

**Interpretation of AAPI Outcomes**

The first hypothesis for this study was that compared to parents who received the comparison treatment, parents who received PCIT will show greater improvement on the 4 AAPI subscales. In general, the AAPI baseline, or pre-treatment, scores and race were significant predictors of the post-treatment scores. Post-treatment scores were higher for people with higher baseline scores. Overall, regardless of treatment type, Latino families had the highest baseline post-treatment scores on every subscale. Treatment type on its own was only significant for the Lack of Empathy subscale. Interestingly, the comparison group experienced greater score improvements than the PCIT group. For the Inappropriate Expectations subscale, the interaction of treatment type by race was significant. This meant that, on average, treatment type had different effects for clients of different races. Specifically, Caucasian and Latino clients showed greater change in the direction of learning appropriate expectations when they participated in PCIT versus the comparison intervention. In contrast, Latinos did better on the Inappropriate Expectations subscale with the comparison treatment. For Physical Punishment and Role Reversal, treatment type did not seem to make a significant difference in effect, even when considering how different racial groups were affected by different treatment types. However, it should be noted that Caucasian clients were the only race group that did see about 1 point improvement from PCIT on each subscale, whereas they did not achieve any improvements from the comparison treatment. The following subsections will describe the outcomes on the 4 subscales of the AAPI in greater detail.
Inappropriate Expectations

In the Inappropriate Expectations subscale, Latino families achieved less improvement from PCIT than from the comparison treatment. On the other hand, African American and Caucasian clients benefited from receiving PCIT. In the case of African American families who received the comparison treatment, the scores in this subscale, on average, decreased from pre-treatment to post-treatment. The scores improved with the African American clients in the PCIT group. This may partly be explained by the design of PCIT. It is intended to teach parents what to expect from their children at their age level. This may not have happened for the Latino clients for many reasons. Latino families had higher baseline scores than other racial groups. In addition, it is also possible that because the Latin American culture is highly relationship-based, PCIT, a behavioral modification therapy, may not have been a good cultural fit.

Therefore, it may be concluded with caution that PCIT seems to help parents of some races who have little understanding of age-appropriate child gain significantly greater understanding and knowledge about how children normally act at their level of development. Parents with inappropriate expectations of children, who lack understanding of age-appropriate child behavior, can cause children to feel debilitated and worthless (Bavolek & Keene, 1999). By teaching parents how to respond to children’s misbehavior and coaching them to vocally reflect on children’s behavior, PCIT may be allowing parents to understand their children from their perspectives.

Lack of Empathy

The rationale behind the Lack of Empathy subscale was that empathetic parents are sensitive towards their children’s emotions and demands and children thereby learn to be empathetic towards others. On the other hand, children who lack empathy are not cooperative
or attentive to others’ feelings and fail to develop a solid set of morals and values (Bavolek & Keene, 1999). In this subscale, Caucasian and Latino families achieved greater pre- to post-treatment score increases from the comparison treatment than from PCIT. For African Americans, there was little difference between the two treatments. Therefore, overall, PCIT was not as effective as the comparison treatment in teaching parents to be more empathetic or sensitive to their children’s emotions and needs. PCIT is aimed at teaching parents skills to cope with their children’s behavioral disorders and respond to them in ways that are consistent. It does not teach them to recognize their children’s behavioral signs or understand the reasons for their children’s behavior. Therefore, PCIT may have shown less improvement in this realm for this reason.

