Journal of Consulting and Clinical Psychology
2011, Vol. 79, No. 1, 84–95

0022-006X/10/$12.00 DOI: 10.1037/a0021227

A Combined Motivation and Parent–Child Interaction Therapy Package Reduces Child Welfare Recidivism in a Randomized Dismantling Field Trial

Mark Chaffin, Beverly Funderburk, and David Bard
University of Oklahoma Health Sciences Center

Linda Anne Valle
Centers for Disease Control and Prevention

Robin Gurwitch
Cincinnati Children's Hospital Medical Center

Objective: A package of parent–child interaction therapy (PCIT) combined with a self-motivational (SM) orientation previously was found in a laboratory trial to reduce child abuse recidivism compared with services as usual (SAU). Objectives of the present study were to test effectiveness in a field agency rather than in a laboratory setting and to dismantle the SM versus SAU orientation and PCIT versus SAU parenting component effects. ***Method:*** Participants were 192 parents in child welfare with an average of 6 prior referrals and most with all of their children removed. Following a 2 × 2 sequentially randomized experimental design, parents were randomized first to orientation condition (SM vs. SAU) and then subsequently randomized to a parenting condition (PCIT vs. SAU). Cases were followed for child welfare recidivism for a median of 904 days. An imputation-based approach was used to estimate recidivism survival complicated by significant treatment-related differences in timing and frequency of children returned home. ***Results:*** A significant orientation condition by parenting condition interaction favoring the SM + PCIT combination was found for reducing future child welfare reports, and this effect was stronger when children were returned to the home sooner rather than later. ***Conclusions:*** Findings demonstrate that previous laboratory results can be replicated in a field implementation setting and among parents with chronic and severe child welfare histories, supporting a synergistic SM + PCIT benefit. Methodological considerations for analyzing child welfare event history data complicated by differential risk deprivation are also emphasized.

Keywords: parenting, PCIT, child abuse, child neglect

Parent training programs focused on teaching behavior management and parent–child interaction skills are a common part of child welfare service plans for maltreating parents. In a large, nationally representative sample of parents in child welfare, parenting programs were the most common type of service brokered by child welfare (NSCAW Research Group, 2005). Adaptations of evidence-based behavioral parenting programs such as parent–child interaction therapy (PCIT), The Incredible Years, or Triple-P are increasingly promoted for child welfare applications (California Evidence-Based Clearinghouse for Child Welfare, 2009; Centers for Disease Control and Prevention, National Center for Injury Prevention and Control, 2004). These models were not originally designed for child welfare populations, nor were they designed to deliver child maltreatment recidivism reduction. They were originally developed as parent-mediated treatments for childhood disruptive behavior disorders, and their adoption to achieve child maltreatment reduction outcomes represents a significant departure from their original purpose. Parents in child welfare are referred to parenting programs primarily to achieve *parent* behavior change, not *child* behavior change.

The current study focuses on PCIT (Eyberg & Boggs, 1998). The efficacy of PCIT as a treatment for its original target problem—externalizing child behavior disorders—has been well established (Eyberg, Nelson, & Boggs, 2008). In child welfare adaptations of the model, the ultimate desired outcome is reduction of harsh, violent, or neglectful parenting practices (e.g., Barth et al., 2005; Pinkston & Smith, 1998), suggesting the need to evaluate PCIT closely within the child welfare context and specifically for child welfare outcomes.

This article was published Online First December 20, 2010.

Mark Chaffin, Beverly Funderburk, and David Bard, Department of Pediatrics, University of Oklahoma Health Sciences Center; Linda Anne Valle, Centers for Disease Control and Prevention; Robin Gurwitch, Cincinnati Children's Hospital Medical Center.

This research was supported by cooperative agreement R49CE622338 from the U.S. Centers for Disease Control and Prevention. We wish to acknowledge the contributions of Carol McCoy, Desi Vasquez, Allison Garrett, Christi Madden, Carol Moore, and Rebecca Hurst. We especially recognize the efforts of staff at the implementation agency and the cooperation of the Oklahoma Department of Human Services. The findings and conclusions in this article are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Correspondence concerning this article should be addressed to Mark Chaffin, University of Oklahoma Health Sciences Center, P.O. Box 26901, Oklahoma City, OK 73126-0901. E-mail: mark-chaffin@ouhsc.edu

A rationale for adapting PCIT as a treatment for abusive or neglectful parenting was initially articulated by Urquiza and McNeil (1996). Child physical abuse often occurs in a context of child discipline and a dysfunctional or detached parent–child relationship. This can take the form of escalating coercive cycles similar to but more extreme than those thought to characterize the development of externalizing child behavior problems along with a low level of positive parent–child engagement and attachment (Patterson, 1976, 1982; Patterson & Reid, 1984; Urquiza & McNeil, 1996). Unresponsiveness to appropriate child behavior or child needs, deteriorated relationship quality, and withdrawn or weak parent–child attachments form a relational context for physical abuse and neglect (Stith et al., 2009). This pattern of negative or disengaged sequential interactions has been found to change rapidly and consistently among parents in child welfare who receive PCIT, and coded parent–child interaction changes observed in the laboratory have been found to partially mediate long-term reductions in abuse case recidivism (Chaffin et al., 2004; Hakman, Chaffin, Funderburk, & Silovsky, 2009).

The parenting skills taught in PCIT are similar to those taught in many other well supported behavioral parenting models. Most of these models directly target coercive cycle behaviors, relationship deterioration, and disengaged aspects of parent–child interactions. In contrast to their didactic parenting model counterparts, behavioral parenting models teach clear, observable behavioral skills and often involve direct skill modeling, practice, and observation. A defining aspect of PCIT is its live coached delivery approach. Skills are taught by coaching the parent to criterion during live parent–child interactions. Interactions are observed by the therapist through a one-way mirror, and coaching is accomplished via a wireless earphone worn by the parent. Live parent–child skill practice approaches are more labor intensive than group didactic delivery modes but have been associated with larger effect sizes (Kaminski, Valle, Filene, & Boyle, 2008) and may be worthwhile for parents in child welfare where large behavior changes are usually desired.

The main modification made to PCIT in our 2004 randomized trial (Chaffin et al., 2004) was the addition of a short, self-motivational (SM) orientation intervention. This modification was introduced on the basis of the assumption that more active behavioral parenting models like PCIT require a higher degree of client motivation in order to be feasible. Unlike didactic parenting group models, where content can be passively consumed, PCIT requires active in vivo skill demonstration, practice, and criterion attainment.

In regular PCIT for child behavior problems, parents are usually voluntary service seekers and therefore may be well motivated to participate in services and apply skills in the home environment. This is contextually different from parents in child welfare who are typically coerced into services and may not be self-motivated. Some evidence supports the benefit of the SM component. In an earlier report from this same project, SM plus PCIT (SM + PCIT) was found to yield better program retention rates (85% vs. 61% cumulative program retention) compared with other component combinations and was especially effective at retaining the initially lowest motivated parents (Chaffin et al., 2009).

The 2004 outcome trial yielded encouraging findings. Among 110 physically abusive parents, child welfare recidivism survival over a 2.5-year follow-up was reduced from 49% to 19% compared with a services as usual (SAU) didactic parenting package (Chaffin et al., 2004). Parents in the 2004 laboratory trial had an average of two prior child welfare referrals and had entered child welfare because of physical abuse. Based on the favorable findings, PCIT was evaluated as one of the potentially most cost-effective models available for reducing physical child abuse (Lee, Aos, & Miller, 2008). In the 2004 laboratory trial, the combined SM + PCIT package was tested compared with a SAU package consisting of an informational SAU orientation intervention combined with a didactic SAU parenting group. The experimental design permitted direct comparison of the two packages but did not allow separate estimation of the orientation components (SM vs. SAU orientation), the parenting components (PCIT vs. SAU parenting), or the combinatory and synergistic effect that was hypothesized to exist. In the present study we used a crossed 2×2 sequentially randomized design to dismantle the orientation and parenting components of the packages used in the 2004 study. In this article we examine child welfare recidivism outcomes at follow-up, with attention to component main and interaction effects.

