Journal of Clinical Child and Adolescent Psychology
2007, Vol. 36, No. 3, 378–391

Treatment for Preschool Children With Interpersonal Sexual Behavior Problems: A Pilot Study

Jane F. Silovsky
Department of Pediatrics, University of Oklahoma Health Sciences Center

Larissa Niec
Department of Psychology, Central Michigan University

David Bard and Debra B. Hecht
Department of Pediatrics, University of Oklahoma Health Sciences Center

This pilot study evaluated a 12-week group treatment program for preschool children with interpersonal sexual behavior problems (SBP; N = 85; 53 completed at least 8 sessions). Many children presented with co-occurring trauma symptoms and disruptive behaviors. In intent-to-treat analysis, a significant linear reduction in SBP due to number of treatment sessions attended was found, an effect that was independent of linear reductions affiliated with elapsed time. Under the assumption that treatment can have an incremental impact, more than one third of the variance was accounted for by treatment effects, with female and older children most favorably impacted. Caregivers reported increase in knowledge, satisfaction, and usefulness of treatment. In addition to replication, future research is needed to examine (a) effects of environment change and time on SBP, (b) stability of treatment effects, and (c) best practices to integrate evidence-based treatments for comorbid conditions.

Juvenile courts and child protective services have experienced growing caseloads of children younger than 12 who have been sexually aggressive (Araji, 1997). Concurrently, clinical child psychologists are more likely then ever to encounter sexual behavior problems (SBP) as a treatment issue. Although the literature on childhood SBP has expanded in the last decade, gaps in treatment outcome research are notable, particularly for preschool children (Association for the Treatment of Sexual Abusers, 2004).

Child interpersonal SBP are sexual behaviors (behaviors that involve sexual body parts, i.e., genitals, anus, buttocks, or breasts) that are developmentally inappropriate or potentially harmful and are initiated by a child toward another (Silovsky & Bonner, 2003). Interpersonal SBP are distinguished from normal childhood sexual play and exploration (i.e., sexual behaviors that occur spontaneously and intermittently, are mutual and noncoercive, and do not cause emotional distress; Chaffin et al., 2006). Preschool children have been found to demonstrate interpersonal SBP, including attempting sexual intercourse and oral–genital contact (Friedrich & Luecke, 1988; Johnson, 1988; Silovsky & Niec, 2002). The primary motivation driving childhood SBP is not necessarily sexual gratification and instead may be related to curiosity, anxiety, imitation, attention seeking, self-calming, or other reasons (Silovsky & Bonner, 2003).

SBP are deviations from the normal course of sexual development (Araji, 1997). Social context, individual characteristics, disruptive experiences, and the interactions of these factors impact the course of sexual development (Araji, 1997). Sexual abuse is one type of disruptive experience impacting sexual development. Children who have been

This study was supported in part by a grant to Jane Silovsky from the Children's Medical Research Institute, Oklahoma City, Oklahoma. Clinical services were supported through a contract with the Oklahoma Department of Human Services. We acknowledge the contributions of Barbara L. Bonner, C. Eugene Walker, William N. Friedrich, and Mark Chaffin to this project. We also express our sincere thanks to the families, therapists, and research assistants involved with this project.

Correspondence should be sent to Jane F. Silovsky, Department of Pediatrics, University of Oklahoma Health Sciences Center, P.O. Box 26901, Oklahoma City, OK 73190. E-mail: jane-silovsky@ouhsc.edu

sexually abused are more likely to demonstrate SBP than children without such a history (Friedrich et al., 2001). Sexually abused preschool children appear to be at particularly high risk for developing SBP. In a meta-analysis of studies of sexually abused children, 36% of preschoolers were found to have SBP compared to 6% of school-age children (Kendall-Tackett, Williams, & Finkelhor, 1993).

Although sexual abuse impacts the development of SBP, many children with SBP have no known history of sexual abuse (Bonner & Fahey, 1998; Bonner, Walker, & Berliner, 1999; Johnson, 1988; Silovsky & Niec, 2002). The development of SBP appears to have multiple origins. Parenting practices, exposure to sexual material, absent or disrupted attachments, exposure to family violence, physical abuse, and the development of nonsexual behavior problems contribute to the development and maintenance of childhood SBP (Friedrich, 2002; Friedrich, Davies, Feher, & Wright, 2004; Gray, Pithers, Busconi, & Houchens, 1999).

Regardless of the causal pathway, demonstrating SBP as a young child is associated with a variety of problems in adjustment and development. Children with SBP often exhibit other behavior problems (Bonner et al., 1999; Friedrich & Luecke, 1988; Gray et al., 1999). Poor impulse-control skills, aggressive behaviors, and inaccurate perceptions of social stimuli hinder social relationships and cause problems at school (Araji, 1997; Bonner et al., 1999; Friedrich & Luecke, 1988; Gil & Johnson, 1993; Horton, 1996). Socialization difficulties and stigmatizing responses from peers and adults may impede developing self-concepts (Heiman, 2001). Poor boundaries and indiscriminate friendliness may increase risk of future victimization (Pearce, 2003; Silovsky & Niec, 2002). Further, children with SBP are at risk of separation from parents and of placement disruptions (Baker, Schneiderman, & Parker, 2001; Bonner et al., 1999; McKenzie, English, & Henderson, 1987; Silovsky & Niec, 2002). The impact of SBP is potentially wide ranging, and thus the importance of intervention is considerable.

Despite the need for efficacious interventions, treatments for preschool children with SBP that account for diverse etiological factors and address multiple areas have not yet been systematically evaluated. In response to this need, we established a two-pronged research program to extend the knowledge about the characteristics of preschool children with SBP and evaluate an intervention based on existing research. This study focused on the second prong by piloting a group treatment program designed to reduce SBP in preschool children. Three areas of treatment-outcome research informed our development of the treatment model: (a) sexual abuse treatment for young children, (b) treatment of school-age children with SBP, and (c) treatment for disruptive behavior disorders in preschool children.

Sexual abuse treatments for preschool children have included components that directly address SBP. Pre- to posttreatment reductions in SBP have been found in evaluations of cognitive-behavioral (CBT) abuse focused group therapy for sexually abused children with concurrent treatment for nonoffending parents (Hall-Marley & Damon, 1993; Stauffer & Deblinger, 1996). Reductions in SBP were not found during an initial wait period (Stauffer & Deblinger, 1996). In a randomized controlled study with sexually abused preschoolers, Cohen and Mannarino (1996, 1997) found CBT more efficacious in reducing SBP than nondirective supportive treatment and improvements were maintained at 1-year follow up (Cohen & Mannarino, 1997). In contrast to these findings, sexual abuse treatments that do not include components addressing SBP have not been found to reduce SBP (Finkelhor & Berliner, 1995).