**Physical Punishment**

In the Physical Punishment subscale, only Caucasian families benefited more from PCIT than from the comparison treatment. The pre- to post-treatment score increase went from 0 to 1, which meant that Caucasian parents’ reliance on physical punishment for discipline did not change at all through the comparison treatment, but PCIT was able to bring up the scores substantially. The other two racial groups benefited equally from both types of treatment and therefore the interaction of treatment type by race was not significant. Abusive parents tend to also resort to corporal punishment, which is likely to teach the children to use violence for solving problems (Bavolek & Keene, 1999). Using more nurturing and less-violent measures of punishment is likely to prevent such violent behaviors in children. It may be that although PCIT attempts to change parents’ methods of discipline, it does not change their beliefs about physical punishment, any more effectively than does the comparison treatment. Nevertheless, this is not to say that PCIT will not result in any changes in beliefs about physical punishment. As the parents begin to respond in different ways to children’s
disordered behavior, they may begin to observe the benefits of the new discipline strategies. This may in turn result in attitude changes in the long run.

**Role Reversal**

In the Role Reversal subscale, only Caucasian families achieved greater improvement from PCIT than from the comparison treatment. For the other two racial groups, PCIT and the comparison treatment served families equally well, which may explain why the interaction of treatment type by race was not significant in the GLM model. Sometimes parents are overly dependent on their children and expect them to play the role of the parent. Children assume authority to make decisions instead of the parents. This leaves children feeling inadequate and at the same time does not allow them to develop at their age-appropriate levels. Research has shown that there is significant correlation between role reversal and developmental deficiencies in children (Bavolek & Keene, 1999).

From these findings, it can be concluded that the first hypothesis for this study was not supported by the results, for PCIT only resulted in significant improvement over the comparison treatment for the Inappropriate Expectations subscale. Even for this subscale, results for different racial groups were mixed. Only African American and Caucasian groups showed some minor increase in average improvements in the Inappropriate Expectations subscale from switching to PCIT, whereas Latino group experienced a drop in improvement. PCIT clients showed less improvement in the Lack of Empathy scale, with the exception of African American clients. For the Physical Punishment and Role Reversal subscales, the two treatments were roughly equivalent. Still, there were some racial differences in relative effectiveness.

The second hypothesis was also not supported. PCIT did not serve all racial groups equally, as improvements on the AAPI scores from pre-intervention to post-intervention
varied significantly from group to group. The comparison treatment also served some racial groups better than others. In addition to this finding, it was also seen that different racial groups responded better to different treatment types.

Latino families had higher post-treatment scores than the other two races on all AAPI subscales. However, this may not be the result of treatment effect. Latino families achieved less or almost equal improvement from PCIT than from the comparison treatment in all subscales. Therefore, PCIT may not be the best fit for Latino families in this sense. As mentioned briefly above, the Latin American culture places importance on relationships. Therefore, the comparison treatment, where the practitioners interacted more with clients directly rather than instruct and coach them using remote devices, may be a better fit for this group. Also, they may have different underlying concerns than other racial groups because of their immigrant status that individualized services could better provide.

With African American clients, the PCIT group and the comparison group had similar results, except for the Inappropriate Expectations subscale, in which PCIT served African American clients slightly better. However, even this group showed less improvement in the Physical Punishment subscale and in the Role Reversal subscale from the shift to PCIT (though differences were nonsignificant).

Caucasian clients experienced modestly greater improvement on the Inappropriate Expectations subscale and significantly less improvement on the Lack of Empathy subscale from the switch to PCIT. On the other hand, although the data analysis may not provide enough information, it can be said with caution that Caucasian clients seemed to benefit more from PCIT on the Physical Punishment and Role Reversal subscales than from the comparison treatment.
Limitations

One caveat to the interpretation of AAPI results provided here is that since clients and providers were not randomized, there may be other factors influencing this result. PCIT was administered by relatively diverse providers, with differing levels of expertise. A greater proportion of the PCIT clients were served by interns. On the other hand, the comparison treatment was mostly provided by an experienced practitioner. While this may hint at the possibility that PCIT may be more replicable than treatments without instruction manuals and standardized training courses—since even with the greater proportion of interns as providers for PCIT, there were no significant provider effects on AAPI outcomes—it may be that the low sample size and significant amount of provider variability masked effects in this area. In addition, because the information about fidelity of implementation was not provided, provider effect was solely analyzed using the number of days a provider had worked with the Family Center. This may be an inappropriate proxy variable for practitioner experience or fidelity of implementation.