Another main aim of this trial was to determine whether the recidivism reduction benefits found in the laboratory trial could be replicated in a field setting. It is well known that laboratory trial efficacy may be either absent or severely attenuated when interventions are transported to actual field clinical settings (Weisz, Weiss, & Donenberg, 1992). In our 2004 trial, the SM + PCIT package was delivered in a university-based laboratory setting, whereas the SAU package was delivered in a field agency setting with equivalent clients but with a different set of clinical resources, competing demands, and provider characteristics. In the present study, both of the orientation components (SM vs. SAU) and both of the parenting components (PCIT vs. SAU) were delivered in the same field agency setting, with agency staff providing all services. This effort followed the general translational research agenda for moving from more internally valid efficacy studies to comparative effectiveness trials that test generalizability in community settings (Wells, 1999).

Method

Participants

Inclusion and exclusion criteria. Participants were 192 maltreating biological parents, stepparents, or primary caregivers referred for parenting services at a small, inner city, nonprofit, community-based agency operating a parenting program under contract with the single state child welfare system. Study inclusion criteria for parents included a referral to the program by child welfare for neglect and/or physical abuse, at least one child between ages 2.5 and 12 years who was available to participate in PCIT if so randomized, and a Kaufman Brief Intelligence Test (KBIT; Kaufman & Kaufman, 1990) parent IQ score of at least 65, which was required to complete study measures. Parents were included if they had access to at least one child, including children in foster or kinship care, if interaction with that child in PCIT sessions would be permitted. Parents were excluded if they had ever received PCIT in the past or if all parental rights on all children were being terminated. There were provisions to exclude parents if either the treating clinician or researchers judged that

parent–child contact during PCIT was distressing to the child, but this option never had to be exercised. Parents who became ineligible over the course of the study (e.g., had all parental rights terminated) were withdrawn from the study, and this was done equivalently across assigned study conditions. Ineligible, excluded, or withdrawn participants retained access to open services at the agency.

Demographic data. The participant group was 75% women, with a mean age of 29 (SD = 6.5, range = 20–57). Sixty percent were non-Hispanic Caucasian, 19% were African American, 9% were Native American, 7% were ethnically Hispanic, and 6% were of another race/ethnicity. Thirty-five percent were married or cohabitating, 29% were never married, 18% were divorced, and 17% were separated. Twenty-nine percent had less than a high school or equivalent education, 40% were unemployed, and 48% were receiving some form of public assistance. Seventy-five percent of households fell below the federal poverty threshold, and the median household income was around $900/month. Parents had an average of 2.6 children in their household, and 10% of the women reported that they were currently pregnant. Most parents (73%) had at least one child in the preschool age range.

Measures

Assessment information for this study was drawn from three sources—self-report questionnaires administered via audio-assisted computerized self-interview (ACASI) using touch-screen computers, observational coding of parent–child interactions, and administrative data from the state child welfare database. Computer interviews were administered at baseline, at completion of the six-session orientation pretreatment, and at completion of the subsequent 12-session parenting program. The aims of the present report rely on baseline data and recidivism follow-up data, both of which were available for all participants. Research assistants capturing baseline data demonstrated the ACASI system and observed practice items to ensure that participants understood its proper use. Computer interviews were conducted in private. Research assistants normally did not view participants' responses unless assistance was requested. A federal certificate of confidentiality was obtained to protect participant responses, and participants were assured that their answers would not be shared with child welfare authorities or the service agency. Observational coding of parent–child interactions was scheduled at the same landmark points as the questionnaire data and was coded live from behind a one-way mirror by a research assistant. Median time of active study involvement (baseline until posttreatment data collection) was 247 days. Database matching for child welfare recidivism was performed for all participants at a median of 904 days from baseline (range = 229 to 1,282 days).

Demographic questionnaire. A basic demographic questionnaire was used to capture parent and family characteristics, referral information, and family background and economic information. An earlier version of the questionnaire was pilot tested on 100 parents in similar programs, and items answered inconsistently or indicated by parents to be confusing were modified prior to use in the present study.

Readiness for Parenting Change scale (REDI). The REDI is a measure of motivation to change parenting behavior originally developed by Mullins, Suarez, Ondersma, and Page (2004) for use with substance-abusing parents in combined substance abuse and child welfare services. The measure was modified and expanded for the present study by rewording items to reflect motivation for participating in a child welfare parenting program and adding items reflecting parenting attitudes and attitudes toward being mandated to attend the program. The expanded item pool was pretested with 122 anonymous nonstudy clients at the study site prior to beginning the randomized trial in order to identify confusing items and test the modified scale. Items rated as confusing by clients were dropped. A confirmatory maximum likelihood factor analysis was executed using Mullins et al.'s original REDI subscales (Readiness to Change, Self-Efficacy, and Problem Recognition) plus groups of newly added items reflecting belief in harsh discipline and attitude toward the program. Items with information values less than .25 were excluded, resulting in a 23-item instrument with an overall omega value of .92 and a correlated subscale structure. Overall internal consistency was rechecked using the present study sample, and the overall alpha was .84. The REDI has previously been reported to predict parenting program retention in this study sample (Chaffin et al., 2009). The REDI was used in this study as a manipulation check for the SM orientation.

Child Abuse Potential Inventory (CAP). The CAP (Milner, 1986) is a widely used 160-item agree–disagree format questionnaire developed to estimate risk for committing child physical abuse. Items on the instrument primarily reflect personal characteristics and parenting attitudes associated with maltreatment, not maltreatment behavior itself. The CAP Abuse Scale has been reported to have high internal consistency (KR-20 of .92 to .95), a 1-month test–retest stability of .83, good discriminant validity, and adequate future predictive validity for both physical abuse and neglect (Chaffin & Valle, 2003; Milner, 1986, 1994). Alpha for the CAP Abuse Scale among participants in the present study was .92. The CAP was used in this study to test prerandomization equivalence and as a covariate.

Dyadic Parent–Child Interaction Coding System—II (DPICS–II). The DPICS–II is a system for coding observed parent–child interactions during a structured three-part task: child-directed activity, parent-directed activity, and clean-up from the activity (Eyberg, Bessmer, Newcomb, Edwards, & Robinson, 1994). The tool was developed specifically for PCIT and corresponds to the parent behaviors that are coached in PCIT. Parent verbal behavior (e.g., commands, praises, criticisms, reflections, sarcasm), physical behavior (e.g., physical positives such as hugs or pats; physical negatives such as slaps), and unresponsiveness to the child's behavior are coded. Interrater and test–retest reliability as well as discriminant validity between referred and nonreferred children are satisfactory (Bessmer, 1998; Robinson & Eyberg, 1981). Research assistants conducted the observation sessions and coded the instrument. Coder training required attaining a criterion of at least 90% correct with standard video stimulus sets prior to coding study data. Periodic random review of coding sessions was conducted by a study investigator with many years of experience as a DPICS–II trainer in order to ensure that no coder fell below this criterion. The trainer would independently code the session, compare codes with the research assistant, and provide any necessary corrective feedback. The DPICS–II was used in this study as a manipulation check for PCIT.

Child welfare database matches. Administrative data for child welfare reports and child placements were captured from service agency and state child welfare agency records and databases. The state maintains a central database that logs all reports, allegations, dispositions, and placements. Because all study participants were already in the child welfare system at enrollment, unique parent identifiers from the state child welfare database were collected at enrollment and used for subsequent matching. Matches were made to capture future events where the parent in the study was identified as the person committing maltreatment of any child in the family. The identity of the research participant was confirmed using both the unique database identifiers and either social security number match or date of birth combined with name match. Reports that were ruled out or screened out by child welfare were excluded, and the remaining reports were aggregated across report dates and children involved in the report to yield unduplicated event histories. There were no identified surveillance reports (i.e., reports made by the treatment agency).

Placement data on children in the family were aggregated into placement episodes. A placement episode was defined as a sequence of placements punctuated by a terminal event such as reunification. Transient placement episodes lasting fewer than 10 days (e.g., a brief shelter placement followed by reunification) were not counted as an out-of-home placement. Placement episode start dates were juxtaposed against new report dates and attributed to a new report if the episode began on or shortly after the report date.