Although treatments for sexual abuse have been found to reduce SBP in young children, such abuse-focused treatment would be inappropriate for children with no history of sexual abuse. Treatments specifically for SBP have been developed (Araji, 1997) and evaluated with two randomized controlled trials of SBP treatment for school-age children (Bonner et al., 1999; Pithers & Gray, 1993). Both studies randomly assigned over 100 families each to one of two group treatment programs. Bonner et al. (1999) compared CBT group treatment program to a play therapy group. The other study compared an expressive therapy to a relapse prevention program (Pithers & Gray, 1993; Pithers, Gray, Busconi, & Houchens, 1998). Pithers et al. found that a subgroup of the children (Highly Traumatized) demonstrated greater benefit from the relapse prevention treatment at midtreatment. Posttreatment reductions in SBP were found regardless of treatment modality in both randomized trials (Bonner et al., 1999; Pithers and colleagues' results as cited in Bonner & Fahey, 1998). However, a 10-year follow-up of Bonner et al.'s (1999) participants that included a clinic comparison group utilizing administrative data found the CBT group had significantly less sex offenses than the play therapy group (2% vs. 10%), and the CBT group did not differ from the general clinic comparison (3% Carpentier, Silovsky, & Chaffin, 2006).

One consistency across all SBP interventions has been the direct involvement of the caregivers in the treatment. Support for the involvement of

parents in the treatment of children with SBP is also found in the literature on disruptive behavior disorders (Brestan & Eyberg, 1998). SBP are similar to other behavior problems in a variety of ways: The behaviors involve behavioral disinhibition and social rule breaking, and they may include aggressive acts toward self or others. Further overlap is found in the factors that contribute to the development and maintenance of both disruptive behavior problems and SBP (e.g., history of violence exposure, poor supervision/monitoring, negative parent–child relationship; Burke, Loeber, & Birmaher, 2002; Friedrich et al., 2004). Research on treatment for disruptive behaviors has consistently identified behavior management training as an effective modality (see Brestan & Eyberg, 1998; Nixon, 2002).

Drawing from the existent literature, several factors emerge as important for treatment for preschool children with SBP. Treatment needs to (a) directly address SBP using behavioral, CBT, and psychoeducational approaches; (b) address the child's social problems, impulse-control, coping strategies, boundary issues, and caregiver–child relationship; (c) directly involve the parent/caregiver in treatment; and (d) include a behavior management training component. The treatment must be broad enough to address the needs of children with SBP who do not have a history of sexual abuse. Further, treatment should be developmentally sensitive and consider the cognitive, emotional, and behavioral capacities of young children.

Our study investigated rates of SBP prior to, during, and after intervention with preschoolers who have SBP. To mitigate the effects of history and maturation inherent in a single group pre–post design, we evaluated SBP during the natural wait period between referral and the start of the next treatment group. We predicted that levels of SBP would not change over the wait period and would significantly decrease from pre- to posttreatment. Trauma history, including sexual abuse history, and evidence of other externalizing and internalizing behavior problems were evaluated to inform treatment specificity and treatment development.

Method

Participants

Participants were referred to this study primarily from the state's child protective services (CPS) agency or related programs (44%) and from other mental health service providers (38%). Other sources of referral were self-referral (8%), medical professional (2%), or other or unknown (8%). Over the 3-year recruitment period, 155 children were referred. Only children between 3 and 7 years of age at intake were included (two excluded due to being too young). Children were classified as having a sexual behavior problem by caregiver report on the Child Sexual Behavior Inventory–III (CSBI–III; a significant T score > 65 on either the Total Sexual Behavior Problem Scale or the Sexual Abuse Specific subscale; Friedrich, 1997); 25 were excluded due to not having SBP (typically these were sexually abused children with no SBP). The sexual behavior exhibited had to include interpersonal sexual behaviors and not be exclusively problematic self-touch sexual behavior. Exclusion criteria were as follows: (a) demonstration of severe psychiatric or developmental problems (none excluded), (b) caregiver could not accurately complete the assessment forms (e.g., for developmental, language, or other reasons; one excluded), and (c) caregiver was a known or suspected perpetrator of child sexual abuse (none excluded). Twenty-two families could not be reached from the information provided in the referral, 10 lived too far away to participate in services, and 10 were not interested in services or did not participate for other reasons (these 42 children did not differ in age or gender from those who participated). Eighty-five families met inclusion and exclusion criteria, and all consented to participate in the study. No other agency in the metropolitan area provides specialized services for children with sexual behavior problems. The sample of children appears to be representative of clinic-referred young children with interpersonal SBP in the community, as there was a range of referral sources (including other mental health facilities), there were no other agencies specialized in this area, and all the qualified families agreed to participate in the study.

Measures

Child sexual behavior inventory–III. The CSBI–III is a caregiver-completed instrument that assesses the presence and frequency of a range of child sexual behaviors (Friedrich, 1997). Norms for children 2 to 12 years of age are provided, separated by age groups (2–5, 6–9, and 10–12) and gender. The CSBI–III yields three standardized scores: CSBI Total and two subscales (Developmentally Related Sexual Behavior and Sexual Abuse Specific Items scales). The manual provides support for the CSBI–III internal consistency (.72 for normative sample and .92 for sexually abused sample); stability (.91); and convergent, discriminant, and construct validity. The CSBI–III has demonstrated sensitivity to treatment effects (e.g., Cohen & Mannarino, 1996, 1997). Additional items assess children's exposure to (a) stressful events (e.g., physical abuse,

parent's divorce), (b) sexual behavior of adults, and (c) sexual information and materials. Nine items, added since the third revision of this inventory, assess aggressive and/or coercive sexual behaviors as well as withdrawal and avoidance of sexual content. In a normative sample, none of the items assessing aggressive sexual behaviors were endorsed (Friedrich, 2002).

The internal consistency of the CSBI–III was further supported with a sample of 77 preschool age children referred to our treatment clinic (.89). The standard scores on the CSBI–III Total Score at intake were highly skewed, because the majority of the children's scores were at the 99th percentile, whereas the raw scores were more normally distributed (CSBI–III Total Standard Scores: $Mdn = 110$, $M = 99.8$, $SD = 15.3$; CSBI–III Total raw scores: $Mdn = 31.5$, $M = 32.7$, $SD = 12.9$). Therefore, the raw scores for the CSBI–III were used in the analyses.

Child behavior checklist. The Child Behavior Checklist (CBCL; 4–18 years) is a 134-item standardized checklist of childhood behavior problems and social competence (Achenbach, 1991) completed by parents. The CBCL has been used extensively in clinical research with children and has demonstrated internal consistency (.95), stability (one week; Total Problems scale = .93), and construct validity (Achenbach, 1991). For the purpose of analysis, the Total Problems scale was used in this study. Scores greater than 65 are considered clinically significant.

Diagnostic interview schedule for children. The Diagnostic Interview Schedule for Children (DISC–IV; National Institute of Mental Health, 1998) is a structured diagnostic interview based on the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; American Psychiatric Association, 2000) criteria of childhood mental disorders. It was developed for use by nonclinicians for screening and research purposes. Test stability and internal consistency and validity of parent report on school-age children have been supported, although information about the posttraumatic stress disorder (PTSD) scale was not reported (Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000). The DISC–IV was chosen to provide a more detailed evaluation of PTSD and major depressive disorder (MDD). DISC–IV data used in analyses included whether the child met criteria for PTSD or MDD and the number of symptoms reported for each diagnosis. At intake, the number of PTSD and MDD symptoms positively correlated with internalizing symptoms on the CBCL (.29 and .27, respectively, $p < .05$).

Abuse dimensions inventory. The Abuse Dimensions Inventory was designed to record and classify child abuse victimization experiences (Chaffin, Wherry, Newlin, Crutchfield, & Dykman, 1997). Interrater reliability (Kappa) of the measure has ranged from .65 to 1.00 on individual scales (Chaffin et al., 1997). For our study, data on whether sexual abuse was suspect and/or confirmed were included.