There are some additional caveats. There were 109 missing AAPI data in total. The completion of pre- and post-treatment AAPI was not independent of the completion of stated goals ($p < 0.0001$). Of clients who completed state goals, 93.68% completed the AAPI, whereas 30.65% of the non-completers took both AAPIs. Treatment effects based on AAPI results may be overestimated because clients who completed stated goals and received greater number of sessions were more likely to complete AAPI surveys. People who completed the AAPIs ($n=206$) had received a mean of 15.80 sessions ($SD=0.66$), while people who did not complete the AAPI ($n=75$) had received a mean of 6.32 sessions ($SD=6.36$). However, the number of completed AAPI data was equal for PCIT and the comparison treatment, and the dropout rate was similar for the two treatments. Therefore, although one-third of the data were missing, the analysis is not likely to be biased toward either treatment type.
On other note, it may be that AAPI is not a good measure of PCIT effectiveness. PCIT may not be effective in changing parental attitudes toward parenting and child behavior and may seem to be ineffective according to the standards of the AAPI. However, by inducing behavioral changes in parents when coping with children with behavior issues, PCIT may still achieve its intended effect of helping children overcome these problems through positive parental influence.

**Interpretation of Satisfaction Survey Results**

The analysis of satisfaction survey results showed that more clients in the PCIT group felt that their provider respected the family’s ways of doing things and was knowledgeable. It may be that because PCIT, as with many evidence-based programs, has fixed components that must be consistently implemented and must follow a rigid structure, providers seem to be more knowledgeable and in control. If this is so, PCIT may be more effective than the comparison treatment because it allows providers to gain greater respect and trust of their clients.

**Conclusion**

As mentioned in the first few sections of the paper describing evidence-based practice and policy, there are many barriers to the broad implementation of EBPs. There are issues that arise from the rigor required of the evidence to support EBPs, the rigid regimen for implementation of practices, and political interests of elected policymakers and ensuing tension in this community. These challenges thwart the implementation of EBPs as widespread, standard practice. This paper specifically focuses one aspect of this overarching problem: barriers to reaching full effectiveness in carrying out EBPs in the real world.

Clients do not respond universally to treatments. Fidelity of implementation,
practitioner experience, and client characteristics and values are some of the factors that can influence outcomes. EBPs must be implemented correctly in order to have the anticipated impact, as tested through rigorously designed studies like RCTs. Good evaluations are also necessary because they allow for assessment of client progress, treatment effectiveness, and implementation fidelity.

One concern with EBP was that practitioners draw on their own experiences and accumulated knowledge regardless of strict implementation guidelines and training. Therefore, this study also tested provider impact on client outcomes. It also assessed whether clients were receptive to PCIT and believed that the treatment was helping them. Unfortunately, the fidelity question could not be answered due to a lack of data.

The central question that the study attempted to answer was whether the switch to an EBP made a significant difference in client outcomes. Results on evaluation measures used by the Family Center, specifically AAPI and two satisfaction, or exit, surveys, were analyzed to investigate this.

The main finding of this study was that while PCIT does significantly improve clients on all subscales of the AAPI for all race groups—whereas the comparison treatment, or treatment as usual, did not always see improvements in these scores—it does not show significantly greater improvement than the comparison treatment in all subscales of the AAPI. This may be because PCIT is intended to change parents’ behavioral responses, not attitudes or emotional responses, to their children’s oppositional conduct. Therefore, they learn what kind of child behavior to condone at their age levels and what actions to discourage using which parenting methods. This explains why Caucasian and African American clients especially experienced bigger leaps in the Inappropriate Expectations subscale of the AAPI through PCIT than the comparison treatment. However, the Latino group, which had a higher average baseline score than the other two race groups, did not achieve significantly greater
improvements through PCIT than through the comparison treatment. This may mean that PCIT is appropriate only for some racial groups with certain cultural characteristics. Based only on the AAPI results, Latino clients would benefit more from the comparison treatment or other treatments that involve more direct and personal interaction with providers. On the other hand, African American and Caucasian clients may be better served by PCIT, although the AAPI results did not provide sufficient evidence for making conclusive recommendations. In addition, the results also suggested the possibility that some subscales of the AAPI may not be a good measure for PCIT.