Configuring child welfare report data to account for risk deprivation. Configuring and analyzing child welfare report events and survival times in this study presented a particular challenge. First, as previously noted, two thirds of the study participants had all of their children removed at enrollment. A parent with no children in his or her home has virtually no opportunity for a new child welfare report, which we refer to as *risk deprivation*. Confusion between periods of risk deprivation and bona fide event-free survival would introduce substantial bias if this was not equivalent across treatments (i.e., if reunification events are conditional postrandomization outcomes). For example, treatments that are more successful in helping parents keep or regain their children might accrue greater risk exposure. On the other hand, treatments with low rates of family reunification could present the illusion of lower maltreatment risk when in fact there was simply less opportunity for maltreatment to occur.

As parents progressed through their treatment program, most had children returned to their homes and thereby became eligible for a new report, but this occurred at different time points if it happened at all. Three data sources were used to identify if and when risk deprivation ended. Official placement and reunification data were used to identify the earliest date at which any child was returned. Self-report data about having children in the home were included in the computer interviews during the service interval. Narrative notes made by data collectors were examined. The earliest point at which a parent was documented to become eligible for a new report across these three sources was used. Pregnant participants were not treated as risk deprived.

To examine whether risk depravation significantly affected recidivism survival, we adjusted survival and follow-up times by removing periods of risk deprivation, and the resulting times were compared with unadjusted figures. Accounting for risk deprivation resulted in substantial survival and follow-up time changes. For example, the median unadjusted follow-up time was 904 days; however, the median risk-deprivation-adjusted follow-up time was 730 days. The median unadjusted recidivism survival time was 640 days, but the median adjusted survival time was 393 days. Some parents were determined to have never had children returned during the follow-up interval. Ultimately, 22% of the participants experienced termination of parental rights at some point during the entire treatment and follow-up interval, which is an extraordinarily high number (termination of parental rights rates are difficult to estimate for entering child welfare cohorts, but the observed rate in this sample may be as much as 10 times the average). Termination of parental rights is the most drastic action available to child welfare and the juvenile court, and again emphasizes the severity of the study sample.

Risk deprivation differed across cells of the experimental design for both partial and total risk deprivation, as shown in Figure 1. It is important to note that risk deprivation includes both a baseline element (i.e., having children in the home at baseline, which did not differ significantly) and a posttreatment outcome or principal stratification element (child return), and it was on the latter that the most substantial frequency and timing differences were observed. In effect, differences in one possible treatment outcome (differential patterns of child return) created a conditional and hidden missing data problem in the main treatment outcome of recidivism. We discuss in a subsequent section how this challenge was addressed analytically.

Procedure

Enrollment was conducted at the service agency by research assistants between January 2004 and August 2006. All procedures were approved by both a local university and a federal government institutional review board. A total of 291 parents were approached to explain the study and screen for eligibility. Of those, 42 declined and 38 were excluded on the basis of eligibility criteria, leaving a preliminary eligibility pool of 211. Of those 211, 19 never completed the baseline assessment and were withdrawn. Two additional participants were withdrawn after baseline assessment, but before initial randomization, because they lost eligibility due to termination of parental rights. This resulted in a first randomization sample of 192 participants who were randomized using an unblinded sequential computer-generated randomization list to one of the two orientation conditions (SM = 99, SAU = 93). Of these 192, 21 lost study eligibility during the orientation treatment phase because they lost access to all of their children (termination of rights and/or being prohibited from having contact with any of their children) and were withdrawn. An additional 11 who were randomized to an orientation condition never returned to the agency for any services (i.e., treatment refusers), and 7 more dropped out of orientation after engaging, leaving 153 available for the second randomization to a parenting condition. These 153 were randomized (PCIT = 70, SAU = 83) with a sequential randomization list, and they form the final analysis sample.

The main source of preinclusion and postinclusion attrition was involuntary withdrawal due to loss of access to children, which precluded assignment of a parenting condition. Recall that a parent who cannot see any of his or her children cannot be assigned to PCIT because it is a dyadic treatment model, and so these partic-

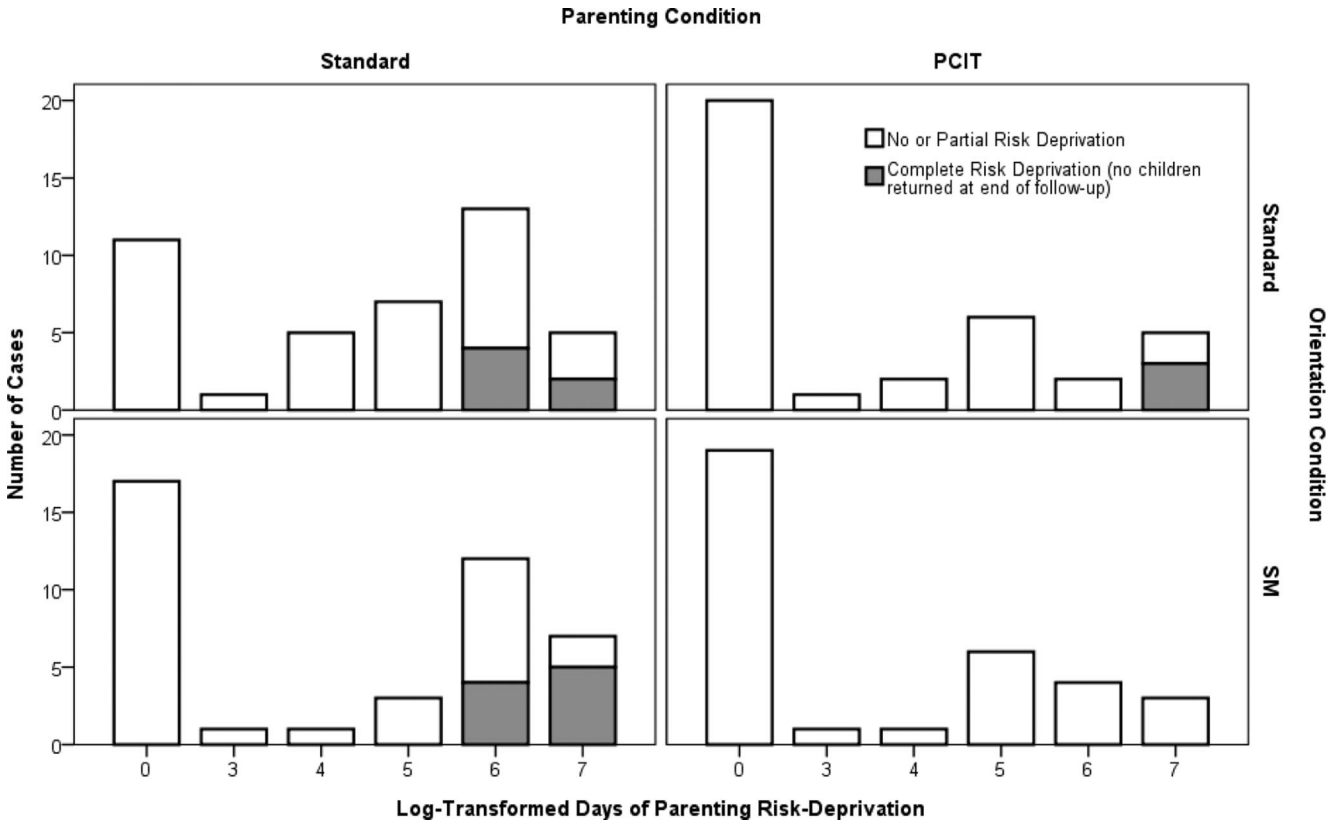

Figure 1. Differential risk deprivation outcomes across 2 × 2 design cells. Gray sections of bars indicate participants who never had any children returned during the follow-up period. All participants in the self-motivation (SM) with parent–child interaction therapy (PCIT) condition reunified. A single day was added to each risk-deprivation length so that zero could indicate no child removal at enrollment.

ipants could not be randomized. A substantial number of parents lost access to their children around or shortly after enrollment. For example, termination of parental rights often was under consideration, but no official petition had been filed at the point of recruitment, and so the participant was enrolled. When an official termination of parental rights petition was filed, often shortly after enrollment, the participant had to be withdrawn from the study. Involuntary withdrawal due to loss of eligibility is an unusual source of attrition, but the unique circumstances of this trial made it more common. We do not believe this poses a high threat to generalizability because individuals with no children are not a population for whom parenting programs are designed.