Peabody picture vocabulary test–third edition. The Peabody Picture Vocabulary Test–Third Edition (PPVT–III) is an individually administered measure designed to assess children's receptive vocabulary (Dunn & Dunn, 1981). The PPVT–III has demonstrated adequate reliability and validity. It has been found to significantly correlate with the Wechsler Intelligence Scale for Children-Third Edition Full Scale score ($r = .90$; Williams & Wang, 1997). The PPVT–III was administered to obtain an estimate of the children's receptive vocabulary for descriptive purposes.

Parenting stress index–short form. The Parenting Stress Index–Short Form (PSI–SF) is a 36-item parent self-report instrument designed to measure the relative degree of stress in a parent–child relationship and to identify the sources of distress (Abidin, 1995). Scores above the 85th percentile are considered to be clinically significant. The PSI–SF has been found to correlate high with the long form of the PSI (.87; Abidin, 1995). Haskett, Ahern, Ward, and Allaire (2006) found adequate 1-year test–retest stability (.61, $p < .005$) as well as support for internal consistency ($\alpha = .78$ and .91) and predictive validity. The PSI–SF was administered to assess the initial level of and treatment impact on parenting stress.

Treatment history form. The Treatment History Form was designed by the investigator and used to record type and timing of services provided to the child and family outside of the studied group treatment program. Services tracked included clinic, home, and school-based treatment for the child, parent, and family.

CSBP preschool group satisfaction and social validity questionnaire. The CSBP Preschool Group Satisfaction and Social Validity Questionnaire

(CSBP–PGSQ) is a 38-item measure of caregiver satisfaction and social validity, modified from the Child Project SAFE Evaluation Form used to evaluate a sexual abuse treatment program (Futa, Hecht, & Hansen, 1998). Caregivers rate specific aspects of the group; knowledge before and after group on the topics addressed; helpfulness of topics; qualities of the therapists; impact of the group on their child, self, and family; and overall satisfaction at the end of the treatment program. Internal consistency for the subscales of Group Components, Knowledge, Helpfulness, Therapist Qualities, and Impact ranged from .79 to .94.

Treatment Protocol

The Group Treatment for Preschool Children with Sexual Behavior Problems is a manualized 12-session treatment protocol that uses behavioral, cognitive-behavioral, and psychoeducational approaches in parallel child and caregiver groups, with time during each session in which the children and caregivers are together (Silovsky & Niec, 1998). The treatment focuses on eliminating SBP and replacing these behaviors with prosocial behaviors and coping skills. The targets of the intervention are child behaviors (e.g., maintaining physical boundaries, impulse control, social behavior), caregiver behaviors (e.g., those that inadvertently reinforce SBP and other behavior problems), child cognitions (e.g., maladaptive beliefs about appropriate touch), caregiver cognitions (e.g., beliefs that the child will become a pedophile), and the quality of the child-caregiver relationship.

The children's group addresses (a) body awareness and "safe" and "unsafe" touching, (b) maintaining physical boundaries, (c) relaxation, (d) impulse control skills, (e) abuse prevention skills, and (f) feeling identification and expression skills. These areas are taught and practiced with developmentally appropriate activities (e.g., singing, coloring, puppet play, games) to facilitate learning concepts and practicing skills. The caregivers' treatment is designed to increase awareness of the factors that may be related to SBP (e.g., exposure to sexualized material, sexual abuse, other trauma) and develop new family rules and patterns of interactions. Caregivers are taught about sexual development and SBP, methods of responding to and preventing sexual behavior, strategies to enhance parent–child interactions, and behavior management techniques consistent with treatments with demonstrated efficacy (i.e., parent–child interaction therapy [PCIT]; Hembree-Kigin & McNeil, 1996). The caregivers group includes time to facilitate support and give opportunities for active learning.

Each session includes 20 to 30 min at the end of the 90-min session for group family work to practice skills under the guidance of the therapists to facilitate changing patterns of interactions. For eight sessions during group family work, the children describe, demonstrate, and practice skills, and the caregivers are encouraged to support the use of the skills with modeling and home practice. For four sessions the caregivers practice behavior management and relationship-building skills with their children while being observed and receiving feedback from the therapists. For all the joint activities, therapists model behavior management strategies (e.g., praising the children, providing reinforcement for appropriate behaviors, using selective attention) and provide the caregivers with feedback.

The group treatment program is closed-group 12-session format in which session material builds on earlier treatment sessions. If families missed a session (e.g., due to illness, vacation, transportation issues, etc.), the caregiver and child were provided the session material individually, typically before the next session by meeting with the therapists.

Therapists, training, and treatment integrity. Children groups were lead by a therapist with a master's degree or a doctorate in psychology, with training and experience implementing behavioral and psychoeducational treatments for children. Depending on the size of the children's group (3 to 7 members), an additional one to three cotherapists assisted with the group. Cotherapists were primarily graduate-level practicum students. The first author (a licensed doctoral-level clinical psychologist) observed most of the children's group sessions and joint parts of the sessions through a one-way mirror and provided weekly supervision and training to ensure adherence with the treatment manual. The caregivers' group was led by either the second or fourth author, both doctoral-level clinical psychologists. Adherence to the treatment manual was reviewed with the first author before and after each session.

Procedures

Approval from the university's Institutional Review Board was obtained. Family services were not dependent on their participation in the research. All caregivers of children referred to the Preschool Program for Children with Sexual Behavior Problems were approached and agreed

to participate in the research. Informed consent was obtained from the legal guardian. CPS's investigations and protective acts (e.g., placement in foster care) had to be completed prior to inclusion in the program. Consent to contact CPS was obtained before completing the Abuse Dimensions Inventory through an interview with the CPS worker. Information on the children's trauma histories and other prior and concurrent treatment was collected but did not impact inclusion in the study, because this was a pilot study and additional services were often required by CPS or the placement.

The primary caregiver and child participated in a 3-hr intake evaluation. During the intake evaluation, the caregiver completed an assessment battery (CSBI–III, CBCL, DISC–IV, and PSI–SF) and the PPVT–III was administered to the child. The wait period varied depending on the timing of the referral to the start of the next group and ranged from 0 to 16 weeks ($M = 6.5$, $SD = 4.6$). The families did not receive any financial reimbursements for their participation in the research.

The assessment battery (except the DISC–IV) was readministered to the caregiver at pretreatment (unless the intake was within 3 weeks of the first session), and the entire battery was completed at posttreatment. For this pilot study, the DISC–IV was not readministered at pretreatment for two reasons: It assesses the presence of symptoms over a longer period than many children spent on the wait list and therefore would not be sensitive to change during the wait time, and time constraints prevented the readministration of such a lengthy measure at the first session. At posttreatment, the CSBP–PGSQ was completed with no identifying information and placed in an unmarked envelope, to reduce the impact of social desirability when providing social validity feedback.

Data Analyses

To evaluate treatment effects, repeated measures analysis of variance (ANOVA) was planned with CSBI–III scores as dependent variables (intake, pretreatment, and posttreatment) while adjusting for total number of study weeks. The impact of total stressful events and of concurrent treatment was tested as a main effect and as an interaction with CSBI–III over time. Examination of specificity of treatment effects was planned, using repeated measures analysis for PSI–SF, CBCL, and DISC–IV.