On the other hand, the satisfaction surveys offered valuable information about client responses to the two treatments. PCIT, as with other EBPs, has manuals that deliver clear and detailed instructions to all providers. Clients’ perception that providers were knowledgeable and respectful of their values and tradition were significantly greater with the PCIT group than with the comparison group. PCIT provides a strictly planned out schedule of clinical sessions with educational components that allow parents to understand the activity and follow-up coaching where parents are given individual attention. Therefore, parents were more likely to feel that the provider was organized in delivering the services and that they showed consideration for the families’ ways of doing things.

Based on these findings, PCIT can be said to have some benefits at the local agency level, though the analysis of AAPI results provided limited information about the benefit of PCIT relative to the comparison treatment. First, PCIT helped the clients of the Family Center gain more accurate knowledge about age-appropriate child behavior. Second, it significantly changed Caucasian parents’ belief in the use of physical punishments. It encouraged more nurturing ways of penalizing children for bad conduct. Third, more clients who received PCIT believed their provider was knowledgeable and respectful of their familial values and traditions than clients who received the comparison treatment.
The overview of the Family Center’s practice and the study findings about the effect of switching from treatment as usual to an EBP may offer some insight about the use of EBPs and evidence-based policymaking in public health and their possible consequences. The study has several lessons.

First, tests of fidelity are crucial to EBPs. The Family Center administered a variation of PCIT intended to be administered in home settings. In reality, clinics often may not be equipped with one-way mirrors or high technology devices which are costly. Because of the variations, it was difficult to extend the findings of the study, such as the assessment of the effectiveness of PCIT, to regular PCIT. This is often the case with EBPs. Because of real-world barriers to preserving the treatments in their original format and structure, it is hard to ensure that they are making impact. Therefore, providers of EBPs will likely benefit from tests of fidelity unique to every treatment.

Second, there may be a critical downside to allowing for modifications to the original treatment. The impact of in-home PCIT has only been explored through one case study with fewer than 10 subjects. It only provided inconclusive information about the treatment outcomes because the results of this case study could not be generalized to different populations. It is strongly suggested that RCTs or other studies that provide supporting evidence for EBPs identify the core components of implementation. This would inform providers what they can and cannot modify to achieve intended outcomes. However, there is clearly a limit to separately studying the impact of all the different components of one treatment.

Third, EBPs may give providers the sense that they are well-informed about the treatment and its impact on clients because they know the evidence base of the treatment they are administering. Providers may also take more control of the clinical sessions and settings and more efficiently make quick judgments when necessary as they are better informed. It
may be this confidence or assurance that gives clients the feeling that their providers are knowledgeable, as suggested by the data analysis.

In short, to ensure efficient implementation of EBPs, it is recommended that tests of fidelity and appropriate outcome assessment measures be incorporated whenever possible. The data analysis also suggested that clients of EBPs may be more likely to feel they are receiving treatment from knowledgeable providers because the treatments are thoroughly planned and highly structured. In the case of PCIT, which does not educate families on how to think or feel, clients were more likely to feel that providers respected their values and tradition. However, much effort should be put into identifying the core components of EBPs for high-fidelity implementation and efficient assessment. This remains a critical challenge for the research and service community.
References


Evidence-Based Practice: a Critical Appraisal (pp. 89-110). Oxford: Blackwell.


doi:10.1016/j.chiabu.2008.08.003


http://www.fpg.unc.edu/~nirn/resources/other/NIRNservices2006.pdf

http://www.fpg.unc.edu/~nirn/implementation/01_implementationdefined.cfm