Overall rates of attrition and reasons for attrition (refusal, withdrawal due to loss of eligibility, drop-out) did not differ significantly across the randomized conditions. The final analysis sample (n = 153) represented 53% of those who were initially approached and screened for preliminary eligibility, 80% of those who completed the baseline assessment and were determined to be fully eligible, 89% of those remaining eligible by the end of orientation, and 100% of those who were still available and eligible for the second randomization. Once parents were randomized to a parenting treatment condition, completing their membership in one cell of the 2 × 2 experimental design, intention to treat principles were followed. A participant flow diagram can be seen in Chaffin et al. (2009).

SM orientation condition. The SM orientation condition was a manualized group program using the protocol from the Chaffin et al. (2004) laboratory PCIT study (designed dose = 6 sessions; mean received dose = 5.2 sessions, SD = 1.3). The SM protocol was derived from general motivational interviewing principles (Miller & Rollnick, 1991) and included sessions involving hearing testimonials from parents who previously completed the parenting programs, performing decisional balance exercises weighing the pros and cons of harsh physical discipline and the pros and cons of change, encouraging parents to develop their own list of parenting and parent–child relationship goals, elaborating discrepancy between current parent–child interaction patterns and personalized goals, and encouraging commitment to change. Although motivational interviewing is normally individualized and delivered in one-on-one sessions, a group approach was selected so that the SM and SAU orientation conditions in the study would share a common delivery mode.

SM was delivered by master's-level agency therapists (n = 4; 1 man, 3 women) who were initially trained in the protocol by investigators and were monitored periodically by study staff for fidelity, using session checklists. These were reviewed in weekly clinical supervision meetings conducted by investigators with the therapists, and any fidelity or implementation problems were addressed. Fidelity criterion was set at 90% of checklist items. All therapists were able to maintain reasonable fidelity, and none were

removed or placed in remedial training. Manuals for this protocol are available from Beverly Funderburk.

SAU orientation condition. The SAU orientation condition was a manualized group orientation program that had been routinely implemented at the field agency for many years (designed dose = 6 sessions; mean received dose = 5.2 sessions, *SD* = 1.4). The focus was primarily informational and educational and included information about the roles of child welfare and the agency, definitions of child maltreatment and how it affects children, information about agency services, and information about the possible insight-oriented links between a parent's own childhood experiences and current parenting practices. The SAU orientation group program was delivered by master's-level therapists (*n* = 11; 1 man, 10 women) who were not fidelity monitored by the study, but the agency provided a schedule of weekly clinical supervision comparable in dose to the SM condition. Cross-contamination between the two orientation conditions was considered a potential threat to internal validity, and so the agency agreed that their therapists would deliver one orientation condition or the other but not both. SM materials and training were not shared with SAU therapists until after the study was completed.

PCIT parenting condition. The PCIT parenting condition followed the same protocol used in the Chaffin et al. (2004) laboratory study (designed dose = 12–14 sessions; mean received dose = 10.1 sessions, *SD* = 3.6). In PCIT, parents learn a specific behavioral skill set during two sequential program phases. During the first phase, Child-Directed Interaction (CDI), parents learn to follow their child's lead in dyadic play and provide positive attention for desirable behavior combined with active ignoring of minor misbehavior. Major objectives of CDI are to improve the quality of the parent–child relationship, improve attachment and engagement, and establish attention to positive child behavior. In CDI, parents are taught to use the P.R.I.D.E. skills—Praise, Reflection, Imitation, Description, and Enthusiasm—to reinforce children's positive behaviors. Parents also are taught to avoid specific types of parenting behaviors such as threats, physical aggression, commands, questioning, criticism, and sarcasm.

During the second phase of PCIT, Parent-Directed Interaction (PDI), parents learn to give effective commands and instructions, to use a consistent step-by-step time-out protocol in response to child noncompliance, and to properly reinforce child compliance. The time-out protocol is highly structured and includes specific solutions to possible parenting challenges (e.g., managing child behavior in public places, managing child escape from time out). The regular PCIT time-out protocol was adapted for parents in child welfare to include (a) basic self-control techniques such as deep breathing and (b) elimination of corporal punishment as a back-up for time out. Adaptations were made to accommodate children up to age 12 years rather than the regular PCIT age ceiling of 7 years. The child age ceiling of regular PCIT was felt to be less relevant for PCIT in the child welfare context given that our focus is on changing parent behavior rather than changing child behavior problems, and thus children's participation is collateral. Additional adaptations were required for parents who had all of their children placed in foster or kinship care, mostly pertaining to homework assignments. These included using role plays to practice skills and practicing CDI skills during scheduled visitations.

PCIT was delivered by master's-level agency therapists (*n* = 7, all women) who were initially trained by study staff. PCIT sessions were periodically observed and coded by investigators using a checklist and were discussed in weekly clinical supervision. All therapists were able to maintain reasonable fidelity, and none were removed or placed in remedial training. Manuals for this protocol are available from Beverly Funderburk.

SAU parenting condition. The SAU parenting program consisted of a weekly didactic parenting group in which parents learned about child development and developmentally appropriate expectations, principles of discipline, use of praise, communication strategies, stress management, and the ways in which parental personal problems affect children (designed dose = 12 sessions; mean received dose = 9.0 sessions, *SD* = 4.4). Special needs and crises presented by parents also were addressed during group discussion. The treatment utilized an unpublished curriculum and manual developed by the field agency, with elements similar to common didactic parenting books and curricula.

The SAU program also employed several additional modules after the basic parenting curriculum was complete. These included a proprietary commercial curriculum on compassionate parenting designed to promote parental empathy, and several ancillary counseling services. These additional services were not used with PCIT recipients but were used liberally with SAU parenting participants, who received an average of 10.4 additional sessions (*SD* = 4.7) of service beyond the 12-session parenting curriculum. Thus, the total dose of agency services received for those randomized to SAU was almost double that received by those randomized to PCIT.

Our review of the SAU model suggested that it was similar in content to many other didactic parenting classes and anger management programs used with parents in child welfare. The primary focus of the program was on attitudes, beliefs, feelings, and concepts surrounding parenting, along with experiences in the parent's own childhood. This contrasted sharply with the briefer but far more focused, less abstract, and more concrete behavioral skill approach used in PCIT. The SAU parenting condition was delivered by master's-level therapists (*n* = 18; 2 men, 16 women) and was not fidelity monitored by study staff, but the agency provided a schedule of clinical supervision comparable to that of the PCIT condition.

Given that the formats of the two parenting conditions were so fundamentally different (individual parent–child dyad format for PCIT vs. parent group discussion format for SAU), there were limited opportunities for cross-contamination between PCIT and SAU. The limited risk of cross-contamination was viewed as outweighed by the advantages of therapist counterbalancing, and so all of the PCIT-trained therapists also delivered SAU parenting.

Results

Participants

The mean and median KBIT IQ score was 92 (range from the exclusion floor of 65 to 116). On average, parents had a large number of prior referrals into the child welfare system, suggesting an unusually chronic child welfare population. The 192 parents had accumulated a total of 1,142 unduplicated prior household referrals to the state child welfare system that were not ruled out or screened out, with a mean of six prior referrals per household (*SD* = 4.8, median = 4, range = 1–39). The majority (70%) of past household referrals involved child neglect, followed by phys-

ical abuse (23%) or sexual abuse (6%). Most parents (76%) had one or more children placed in foster care at enrollment, which is almost four times the average rate for all child welfare cases (U.S. Department of Health and Human Services, Administration for Children, Youth and Families, Children's Bureau, 2010), and 66% had *all* of their children removed to foster care. Given that removal of children requires a juvenile court determination of acute and significant danger to the child, this suggests that the maltreatment often was severe. The mean CAP score in the study sample was 156 ($SD = 107$), and it was 202 ($SD = 104$) for valid profiles, compared with a general population norm for valid profiles of 91 ($SD = 75$; Milner, 1986). Few parents considered themselves voluntary service seekers. Eighty-one percent self-reported that their motivation for seeking parenting services was either "told to come by child welfare" or "ordered to come by the court."