Our main analysis, an intent-to-treat multilevel analysis, was constructed to explore the trend in SBP (again, modeling CSBI–III scores) during the random wait period and treatment intervention time segments. The intent-to-treat models constructed assessed separate effects due to treatment and time (time since intake) under various assumptions about the impact of each.

Results

Intake Data

Demographic and clinical information of 85 children meeting inclusion and exclusion criteria and their caregivers are provided in Table 1. Most of the caregivers were female, with the exception of 3 biological fathers, 2 adoptive fathers, and 1 foster father. The median annual income was around $25,000, ranging from less than $10,000 to more than $80,000. For children who were not

Table 1. *Demographic and Descriptive Information on Participants at Intake*

Age (*M*, *SD*)	4.9	1.1
PPVT–III (*M*, *SD*)	89.5	13.1
Total Stress Events (*M*, *SD*)	4.7	2.1
Child Female Gender	49	58%
Child's Race Minority	28	33%
Caregiver's Relationship		
Biological Parent	35	41%
Foster Parent	26	31%
Other Female Relative	15	18%
Adoptive Parent	5	6%
Step-Parent	4	5%
Sexual Abuse of Child		
Confirmed	30	35%
None/Ruled Out	28	33%
Uncertain Findings	14	16%
Unknown	13	15%

Note: $N = 85$. Values are *n* and % unless noted. PPVT–III = Peabody Picture Vocabulary Test–Third Edition.

living with their biological parent, the average number of months the child lived in the home was 19 months ($SD = 19$, range = 1–60) though the caregivers typically knew the child longer (how long caregiver knew child: $M = 28$ months, SD = 25, range = 1–72).

The children's mean CSBI–III *t* score was greater than the 99th percentile ($M = 98.9$, $SD = 14.6$). All 85 children demonstrated interpersonal sexual behaviors, and many had sexually aggressive behaviors. Twenty-seven children (32%) tried to undress other children against their will, and 18 (16%) were reported to force other children to do sexual acts.

Many of the children had experienced multiple stressful/traumatic events (rates of specific types of events are similar to Silovsky & Niec, 2002). The children experienced an average of almost five stressful events, and 47 (55%) children responded to traumatic events with intense fear, helplessness, horror, or disorganization by caregiver report (i.e., they met the DISC–IV–TR Criteria A for PTSD diagnosis; American Psychiatric Association, 2000). These children averaged 5.1 ($SD = 3.6$) PTSD symptoms (Summing Criteria B—reexperiencing of trauma symptoms, C—avoidance symptoms, and D—arousal symptoms; American Psychiatric Association, 2000), and 6 (13%) children met full criteria for PTSD diagnosis at intake. The average number of depression symptoms was 2.7 ($SD = 2.2$, range = 0–8), and 17 (20%) children met criteria for MDD diagnosis at intake.

Thirty (35%) of the children attended individual therapy (typically services required for their therapeutic foster care placement), 3 families attended PCIT, and 4 caregivers attended individual therapy. None of the children or families began other treatment at or around the same time that the SBP treatment was initiated, and 98% had started the concurrent treatment prior to the intake evaluation.

Treatment Outcome Results

Fifty-three families completed at least 8 sessions of treatment, and typically the content of missed sessions was addressed individually. Thus, these families often were provided the content of all 12 sessions. Twelve families either dropped out after starting treatment or missed more than 4 sessions, 16 never started group, and 4 moved (see Figure 1). Families who completed group were compared to those who did not, with no differences found in demographic characteristics or intake CSBI–III, CBCL, PSI, or DISC–IV scores. The means and standard deviations for the CSBI–III scores at intake, pretreatment, and posttreatment are provided in Table 2.

The software used for the repeated measures ANOVA analysis required complete data at all three waves of data collection. Unfortunately, 16 families did not complete a pretreatment assessment because their wait period was within 3 weeks of intake. To avoid the possibility of fatigue and "practice" effects of a full CSBI–III administration at pretest for these families, the pretreatment battery of tests were not administered to families whose intake was within 3 weeks of the start of the next group. An additional 6 families failed to complete posttreatment measures, leaving 31 families for the repeated measures ANOVA analysis. The CSBI–III within-subject effect was significant, $F(2, 28) = 4.06$, $p < .03$, after adjusting for total number of study weeks, and all Sidak-adjusted pairwise contrasts of intake, pre- and posttest scores (Ms = 33.2, 27.2, and 15.3, respectively, among the study completers) were significant ($\alpha = .05$), indicating decreasing scores over time.[1] At posttreatment, 36% of the children's CSBI–III scores were below the clinical range (<65). Total Stress Events and Concurrent Treatment main effects and interactions with CSBI–III were not significant. However, the interaction of CSBI–III with the number of study weeks (from intake to posttest) was significant, $F(2, 28) = 3.51$, $p < .05$, suggesting length of time in the study was related to the magnitude of differences that existed between the three assessments. To explore the mechanism relating time to these assessment differences, we turned our attention to the intent-to-treat analysis.

The significant interaction between study weeks and the within CSBI–III factor is a complex and intriguing result, implicating numerous time-structured effects. For example, it is possible this effect signals a time-related difference in either (or both) contrast between intake and pretest or pretest and posttest. Given the longitudinal structure of data collection and the time-varying implementation of treatment, we used the intent-to-treat, multilevel modeling technique to explore these possibilities by attempting to disentangle effects specific to time and treatment. The software used for multilevel modeling also allowed us to utilize all existing data, handling planned missingness (exclusion of pretests from 16 families) and attrition through likelihood and multiple imputation procedures (see Collins, Schafer, &

[1]Similar results emerged when considering intake scores as pretests for the 14 individuals who started treatment within 3 weeks of intake. The pretest and posttest means for this sample of 45 individuals were 29.11 and 16.27, respectively.

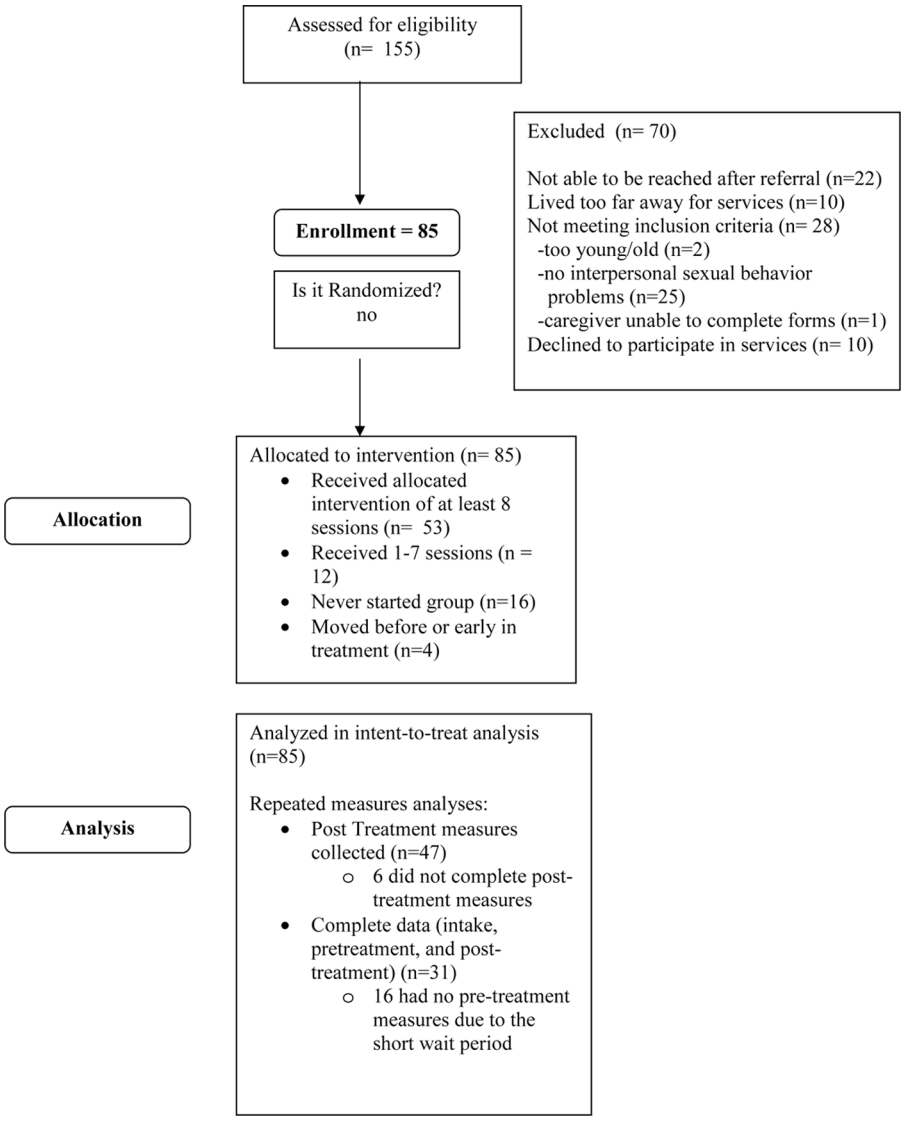

Figure 1. Flow chart for recruitment and participation.

Kam, 2001, for comparison and description of these methods).

To begin the intent-to-treat analysis, similar to Little and Yau (1996), missing CSBI–III data were imputed under a variety of assumptions about the effects of treatment and time. We report the results from two extreme scenarios: a worst-case completers-only treatment effect and a best-case partial treatment impact situation. In both scenarios, the imputed outcomes (CSBI–III) were predicted by

Table 2. *Means and Standard Deviations for Outcome Measures at Intake, Pretreatment, and Posttreatment*

	Intake[a]		Pre-Tx[b]		Post-Tx[c]			
	M	*SD*	*M*	*SD*	*M*	*SD*	F^d	*p*
CSBI–III								
Total (Standard)	98.4	14.9	95.9	17.9	75.0	19.7		
Total (Raw)	31.7	13.7	28.6	13.2	16.2	10.2	$F(2, 28) = 4.06$	$< .03$
PSI–SF Total	76.5	25.7	78.6	25.3	71.0	30.0	$F(2, 27) = 1.87$	.17
CBCL Total Problems Scale	70.0	10.0	70.8	8.0	65.9	9.9	$F(2, 24) = 1.35$	.28

Note: Pre-tx = Pretreatment; Post-tx = Posttreatment; CSBI–III-Child Sexual Behavior Inventory–III;PSI–SF = Parenting Stress Index–Short Form; CBCL = Child Behavior Checklist.
[a]$N = 85$. [b]$n = 40$. [c]$n = 47$. [d]Includes only participants with all three waves of data. Adjusted for total number of study weeks.

two linear piecewise effects, one each for treatment and time. In other words, these effects represent constant increases or decreases in the outcome, per unit increase in the predictor, which are active for a specified portion of the study observation period. For time, the influence was active throughout the entire participation interval (representing constant spontaneous or nontreatment-related improvement in the outcome, per week involved); for treatment, the effect was active in the model (allowed to predict the outcome) for the weeks of therapeutic intervention but only for those specified in the different case scenarios. For the worst-case scenario, outcomes for individuals not completing treatment and consequently missing one or both pre- and posttest CSBI–III scores were imputed assuming no impact whatsoever from the linear treatment effect (only the time effect was allowed to predict their missing data). In the second (best-case) scenario, as long as an individual completed one or more sessions of treatment, imputation of missing data was allowed to incorporate unique linear effects of the treatment impact.

To aid prediction in both scenarios, the imputation procedure began with a likelihood-based exploration of covariates that might explain variation in CSBI–III scores and possibly relate to missingness. Among age at study entry (in years), suspected prior sexual abuse (yes or no), gender, and ethnicity (Caucasian vs. other), only age and gender significantly explained variation in either the treatment or time effect. No covariate effects reached significance for the estimation of initial SBP status (i.e., model-estimated baseline CSBI–III scores). When assessing the time-related influence trajectory, after controlling for a significant piecewise treatment fixed-effect, only a significant linear time effect could be detected (i.e., no quadratic time trend evident). Lack of posttest data among noncompleters prevented meaningful exploration of anything but a linear treatment effect. Once data were imputed (six imputed datasets per scenario) from these covariate and piecewise growth models, a new multilevel growth model was constructed to test the beneficial (or detrimental) contributions of time passage and treatment simultaneously in each scenario. Further details on the imputation procedure are provided in the appendix.

The new multilevel growth models applied to the worst-and best-case imputed datasets were identical in structure. This final model included two Level 1 predictors, piecewise linear effects for time and treatment, four Level 2 predictors, age and gender influences on each piecewise component, and three Level 2 intercepts, average initial status and average time and treatment impact. The piecewise treatment effect in these analyses was estimated using the same design matrix as in the best-case imputation scenario (i.e., allowed for a linear effect of number of sessions attended for all individuals in the study). Applying this model to the worst-case scenario dataset should weaken the effect of treatment, because imputation of noncompleters missing CSBI–III scores were modeled so that only time could influence structured changes.

Thus the variability accounted for by the treatment effect among all noncompleters should be close to zero in this scenario, substantially dampening the overall treatment effect. In both runs of this multilevel growth model, there were insufficient degrees of freedom for estimating variability around all three Level 1 random effects (i.e., initial status intercepts and time and treatment regression coefficients). To explore the variability of each piecewise trend effect, separate models were run, fixing one of the trend-estimate variances to zero, to assess the variability in the other. Finally, all inferential results reported next utilized the robust standard errors and t tests (with approximate df) provided by HLM version 6.0 software (see Raudenbush & Bryk, 2002).

Model effects for each scenario are presented in Table 3. Estimated effects, under the best-case scenario, represent an unbiased look at time and treatment effects, provided treatment sessions linearly influenced CSBI–III measured SBP, and session impact was no less effective for noncompleters than completers. Under these assumptions, we find a significant impact of treatment, $t(10) = -3.86$, $p < .005$, for the average individual in our sample (note that age and gender are mean centered so that the Level 2 intercept term represents the average treatment effect). The variability around this effect was not significant ($p = .38$) when fixing the variability of the time effect to zero. A marginally significant effect for time was found, $t(10) = -1.89$, $p < .09$, with significant variability, $\chi^2(82) = 127.08$, $p < .01$. It is interesting to note that although the competing influences of both treatment and time appear to be significant reducers of SBP, treatment in this model has more than double the impact of time (a $-.96$ unit decrease per weekly session completed vs. a $-.42$ unit decrease per week elapsed). Also of interest, significant influences of age and gender on each piecewise effect were found, and these influences were in opposite directions for time and treatment. Although the time effect was weaker for older children and nearly nonexistent for female children, the opposite was true for the treatment effect. Moreover, male and younger children seemed to be affected slightly more from passage of time than treatment.