Adverse Events

One parent with a child in foster care did request to be withdrawn from PCIT due to finding it emotionally difficult to separate from her child after dyadic parent–child sessions. Other parents with children in foster care reported that the additional parent–child contacts were an incentive to participate in PCIT, suggesting that this feature of PCIT can evoke different responses across different parents.

Randomization Checks

To check the initial (orientation condition) randomization of the 2×2 design, we conducted a series of simple bivariate tests comparing the SM and SAU orientation conditions for baseline differences in parent age, gender, marital status, IQ, education level, race/ethnicity, income, poverty status, number of prior child welfare referrals, self-report of being ordered into services, number of children, number of children in out-of-home placement, and baseline scores (CAP, REDI, and DPICS–II). No statistically significant differences were found. To check the second (parenting condition) randomization of the 2×2 design, we conducted the same set of bivariate tests comparing the PCIT and SAU parenting conditions. No statistically significant differences were found.

Manipulation Checks

A manipulation check for the SM intervention was performed comparing motivation change, as measured by the REDI, from Wave 1 to Wave 2 (median interval = 64 days) using only the initial orientation condition randomization as a between-subjects factor. Using a two-within, two-between repeated measures analysis of variance, we found that REDI scores improved over time across both conditions, $F(1, 125) = 18.1$, $p < .001$, but the scores improved significantly more in the SM orientation condition than in the SAU orientation condition, $F(1, 125) = 7.3$, $p < .01$, Cohen's $d = 0.33$. The SM intervention yielded significantly greater motivational change, as it was designed to do.

A manipulation check for PCIT was performed by comparing DPICS–II scores, which were aggregated according to whether the parent behavior was positive (coached to increase in CDI) or negative (coached to decrease in CDI). This manipulation check included DPICS–II codings at all three waves (median total interval = 247 days) so that parenting condition effects were estimated within the context of the prior orientation manipulation. Change over time was modeled with a mixed-effects approach, including the full sample, with the three time points treated as a repeated effect and no error variance structure imposed. The single orientation condition dummy variable was allowed to affect all three time points, but the parenting group dummy variables were constructed to be time-dependent, such that parenting condition participated only in the Wave 2 and Wave 3 time points. This dummy matrix was configured to correspond to the sequentially randomized experimental design.

As we examined the raw data, it became clear that participants receiving PCIT not only showed the expected changes in DPICS–II scores but also had constricted variability at the post-parenting time point, especially for reduced negative behaviors. This is to be expected given that PCIT coaches these behaviors to criterion. Within-subjects correlation between adjacent DPICS–II negative behavior scores was substantial for SAU parenting participants ($r = .41$) but was small for PCIT participants ($r = .06$) because PCIT Wave 3 DPICS–II scores grouped around criterion, irrespective of baseline or Wave 2 score. Significant main effects in favor of PCIT versus SAU parenting were found for both decreases in negative parent behaviors, $F(1, 103.2) = 36.1$, $p < .001$, effect size = 1.07, and increases in positive parent behaviors, $F(1, 119.2) = 8.0$, $p < .01$, effect size = .64, using the procedures described by Feingold (2009) for estimating effect sizes in mixed models. No orientation or interaction effects approached significance. These findings support a conclusion that the PCIT intervention uniformly taught the in-clinic parenting skills that the model was designed to teach.

Recidivism Outcomes and Imputation Approach

The raw observed recidivism percentages were 38% for the entire sample, 29% for SM + PCIT, 47% for SAU + PCIT, 41% for SAU + SAU, and 34% for SM + SAU. Among the 58 cases with a new event, 12 first events (or 21%) resulted in a new removal from the home (10% of SM + PCIT events, 18% of SAU + SAU events, 24% of SM-SAU events, and 29% of SAU-PCIT events). These raw percentages were known to be biased by differential risk deprivation, particularly given that the design cell with the lowest raw recidivism percentage (SM + PCIT) was also the design cell with the most children returned home and therefore the greatest risk exposure. Two approaches were used to estimate recidivism survival in the context of treatment-related risk deprivation. The first approach we employed was to exclude fully risk-deprived cases (i.e., cases where event history outcomes were treated as completely missing due to never having children in the home across the entire follow-up interval) and to use exposure-adjusted survival times for all other cases (i.e., those that had partial but not total risk deprivation, plus those with no risk deprivation). This approach analyzes only the at-risk intervals that were actually observed and excludes subjects who were never at risk. This modeling approach was used to inform covariate selection for the imputed model, which is described below, and as a primary model against which to contrast the imputation-based approach.

The second approach relied on imputation of survival outcomes during all risk-deprived intervals for both fully and partially risk-

deprived cases. In this approach, all risk-deprived intervals, up to the administrative censoring point at the end of observed follow-up, were treated as missing data for which event history outcomes were imputed. The imputation procedure reflects an attempt to balance the unequally distributed risk deprivation across the experimental design cells and allow treatment condition comparisons to be unbiased by differential patterns of postrandomization risk exposure.

Because imputation was required for a variety of covariate distributions and for event history outcome, we relied on the fully conditional imputation procedure provided by the MICE library in R (van Buuren, 2007; van Buuren & Oudshoorn, 2000). This procedure attempts to approximate the joint multivariate distribution of all variables through an iterative algorithm of sequential conditional regression equations (i.e., variables are imputed one at a time using a sequence of conditional univariate regressions). The procedure allowed us to write our own predictive equations for the event history outcomes. Since there was very little missingness in our covariate data (most of which were collected at baseline), covariates were allowed to predict event histories but event histories did not predict covariates.

Newly imputed event history times were capped at observed follow-up to avoid unwarranted extrapolation beyond our sample of observed event histories (see Faucett, Schenker, & Taylor, 2002, for related analytic rationale), and for partially risk-deprived individuals imputed event times had to exceed their risk-adjusted follow-up interval length because we did not wish to impute an event during an interval known not to include a referral report. The remaining administratively censored event times were not imputed. Thus, we assumed these censorings were unrelated to timing of eventual maltreatment report outcomes, which is a standard event history and survival analysis censoring assumption. The event history imputation model selected was a time-dependent Andersen and Gill (1982) counting process version of the Cox survival model available in the coxph package of R. The time dependency accounted for time-sensitive influences of the two-stage experimental condition randomization on report outcomes. Hazard estimates were imputed for a large number of very small discrete intervals, then hazard distributions were sampled to determine whether an event would be assigned to that interval in each imputed data set. Because we expected that this procedure might yield possible volatility in event history estimates, we chose to exceed recommendations for both the number of MICE Gibbs sampling iterations and the number of multiple imputation datasets. We set our limits of iterations and imputed datasets at 20 each.

In addition to treatment effects, the imputation model included group-invariant covariate prediction of proportional event hazard differences. Due to limited sample size within treatment group combinations (orientation by parenting design conditions), we felt comfortable including only a single design cell-specific covariate effect in the event history prediction model. The most reliable moderated treatment effect present among cases with complete data involved our measure of risk-deprivation length (i.e., time until any child returned home). This treatment interaction was present in our imputation model via two-way and three-way interaction terms with the covariate and both treatment factors (orientation and time-dependent parenting). We chose not to impute the missing risk-deprivation lengths for the fully risk-deprived individuals (i.e., did not impute an expected date of child return) and instead fixed these values at each participant's difference between enrollment and the administrative censoring date. More detailed information about the specifics of the imputation procedure and access to the R syntax of our imputation model are available from David Bard.