Table 3. *Parameter Estimates, Tests, and p Values From Final Intent-to-Treat Analyses*

Scenario	Parameter	Estimate	*SE*	H_0 Test[a]	p
Best Case:	Intercept	31.83	1.54	20.73[c]	< .01
Partial	Time	−0.42	0.22	−.89[c]	0.09
Treatment	Time × Age	0.36	0.19	1.89[c]	0.10
Impact	Time × Gender	0.95	0.30	3.12[c]	< 0.01
	Treatment	−0.96	0.25	−3.86[d]	< 0.01
	Treatment × Age[b]	−0.56	0.25	−2.27[d]	0.04
	Treatment × Gender[c]	−0.93	0.38	−2.42[d]	.02
	Intercept *SD*	11.04		305.61[e]	< .01
	Time *SD*	0.59		127.08[e]	< .01
Worst Case	Intercept	31.76	1.53	20.73[d]	< 0.01
Completers	Time	−0.41	0.16	−2.56[d]	0.02
Treatment	Time × Age	0.36	0.15	2.43[c]	0.04
Impact	Time × Gender	1.02	0.26	3.90[c]	< 0.01
	Treatment	−0.83	0.18	−4.66[d]	< 0.01
	Treatment × Age	−0.49	0.20	−2.42[d]	0.02
	Treatment × Gender	−0.77	0.34	−2.22[d]	0.03
	Intercept *SD*	11.01		303.72[e]	< .01
	Time *SD*	0.63		132.64	< .01

[a] Parameter = 0.
[b] Age was mean centered.
[c] Gender was coding as −.58 for male and .42 for female, resulting in a mean of 0.
[d] Robust *t* test statistic for H_0.
[e] Chi-square test statistic for H_0.

Under the restrictive assumption of the worst-case scenario, treatment effect is still significant, $t(36) = -4.66$, $p < .01$, for the average child in the study, which does not seem to vary substantially across individuals ($p = .29$). This analysis also finds a significant effect of time, $t(35) = -2.56$, $p < .02$, that varies significantly, $\chi^2(82) = 132.64$, $p < .01$, across the sample. These treatment and time effects were modified by gender and age, in a similar pattern to the best-case scenario previously described. In both scenarios, age- and gender-modified time and treatment effects indicate formidable influences on SBP reduction, accounting for 34% and 19% of the Level 1 variability, respectively, under best-case assumptions, and 36% and 35%, under the worst case.

In summary, even in the restrictive worst-case scenario, the SBP program reached significant treatment impact ranges. Under the assumption that treatment can have an incremental impact, more than one third of the variance was accounted for by treatment effects, with female and older children most favorably impacted.

Specificity of Treatment Effects

To examine specificity of effects, a repeated measures ANOVA was conducted to examine changes on the CBCL, DISC–IV, and PSI–SF (see Table 2). There were no significant differences on the DISC–IV Depression scale, CBCL, or the PSI–SF. For children who met Criteria A for PTSD, the number of PTSD symptoms changed from intake ($M = 4.1$, $SD = 2.7$) to posttreatment ($M = 2.4$, $SD = 2.6$), $t(8) = 2.5$, $p < .01$.

Caregiver Satisfaction and Social Validity Ratings

On the CSBP-PGSQ, caregivers recommended continuing all aspects of the group. Caregivers rated their knowledge after treatment as being significantly greater than their knowledge before treatment, $t(34) = -11.65$, $p < .05$. The median and mode rating of quality of the topics was "very useful," and the median and mode rating of impact of the group was "very much improved." The average satisfaction rating on a 7-point scale was 6.35.

Discussion

Preschool children have been found to demonstrate concerning interpersonal sexual behaviors. Prior to participating in treatment, the children in this study had high rates of interpersonal sexual behaviors, and more than one third of the children had aggressive or forceful sexual behaviors. In our intent-to-treat analysis, significant linear reduction in SBP were found due to number of treatment sessions attended, an effect that was independent of linear reductions affiliated with elapsed time.

Moreover, this treatment effect was modified by gender and age (but not sexual abuse history), suggesting stronger treatment impact for older and female children. Caregivers who completed the program reported increased knowledge, that the treatment was useful, and that they were satisfied with the services.

This study piloting a group treatment program was conducted concurrently with a broader evaluation of the clinical presentation of preschool children with SBP (Silovsky & Niec, 2002). The results of these studies reveal that preschoolers referred specifically because of SBP have a concerning high rate of SBP (average CSBI–III scores at the 99th percentile). Further, these children presented with noteworthy rates of stressful events and trauma histories (not just sexual abuse), traumatic stress reactions, high rates of other behavior problems (mean CBCL Total Score in the clinical range), disruptions in the child-caregiver relationship, and caregiver stress. In many ways, these youngsters present with more serious histories and presenting problems than school-age children with SBP (Silovsky & Niec, 2002).

Treatment effects were found with the youth studied; however, reductions in SBP were also found during the wait period. In both the best-case and worst-case intent-to-treat scenarios, nontreatment-related recovery accounted for approximately a .4 unit decrease in raw CSBI scores each week, roughly half the size of the treatment effect in these scenarios. This effect corresponds nicely with the significant observed differences found among study completers in the repeated measures analysis, where the wait period mean difference of 6.0 was roughly half the pre-post difference of 11.9. The only other known tracking of SBP during a wait or no treatment period is Stauffer and Deblinger (1996), who did not find reductions in SBP as measured by the CSBI–III over an average of a 12-week waiting period. However, their sample consisted of children referred due to a history of sexual abuse, not specifically due to sexual behavior problems, and the initial rate of SBP was less than half than the current sample ($M = 15.37$ vs. 32.4 on the CSBI–III).

To facilitate the separation of time and environmental effects, longitudinal research measuring changes in caregiver behavior (e.g., supervision), in opportunities to interact with other children, and in home environment as it relates to frequency of SBP is recommended. Further, gender and age differences in these effects will be important factors to examine in this research. After sexual behavior problems are recognized, environment and supervision changes may immediately occur (prior to the initiation of treatment) and may be required by CPS (e.g., close supervision and a stable environment without exposure to violence). Close parental supervision has been noted to be a necessary but not sufficient method for reduction in sexual behavior problems (Friedrich, 2004). Changes in supervision did occur before treatment with some children in this study who around intake had less exposure to other children (e.g., removed from day care), and their time with other children began to be closely supervised by adults.

The preschool children in this study were often multiply traumatized and had experienced a wide range of stressful events. It is interesting that sexual abuse history does not appear to be an influential factor in treatment response. Overall, PTSD symptoms were reduced, but because the DISC–IV was not administered at pretreatment, the effect of time is unknown. Our caregiver treatment addressed the impact of trauma on children and the children's group included considerable overlap with trauma-focused CBT (e.g., addressing emotion expression skills, coping strategies, and safety education; Deblinger & Heflin, 1996). However, our protocol does not include an exposure-based or trauma narrative component, which is often considered to be critical to the resolution of PTSD (Cohen, 1998) but inappropriate for group-based treatment. For children with comorbid SBP and PTSD, determining the best manner to integrate evidence-based treatments is an important next step in treatment development.

Other child behavior problems remained in the clinically significant range at posttreatment. Although behavior management strategies were included in our piloted treatment protocol, they were presented in an abbreviated format with less time for direct coaching of caregiver behavior than in treatments found efficacious for disruptive behavior disorders in preschool children (e.g., PCIT; Hembree-Kigin & McNeil, 1996). Given the severity of the behavior problems, these children may benefit from more intensive behavior management services as severe disruptive behaviors in preschool children are not likely to dissipate without effective intervention (Boggs et al., 2004; McNeil, Capage, Bahl, & Blanc, 1999).

Results of this study must be interpreted with caution, given the limitations. This relatively small pilot study did not include a control group or a comparison group. This pilot study did not have the resources to collect data on families who did not follow through with services, so complete data were collected on a little over one third of the sample. Further, the asymptotic calculations of the intent-to-treat multilevel standard error estimations (for this unbalanced design) are based on large-sample theory.

Posttreatment data including the CSBP–PGSQ were available only for families who completed the treatment. Of the families who began services, 19% failed to complete services, a dropout rate not unusual for treatment of child behavior problems and low when compared to services with families involved in child welfare (Chaffin et al., 2004; Luk et al., 2001; Nordstrom, 2005). Families who dropped out did not differ from those who completed on demographic or initial clinical severity, but we were unable to examine families' perceptions of the program, barriers, and other reasons for attrition. In addition to replication, future studies would benefit from examining barriers and factors that facilitate engagement in services.

Collection of follow-up data is needed for the evaluation of maintenance of the treatment effect and for new explorations of variables that might predict trajectory shape and age and gender effects. Given the multiple needs of preschool children with SBP, multifaceted treatment appears to be warranted. Determining and evaluating the best method to integrate evidence-based treatment for SBP, trauma, and disruptive behavior problems are important next steps.

References

Abidin, R. R. (1995). *Parenting stress index*. Odessa, FL: Psychological Assessment Resources.

Achenbach, T. M. (1991). *Manual for the child behavior checklist/4–18 and 1991 profile*. Burlington: University of Vermont Department of Psychiatry.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Araji, S. K. (1997). *Sexually aggressive children: Coming to understand them*. Thousand Oaks, CA: Sage.

Association for the Treatment of Sexual Abusers (ATSA) Public Policy Committee. (2004). *Children with sexual behavior problems: Information sheet*. Author.

Baker, A., Schneiderman, M., & Parker, R. (2001). A survey of problematic sexual behavior of children in the New York City child welfare system: Estimates of problems, impact on services and need for training. *Journal of Child Sexual Abuse*, *10*, 67–80.

Boggs, S. R., Eyberg, S. M., Edwards, D. L., Rayfield, A., Jacobs, J., Bagner, D., et al. (2004). Outcomes of parent-child interaction therapy: A comparison of treatment completers and study dropouts one to three years later. *Child & Family Behavior Therapy*, *26*, 1–22.

Bonner, B. L. & Fahey, W. E. (1998). Children with aggressive sexual behavior. In N. N. Singh & A. S. W. Winton (Eds.), *Comprehensive clinical psychology: Vol. 9 Special populations* (pp. 453–466). Oxford, England: Elsevier Science.

Bonner, B. L., Walker, C. E., & Berliner, L. (1999). *Children with sexual behavior problems: Assessment and treatment. Final Report, Grant No. 90-CA-1469*. Washington DC: Administration of Children, Youth, and Families, DHHS.

Brestan, E. & Eyberg, S. (1998). Effective psychosocial treatments of conduct-disordered children and adolescents: 29 years, 82 studies, and 5,272 kids. *Journal of Clinical Child Psychology*, *27*, 180–189.

Burke, J. D., Loeber, R., & Birmaher, B. (2002). Oppositional defiant disorder and conduct disorder: A review of the past 10 years: Part II. *American Academy of Child and Adolescent Psychiatry*, *41*, 1275–1293.

Carpentier, M., Silovsky, J. F., & Chaffin, M. (2006). Randomized trial of treatment for children with sexual behavior problems: Ten year follow-up. *Journal of Consulting and Clinical Psychology*, *74*, 482–488.

Chaffin, M., Berliner, L., Block, R., Friedrich, W., Johnson, T., Lewis, D., et al. (2006). *Report of the ATSA task force on children with sexual behavior problems*. Beaverton, OR: Association for the Treatment of Sexual Abusers.

Chaffin, M. C., Silovsky, J. F., Funderburk, B., Valle, L. A., Brestan, E. V., Balachova, T., et al. (2004). Parent-child interaction therapy with physically abusive parents: Efficacy for reducing future abuse reports. *Journal of Clinical and Consulting Psychology*, *72*, 491–499.

Chaffin, M., Wherry, J. N., Newlin, C., Crutchfield, A., & Dykman, R. (1997). The Abuse Dimensions Inventory: Initial data on a research measure of abuse severity. *Journal of Interpersonal Violence*, *12*, 569–589.

Cohen, J. A. (1998). Summary of the practice parameters for the assessment and treatment of children and adolescents with posttraumatic stress disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *37*, 997–1001.

Cohen, J. A. & Mannarino, A. P. (1996). A treatment outcome study for sexually abused preschool children: Initial findings. *Journal of the American Academy of Child and Adolescent Psychiatry*, *35*, 42–50.

Cohen, J. A. & Mannarino, A. P. (1997). A treatment study for sexually abused preschool children: Outcome during a one-year follow-up. *Journal of the American Academy of Child and Adolescent Psychiatry*, *36*, 1228–1235.

Collins, L. M., Schafer, J. L., & Kam, C. M. (2001). A comparison of inclusive and restrictive strategies in modern missing-data procedures. *Psychological Methods*, *6*, 330–351.

Deblinger, E. & Heflin, A. (1996). *Treating sexually abused children and their nonoffending parents: A cognitive behavioral approach*. Thousand Oaks, CA: Sage.

Dunn, L. M. & Dunn, L. M. (1981). *Peabody picture vocabulary test manual for forms L and M*. Circle Pines, MN: American Guidance Service.

Finkelhor, D. & Berliner, L. (1995). Research on the treatment of sexually abused children: A review and recommendations. *Journal of the American Academy Child & Adolescent Psychiatry*, *24*, 1408–1423.

Friedrich, W. N. (1997). *Child sexual behavior inventory: Professional manual*. Odessa, FL: Psychological Assessment Resources.

Friedrich, W. N. (2002). *Psychological assessment of sexually abused children and their families*. Thousand Oaks, CA: Sage.

Friedrich, W. N. (2004, March). *Treating sexual behavior problems in children*. Paper presented at the 20th National Symposium on Child Abuse, Huntsville, AL.

Freidrich, W. N., Davies, W. H., Feher, E., & Wright, J. (2004). Sexual behavior problems in preteen children: Developmental, ecological, and behavioral correlates. *Annuals of New York Academy of Science*, *989*, 95–104.

Friedrich, W. N., Fisher, J., Dittner, C., Acton, R., Berliner, L., Butler, J., et al. (2001). Child sexual behavior inventory: Normative, psychiatric, and sexual abuse comparisons. *Child Maltreatment*, *6*, 37–49.

Friedrich, W. N. & Luecke, W. J. (1988). Young school-age sexually aggressive children. *Professional Psychology*, *19*, 155–164.

Futa, K., Hecht, D., & Hansen, D. (1998). *Child Project SAFE Evaluation Form*. Unpublished measure.