Before imputing data, we first explored the predictive impact of each of 28 baseline covariates on event histories with the primary model, using only observed at-risk intervals. There were only two significant covariate interactions with treatment main or interaction effects when predicting event history outcomes, one involving age and another involving risk-deprivation length. Model inclusion of the two covariate-by-treatment interactions simultaneously produced unstable estimates, likely due to our limited sample size. The stronger interaction (by p value criteria) was found for risk deprivation. The final imputation model therefore included main effects for all covariates and each of the two treatment factors, a two-way treatment factor interaction, and the two-way and three-way interactions of risk deprivation by treatment factors. Convergence was monitored through visual inspection of iterative trends in univariate between-imputation covariate, event history, and censoring indicator means and standard deviations. Between- and within-imputation variability suggested adequate convergence (i.e., sufficient mixing of mean and standard deviation trajectories over the iteration sequence).

The proc MIANALYZE package of SAS was then utilized to combine analytic parameter estimates and standard errors from each dataset and to provide inferential t tests using standard aggregation rules and test procedures (Rubin, 2004). We began analysis with a look at the predictive impact of each covariate, separately, in a time-dependent (accounting for parenting condition staging times) Cox model controlling for all treatment and risk-deprivation main effects and interactions. We chose to include, simultaneously, all covariates associated with predictive coefficient p values less than .10 in our primary analysis Cox model. These covariates included marital status, education level, number of prior household child welfare referrals, income, poverty status, age, KBIT IQ score, number of children in the home at time of study enrollment, and number of children in placement at study enrollment. Final Cox model results appear in Table 1, along with the results of the primary model, which was executed in identical fashion with the same covariates but not using the imputed data sets and examining only observed at-risk intervals.

A key point in comparing the parameters of these two models is the similarity between hazard ratios for the specific effects. Because the imputation approach better managed bias, we will henceforth report findings from that approach, although findings from both approaches were comparable. Graphical residual diagnostics (see Grambsch & Therneau, 1994) evidenced no serious departures from final model hazard proportionality or covariate functional form, and inspection of dfbeta coefficients did not consistently identify overly influential observations across imputed data sets. Summary and plots for each diagnostic procedure across all 20 imputation datasets are available from David Bard.

Three covariate main effects were significant ($p < .05$) in the final model. When we controlled for other model predictors, higher participant ages were associated with significantly longer event histories (hazard ratio [HR] per year increase = 0.94), greater number of prior referrals was associated with shorter event

Table 1
Cox Regression Coefficients, Standard Errors, Hazard Ratios, and 95% Confidence Intervals of Hazard Ratios

Model	Estimate	*SE*	Hazard ratio	95% confidence interval of hazard ratio
Primary model—observed risk intervals only				
SM orientation	1.01	0.56	2.74	[0.92, 8.23]
PCIT parenting	0.86	0.59	2.36	[0.74, 7.51]
SM × PCIT	−3.22	0.97	0.04	[0.01, 0.27]*
Risk deprivation[a]	0.03	0.11	1.03	[0.83, 1.28]
SM × Risk Deprivation	−0.33	0.16	0.72	[0.53, 0.98]*
PCIT × Risk Deprivation	−0.39	0.18	0.68	[0.48, 0.96]*
SM × PCIT × Risk Deprivation	0.89	0.28	2.43	[1.41, 4.22]*
Marital status	−0.44	0.37	0.64	[0.31, 1.33]
Education	−0.24	0.19	0.79	[0.54, 1.14]
Prior referrals	0.08	0.03	1.08	[1.02, 1.15]*
Income	0.00	0.00	1.00	[1.00, 1.00]
Poverty status	−0.62	0.78	0.54	[0.12, 2.48]
Age	−0.09	0.03	0.91	[0.86, 0.97]*
KBIT	−0.02	0.02	0.98	[0.94, 1.02]
Number of children	−0.45	0.18	0.64	[0.45, 0.91]*
Prior placements	0.01	0.21	1.01	[0.67, 1.52]
Imputed event history model				
SM orientation	0.62	0.51	1.86	[0.68, 5.05]
PCIT parenting	0.71	0.53	2.04	[0.72, 5.75]
SM × PCIT	−2.95	0.91	0.05	[0.01, 0.31]*
Risk deprivation[a]	0.03	0.09	1.03	[0.86, 1.23]
SM × Risk Deprivation	−0.27	0.12	0.76	[0.60, 0.97]*
PCIT × Risk Deprivation	−0.21	0.13	0.81	[0.63, 1.05]
SM × PCIT × Risk Deprivation	0.77	0.23	2.15	[1.38, 3.39]*
Marital status	−0.32	0.35	0.73	[0.37, 1.44]
Education	−0.20	0.18	0.82	[0.58, 1.17]
Prior referrals	0.06	0.03	1.07	[1.00, 1.13]*
Income	0.00	0.00	1.00	[1.00, 1.00]
Poverty status	0.11	0.76	1.11	[0.25, 4.95]
Age	−0.07	0.03	0.94	[0.88, 0.99]*
KBIT	−0.02	0.02	0.98	[0.94, 1.02]
Number of children	−0.40	0.17	0.67	[0.48, 0.94]*
Prior placements	0.09	0.20	1.10	[0.74, 1.62]

Note. SM = self-motivational orientation; PCIT = parent–child interaction therapy; KBIT = Kaufman Brief Intelligence Test.
[a] Log transformed.
* $p < .05$.

histories (HR per referral increase = 1.07), and, perhaps surprisingly, greater number of children in the home was associated with longer event histories (HR per child increase = 0.67). The direction of the latter association likely captures effects related to partial risk deprivation. The modeled treatment factor interaction (Orientation × Parenting Condition) was also significant ($p < .002$), suggesting a clear advantage of the SM + PCIT design cell compared with the mean of the other three cells. Bonferroni-corrected pairwise comparisons were conducted among the four design cells, which suggested longer event-free survival for the SM + PCIT group relative to SAU + PCIT (HR = 0.11, $p < .05$) and SM + SAU (HR = 0.10, $p < .05$), and a sizable trend relative to the SAU + SAU group (HR = 0.20). The three-way interaction between treatment factors and risk-deprivation length was also significant ($p < .001$).

The latter finding demonstrated that the effect of the treatment group combination was not equivalent across levels of risk deprivation. The model was re-estimated at a high level of risk deprivation, and no significant pairwise differences were observed. This finding was driven by moderation of the comparative SM + PCIT benefit, where those experiencing greater lengths of risk deprivation did not realize the SM + PCIT benefit found among those with shorter or no risk deprivation ($p < .05$, HR per increase in log-transformed deprivation = 1.36). In other words, relative SM + PCIT benefits were most clearly concentrated among cases that had children in their home or had children returned to their home earlier rather than later. Risk-deprivation length was either not predictive in other design cells or was slightly beneficial in the other design cell conditions. Model estimated Cox regression curves for all four groups, estimated at the mean of all covariates and at a risk-deprivation length of zero, are shown in Figure 2. Figure 2 displays the expected survival if all participants were at risk for an event during the entire follow-up interval.

Discussion

Findings from this study support the comparative effectiveness of PCIT combined with the SM motivational orientation over SAU for reducing recidivism among maltreating parents in the child welfare system. These findings extend previous studies in three ways. First, the findings suggest that combining PCIT with a motivational orientation program may be key for obtaining better

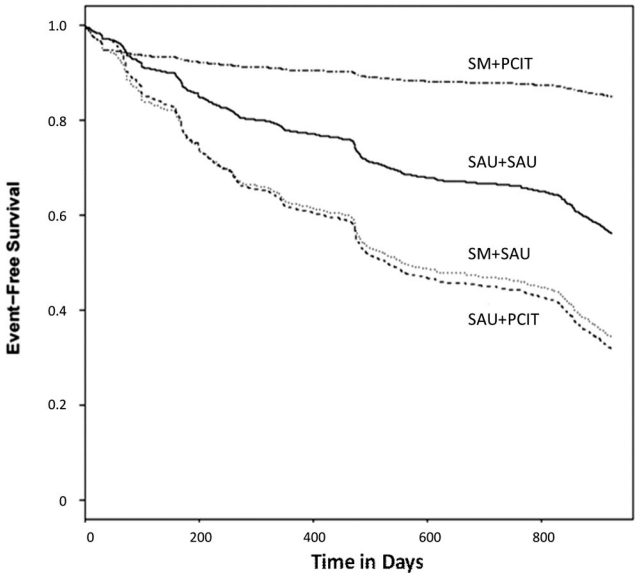

Figure 2. Imputed data set survival at zero risk deprivation and mean of all other covariates. SM = self-motivation; PCIT = parent–child interaction therapy; SAU = services as usual.

benefits with maltreating parents in child welfare. The dismantling design employed in the present study suggests that it is the combination of approaches, not either component main effect separately, that drives the comparative benefits of SM + PCIT found in the 2004 study. Implementing PCIT with a SAU orientation or the SM orientation with SAU parenting did not confer a relative advantage. Additional studies testing PCIT with no orientation or a placebo orientation versus SM might consolidate this conclusion.

Second, the study demonstrated that the favorable laboratory trial results for the SM + PCIT package can be replicated in a field setting. When considering how this finding might generalize to other implementations, readers should note that the training, technology transfer, and implementation quality control efforts used in the study were high compared with our normal field implementation methods but still decidedly short of what is typical for a laboratory trial or a development setting. Many of the laboratory resources present in the 2004 trial were not enjoyed by any intervention condition in the field trial, including transportation support for clients, greater capacity to accommodate client scheduling needs in the laboratory setting, special incentives, reliance on doctoral and postdoctoral psychology trainees as therapists, freedom from billing concerns, session-by-session implementation supervision, availability of on-site model expertise at all times, and far greater scrutiny of model fidelity.

Third, the findings suggest that the benefits reported previously among moderately chronic physical abuse cases can be extended to very chronic, more severe, and more mixed types of maltreatment cases. Sample differences between the two trials were related to shifts in county child welfare practice around the start of the field trial, as child welfare began directing some of their most chronic and severe cases to the parenting agency hosting the study. The numbers of past referrals (M = 6, median = 4), rate of foster placement (76%), rate of termination of rights (22%), and risk factor scores in our sample were high, perhaps even extraordinarily high. This more chronic and severe segment of the child welfare population is rightly viewed as challenging to serve, given their history of recidivism after multiple past intervention efforts.

On the basis of our familiarity with the services available to child welfare clients in the study community, we believe it is likely that these parents had received multiple parenting programs in the past, but they were unlikely to have previously received any evidence-based or behavioral parenting model. This might suggest tempered optimism that evidence-based behavioral parenting services can still yield benefits for parents even in the face of multiple past efforts. An additional observation supporting this optimism comes from the PCIT manipulation checks. In-clinic PCIT skills were acquired with reasonable consistency across cases irrespective of the parent's baseline skill level.

Parents who are chronically in child welfare can and do learn PCIT skills. Unfortunately, skill acquisition by itself did not uniformly translate into lower recidivism rates. The SAU + PCIT combination showed comparable in-clinic skill acquisition but relatively higher recidivism compared with SM + PCIT. Skill acquisition was a main effect of PCIT, whereas reduced recidivism was manifest only as an interaction effect with SM. We speculate that this may be a transfer of learning issue. Generalizing PCIT parenting skills acquired in the clinic to the home environment may have been facilitated by the motivational intervention. This is a bit different from our original rationale for designing a motivational complement to PCIT. The original rationale had to do with motivating parents to participate in the more behaviorally demanding evidence-based model and to learn the skills. Now, we suspect that motivation may be less necessary for participating in PCIT or learning skills but more important for transferring PCIT skills into the home environment and applying them with children. Studies examining in-home skill application would be a helpful test of this hypothesis.

There also was a significant three-way interaction between design cell conditions and risk-deprivation length. The comparative advantage of the SM + PCIT condition was strongest for parents with children in their home at baseline or when their children were returned sooner rather than later. This might reflect these parents' opportunity to practice PCIT skills better when they had children in the home and to generalize skills. Regular in-home skill practice is a normal PCIT requirement, and risk-deprived cases had limited opportunity to fulfill this requirement. Alternatively, because risk deprivation was much rarer among the SM + PCIT recipients, unobserved selection biases could have been operating, or there may be complex treatment moderation processes involving waning parent–child attachment.

Some limitations should be considered. As with any comparative outcome trial without a no-treatment or placebo control condition, it is important to keep in mind that this study examined comparative effectiveness of the interventions, not their absolute efficacy. Results were obtained at a single agency with a small number of part-time clinicians, so generalization to other settings should be made cautiously. The small number of therapists precluded any sound estimation of therapist effects, but counterbalancing therapists in the parenting conditions might partially mitigate this concern.

The main outcome of interest was maltreatment recurrence, which was measured solely by official records. Absolute recurrence rates are probably underascertained by official maltreatment

reports, although there is no evidence that detection bias might have differed across treatment conditions in a way that would affect randomized group comparisons. As with most studies of child welfare services, participants may have received additional procedures apart from the parenting program. These service factors could not be experimentally controlled. We might expect random allocation to distribute the other child welfare procedures evenly across conditions. The level of fidelity control exercised over the interventions in the study was less than what is typical for a laboratory trial, which might reduce internal validity. But this is both a limitation and a strength given that one aim of the study was to examine outcomes under precisely these sorts of more natural and less tightly controlled field conditions.

We would also point out potential limitations of Cox regression-based imputation procedures that include different covariates in the imputation and final analytic models (Faucett et al., 2002). Rubin (1996) and Meng (1994) referred to this situation as *uncongenial multiple imputation.* In our situation, we might be concerned with including proportional hazard predictors in the imputation model but not in our final analytic model. In theory, we would expect the coefficients of included covariates to be biased when predictive covariates are omitted (see Gail, Wieand, & Pantadosi, 1984; Bretagnolle & Huber-Carol, 1988; Struthers & Kalbfleisch, 1986). In practice, however, we would anticipate small differences in coefficients or inferential tests if omitted covariates predict poorly. Although results are not reported here for space purposes, this was true when we compared our multiple imputation coefficients with all or only the selected covariates.

In summary, support for using the SM + PCIT package with child welfare populations as an intervention for maltreating behavior is strengthened by these findings. The results suggest that field implementations of PCIT with child welfare populations should consider adding a motivational element. Support for the SM + PCIT package is extended to field settings and to more severe or chronic child welfare populations, and this latter extension appears particularly encouraging.

Adopters of the model may wish to consider possible clinical implications of the risk-deprivation interaction findings, including the possible importance of parenting skill practice opportunities in moderating SM + PCIT benefits. Future research testing child welfare recidivism outcomes among cases with significant out-of-home placement may need to consider the implications of conditional risk exposure. In many ways, these analytic challenges mirror those involved in managing compliance variations and other forms of postrandomization or posttreatment principal stratification (Jin & Rubin, 2008). Simulation studies of this phenomenon would prove useful in identifying optimal methods for recapturing true effects in the face of discrepant risk deprivation.

References

Andersen, P. K., & Gill, R. D. (1982). Cox's regression model for counting processes: A large sample study. *Annals of Statistics, 10,* 1100–1120. doi:10.1214/aos/1176345976

Barth, R. P., Landsverk, J., Chamberlain, P., Reid, J. B., Rolls, J. A., Hurlburt, M. S., . . . Kohl, P. L. (2005). Parent-training programs in child welfare services: Planning for a more evidence-based approach to serving biological parents. *Research on Social Work Practice, 15,* 353–371. doi:10.1177/1049731505276321

Bessmer, J. L. (1998). The Dyadic Parent-Child Interaction Coding System II (DPICS II): Reliability and validity. *Dissertation Abstracts International: Section B. Sciences and Engineering, 58*(7-B), 3961.

Bretagnolle, J., & Huber-Carol, C. (1988). Effects of omitting covariates in Cox's model for survival data. *Scandinavian Journal of Statistics, 15,* 125–138.