Gil, E. & Johnson, T. C. (1993). *Sexualized children: Assessment and treatment of sexualized children and children who molest*. Rockville, MD: Launch Press.

Gray, A., Pithers, W.D., Busconi, A., & Houchens, P. (1999). Developmental and etiological characteristics of children with sexual behavior problems: Treatment implications. *Child Abuse and Neglect*, *23*, 601–621.

Hall-Marley, S. E. & Damon, L. (1993). Impact of structured group therapy on young victims of sexual abuse. *Journal of Child and Adolescent Group Therapy*, *3*, 41–48.

Haskett, M., Ahern, L., Ward, C., & Allaire, J. (2006). Factor structure and validity of the parenting stress index-short form. *Journal of Clinical Child and Adolescent Psychology*, *35*, 302–312.

Heiman, M. (2001). Helping parents address their child's sexual behavior problems. *Journal of Child Sexual Abuse*, *10*, 35–57.

Hembree-Kigin, T. L. & McNeil, C. (1996). *Parent-child interaction therapy*. New York: Plenum.

Horton, C. B. (1996). Children who molest other children: The school psychologist's response to the sexually aggressive child. *School Psychology Review*, *25*, 540–557.

Johnson, T. C. (1988). Child perpetrators—Children who molest other children: Preliminary findings. *Child Abuse & Neglect*, *12*, 219–229.

Kendall-Tackett, K., Williams, L., & Finkelhor, D. (1993). Impact of sexual abuse on children: A review and synthesis of recent empirical studies. *Psychological Bulletin*, *113*, 164–180.

Little, R. J. & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). Hoboken, NJ: Wiley.

Little, R. J. A. & Yau, L. (1996). Intent-to-treat analysis in longitudinal studies with drop-outs. *Biometrics*, *52*, 1324–1333.

Luk, E. S. L., Staiger, P. K., Mathai, J. Wong, L., Birleson, P., & Adler, R. (2001). Children with persistent conduct problems who dropout of treatment. *European Child and Adolescent Psychiatry*, *10*, 28–36.

McKenzie, W., English, D., & Henderson, J. (1987). *Family centered case management with sexually aggressive youth: Final report of the sexually aggressive youth project*. Washington State Department of Social and Health Services, Division of Children, Youth & Family Services.

McNeil, C. B., Capage, L. C., Bahl, A., & Blanc, H. (1999). Importance of early intervention for disruptive behavior problems: Comparison of treatment and waitlist-control groups. *Early Education & Development*, *10*, 445–454.

National Institute of Mental Health. (1998). *Diagnostic interviewer's manual: Draft*. New York: Columbia University.

Nixon, R. (2002). Treatment of behavior problems in preschoolers: A review of parent training programs. *Clinical Psychology Review*, *22*, 525–546.

Nordstrom, A. (2005). Engagement and retention in family-based preventive interventions: The role of parental attributions and perceived barriers to participation. *Dissertation Abstracts International: Sciences and Engineering*, *65*, 4842.

Pearce, J. (2003, September). *Working with maltreated children: Guidance from attachment theory and research*. Paper presented at the Oklahoma Conference on Child Abuse and Neglect, Oklahoma City, OK.

Pithers, W. & Gray, A. (1993). *Pre-adolescent sexual abuse research project* (National Center on Child Abuse and Neglect Grantees Status Report 90-C0-1470). Washington DC: DHHS, NCCAN.

Pithers, W., Gray, A., Busconi, A., & Houchens, P. (1998). Children with sexual behavior problems: Identification of five distinct child types and related treatment considerations. *Child Maltreatment*, *3*, 384–406.

Rasbash, J., Browne, W., Goldstein, H., Yang, M., Plewis, I., Healy, M., et al. (2002). *A user's guide to MLwiN, 2nd edition*. London: Institute of Education.

Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods (2nd ed)*. Thousand Oaks, CA: Sage.

Schafer, J. L. & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*, 147–177.

Shaffer, D., Fisher, P., Lucas, C., Dulcan, M., & Schwab-Stone, M. E. (2000). NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC–IV): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry*, *39*, 28–38.

Silovsky, J. F. & Bonner, B. L. (2003). Children with sexual behavior problems. In T.H. Ollendick & C. S. Schroeder (Eds.), *Encyclopedia of clinical child and pediatric psychology* (pp. 589–591). New York: Kluwer.

Silovsky, J. F. & Niec, L. (1998). *Group treatment for preschool children with problematic sexual behavior: Program manual*. Unpublished treatment manual.

Silovsky, J. F. & Niec, L. (2002). Characteristics of young children with sexual behavior problems: A pilot study. *Child Maltreatment*, *7*, 187–197.

Stauffer, L. B. & Deblinger, E. (1996). Cognitive behavioral groups for nonoffending mothers and their young sexually abused children: a preliminary treatment outcome study. *Child Maltreatment*, *1*, 65–76.

Williams, K. T. & Wang, J. J. (1997). *Technical references to the Peabody Picture Vocabulary Test-Third Edition (PPVT–III)*. Circle Pines, MN: American Guidance Service.

Received June 10, 2005
Accepted March 5, 2007

Appendix

Before imputation, we estimated the treatment and time trends and explored covariate prediction of each using HLM's (version 6.0, 2004) Restricted Maximum Likelihood procedure for a piecewise regression model assuming normally distributed residuals and Missing at Random (MAR) distributed missing data (see Little & Rubin, 2002). Results from this analysis appear in the main text. The assumption of normally distributed residuals at Level 1 appeared justified based on inspection of Q-Q plots of the final model residual estimates. The viability of normally distributed Level 2 residuals and MAR-distributed missingness, on the other hand, was much more difficult to assess empirically because of the small number of observations per subject (see Raudenbush & Bryk, 2002, p. 274) and the impossible task of verifying violations of MAR (i.e., that missingness depends on covariates not included in the model; see

plausibility of MAR in Schafer & Graham, 2002, p. 152) without subsequent retrieval of data or reasons for missingness. In support of these assumptions, however, the Level 2 normality assumption has no influence on the bias of fixed effect estimates, and test statistics chosen for assessing model inclusion were robust to violations of the random effect residual assumptions. Likewise, MAR-based control for missingness is considered the "practical state of the art" (Schafer & Graham, 2002, p. 173), tends to perform well even when a few covariate predictors are missing from the model (Collins et al., 2001), and carries intuitive appeal in longitudinal contexts when previous outcome responses often predict future outcome missingness.

Retaining all significant covariate effects, modeling shifted focus toward imputation of missing data. First, when pretest and/or posttest assessments were missing, the corresponding time points were imputed using a closest neighbor approach. The closest neighbor was defined as another individual completing their baseline assessment within shortest time proximity, provided this neighbor did not start treatment until after the participant with missing assessments was recruited. Using MLwiN 2.02 (Rasbash et al., 2002), multivariate imputation models were constructed, as described in the text, again under a MAR missing data assumption. In the best-case scenario, CSBI–III missingness was imputed based on the piecewise age and gender-adjusted linear effects of time and number of treatment sessions attended. In the worst-case scenario, only the scores from individuals who met requirements for treatment completion were imputed using both piecewise effects. All noncompleters' missing outcomes were imputed based solely on the time trend effect. Six imputations per scenario were performed, and then data were reanalyzed using the best-case imputation model except this time incorporating HLM's (2004) option for combining and testing effects from multiple imputations.