California Evidence-Based Clearinghouse for Child Welfare. (2009). *Usage guide for the CEBC.* Retrieved from http://www.cachildwelfareclearinghouse.org

Centers for Disease Control and Prevention, National Center for Injury Prevention and Control. (2004). *Using evidence-based parenting programs to advance CDC efforts in child maltreatment prevention: Research brief.* Atlanta, GA: Author.

Chaffin, M., Silovsky, J., Funderburk, B., Valle, L. A., Brestan, E. V., Balachova, T., . . . Bonner, B. L. (2004). Parent–child interaction therapy with physically abusive parents: Efficacy for reducing future abuse reports. *Journal of Consulting and Clinical Psychology, 72,* 500–510. doi:10.1037/0022-006X.72.3.500

Chaffin, M., & Valle, L. A. (2003). Dynamic predictive validity of the Child Abuse Potential Inventory. *Child Abuse & Neglect, 27,* 463–481. doi:10.1016/S0145-2134(03)00036-X

Chaffin, M., Valle, L. A., Funderburk, B., Gurwitch, R., Silovsky, J., Bard, D., . . . Kees, M. (2009). A motivational intervention can improve retention in PCIT for low-motivation child welfare clients. *Child Maltreatment, 14,* 356–368. doi:10.1177/1077559509332263

Eyberg, S. M., Bessmer, J., Newcomb, K., Edwards, D., & Robinson, E. (1994). *Dyadic Parent–Child Interaction Coding System—II manual.* Unpublished manuscript, University of Florida, Gainesville.

Eyberg, S. M., & Boggs, S. R. (1998). Parent–child interaction therapy: A psychosocial intervention for the treatment of young conduct-disordered children. In C. E. Schaefer & J. M. Briesmeister (Eds.), *Handbook of parent training: Parents as co-therapists for children's behavior problems* (2nd ed., pp. 61–97). New York, NY: Wiley.

Eyberg, S. M., Nelson, M. M., & Boggs, S. R. (2008). Evidence-based treatments for child and adolescent disruptive behavior disorders. *Journal of Clinical Child and Adolescent Psychology, 37,* 213–235.

Faucett, C. L., Schenker, N., & Taylor, J. M. G. (2002). Survival analysis using auxiliary variables via multiple imputation, with applications to AIDS clinical trial data. *Biometrics, 58,* 37–47. doi:10.1111/j.0006-341X.2002.00037.x

Feingold, A. (2009). Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis. *Psychological Methods, 14,* 43–53. doi:10.1037/a0014699

Gail, M. H., Wieand, S., & Pantadosi, S. (1984). Biased estimates of treatment effect in randomized experiments with nonlinear regressions and omitted covariates. *Biometrika, 71,* 431–444. doi:10.1093/biomet/71.3.431

Grambsch, P. M., & Therneau, T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika, 81,* 515–526. doi:10.1093/biomet/81.3.515

Hakman, M., Chaffin, M., Funderburk, B., & Silovsky, J. F. (2009). Change trajectories for parent–child interaction sequences during parent–child interaction therapy for child physical abuse. *Child Abuse & Neglect, 33,* 461–470. doi:10.1016/j.chiabu.2008.08.003

Jin, H., & Rubin, D. (2008). Principal stratification for causal inference with extended partial compliance. *Journal of the American Statistical Association, 103,* 101–111. doi:10.1198/016214507000000347

Kaminski, J. W., Valle, L. A., Filene, J. H., & Boyle, C. L. (2008). A meta-analytic review of components associated with parent training program effectiveness. *Journal of Abnormal Child Psychology, 36,* 567–589. doi:10.1007/s10802-007-9201-9

Kaufman, A. S., & Kaufman, N. L. (1990). *Kaufman Brief Intelligence Test manual.* Circle Pines, MN: American Guidance Service.

Lee, S., Aos, S., & Miller, M. (2008). *Evidence-based programs to prevent children entering and remaining in the child welfare system* (Document

08–07-3901). Olympia, WA: Washington State Institute for Public Policy.

Meng, X.-L. (1994). Multiple-imputation inferences with uncongenial sources of input. *Statistical Science, 9,* 538–558.

Miller, W. R., & Rollnick, S. (1991). *Motivational interviewing: Preparing people to change addictive behavior.* New York, NY: Guilford Press.

Milner, J. S. (1986). *The Child Abuse Potential Inventory: Manual* (2nd ed.). DeKalb, IL: Psytec.

Milner, J. S. (1994). Assessing physical child abuse risk: The Child Abuse Potential Inventory. *Clinical Psychology Review, 14,* 547–583. doi:10.1016/0272-7358(94)90017-5

Mullins, S. M., Suarez, M., Ondersma, S. J., & Page, M. C. (2004). The impact of motivational interviewing on substance abuse treatment retention: A randomized control trial of women involved with child welfare. *Journal of Substance Abuse Treatment, 27,* 51–58. doi:10.1016/j.jsat.2004.03.010

NSCAW Research Group. (2005). *National survey of child and adolescent well-being (NSCAW) CPS sample component wave 1 data analysis report, April 2005.* Washington, DC: U.S. Department of Health and Human Services, Administration for Children, Youth and Families.

Patterson, G. R. (1976). The aggressive child: Victim and architect of a coercive system. In E. J. Mash, L. A. Hamerlynck, & L. C. Handy (Eds.), *Behavior modification and families* (pp. 267–316). New York, NY: Brunner/Mazel.

Patterson, G. R. (1982). *Coercive family process.* Eugene, OR: Castalia.

Patterson, G. R., & Reid, J. B. (1984). Social interaction processes within the family: The study of the moment to moment transactions in which human social development is embedded. *Journal of Applied Developmental Psychology, 5,* 237–262. doi:10.1016/0193-3973(84)90021-2

Pinkston, E. M., & Smith, M. D. (1998). Contributions of parent training to child welfare: Early history and current thoughts. In J. R. Lutzker (Ed.), *Handbook of child abuse research and treatment* (pp. 377–399). New York, NY: Plenum Press.

Robinson, E. A., & Eyberg, S. M. (1981). The dyadic parent–child interaction coding system: Standardization and validation. *Journal of Consulting and Clinical Psychology, 49,* 245–250. doi:10.1037/0022-006X.49.2.245

Rubin, D. B. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association, 91,* 473–489. doi:10.2307/2291635

Rubin, D. B. (2004). *Multiple imputation for nonresponse in surveys.* New York, NY: Wiley.

Stith, S. M., Liu, T., Davies, L. C., Boykin, E. L., Alder, M. C., Harris, J. M., . . . Dees, J. E. (2009). Risk factors in child maltreatment: A meta-analytic review of the literature. *Aggression and Violent Behavior, 14,* 13–29. doi:10.1016/j.avb.2006.03.006

Struthers, C. A., & Kalbfleisch, J. D. (1986). Misspecified proportional hazard models. *Biometrika, 73,* 363–369. doi:10.1093/biomet/73.2.363

U.S. Department of Health and Human Services, Administration for Children, Youth and Families, Children's Bureau. (2010). *Child maltreatment 2008.* Retrieved from http://www.acf.hhs.gov/programs/cb/stats_research/index.htm#can

Urquiza, A. J., & McNeil, C. B. (1996). Parent–child interaction therapy: An intensive dyadic intervention for physically abusive families. *Child Maltreatment, 1,* 134–144. doi:10.1177/1077559596001002005

van Buuren, S. (2007). Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research, 16,* 219–242. doi:10.1177/0962280206074463

van Buuren, S., & Oudshoorn, C. C. M. (2000). *Multivariate imputation by chained equations: MICE V1.0 user's manual* (Report PG/VGZ/00.038). Leiden, the Netherlands: TNO Prevention and Health.

Weisz, J. R., Weiss, B., & Donenberg, G. R. (1992). The lab versus the clinic: Effects of child and adolescent psychotherapy. *American Psychologist, 47,* 1578–1585. doi:10.1037/0003-066X.47.12.1578

Wells, K. B. (1999). Treatment research at the crossroads: The scientific interface of clinical trials and effectiveness research. *American Journal of Psychiatry, 156,* 5–10.

Received March 22, 2010
Revision received July 19, 2010
Accepted July 30, 2010 ■