Treatment Outcome and Criminal Offending by Youth With Sexual Behavior Problems

Elizabeth J. Letourneau
Jason E. Chapman
Sonja K. Schoenwald
Medical University of South Carolina

Children and adolescents treated for general delinquency problems and rated by caregivers as having sexual behavior problems (SBP; N = 696) were compared with youth from the same sample with no sexual behavior problems (NSBP; N = 1,185). Treatment outcome through 12-months post-treatment and criminal offending through an average 48-month posttreatment were compared for both groups. It was hypothesized that both groups would improve over time; however, the SBP group would evidence greater psychopathology at follow-up, and these hypotheses were supported. It was further hypothesized that youth with SBP would not differ from youth with NSBP in rates of future sexual or nonsexual offenses. These hypotheses were also supported. SBP group membership was not a significant predictive factor in analyses modeling future offending (any) or future person offenses. Few youth in either group had sexual offenses. The importance of these findings for clinical and policy decision making is discussed.

Keywords: adolescent; child; recidivism; sexual behavior problems; treatment outcome

The scientific literature on children with sexual behavior problems is limited, but of exceptionally high quality. There have, for example, been multiple randomized clinical treatment trials (Bonnier, Walker, & Berliner, 1993, 1999; Cohen & Mannarino, 1996, 1997; Pithers & Gray, 1993; Pithers, Gray, Busconi, & Houchens, 1998), and recently, a 10-year follow-up from one of those trials (Carpentier, Silovsky, & Chaffin, 2006). The results of these and other treatment outcome studies collectively support short-term, sexual behavior–focused cognitive-behavioral therapy (CBT) interventions that substantially include children’s caregivers (Carpentier et al., 2006; Cohen & Mannarino, 1996, 1997; Deblinger & Hefflin, 1996; Silovsky, Niec, Bard, & Hecht, 2007). In a 10-year follow-up, Carpentier and colleagues (2006) reported that youth treated with CBT committed significantly fewer sexual offenses than youth treated with play therapy. It is important to note that the rate of future sexual offending by the CBT condition was indistinguishable from a sample of children originally presenting with nonsexual psychiatric problems (Carpentier et al.,...
2006). Taken together, these results do much to dispel concerns that children with sexual behavior problems follow a developmental path that leads them through adolescent sexual offending and on to adult sexual predatory behaviors—a pathway suggested by retrospective studies linking adult (or adolescent) sexual offending to adolescent (or childhood) sexual misbehavior (Abel et al., 1987; Burton, 2000; Zolondek, Abel, Northey, & Jordan, 2001).

Sexual behavior problems (SBP) have been defined as persistent and developmentally atypical sexual behaviors (Carpentier et al., 2006), including sexual behaviors that (a) occur with unexpected frequency, (b) occur in coercive contexts or between older and younger children, (c) have been resistant to caregiver intervention, (d) interfered with the child’s development, and/or (e) were associated with emotional distress by the child (see Chaffin, Letourneau, & Silovsky, 2002, p. 208). To date, most research on SBP, including the randomized treatment trials cited above, has focused on children age 6 to 12 years. Some investigators have examined SBP in younger children (e.g., Silovsky & Niec, 2002) and across broader age ranges (e.g., Letourneau, Schoenwald, & Sheidow, 2004). The current study seeks to build on this earlier research by examining treatment outcome across a long time period and examining future criminal offending by youths designated as having SBP.

In our earlier study (Letourneau et al., 2004), youth referred for treatment primarily because of delinquent behavior and subsequently identified by their caregivers as having any SBP were compared with similar youth identified by their caregivers as having no SBP (NSBP) on a variety of personal characteristics and with regards to immediate posttreatment outcome. All youth received multisystemic therapy (MST; Henggeler, Schoenwald, Borduin, Rowland, & Cunningham, 1998) after having been referred to community-based clinics for serious antisocial and other externalizing behavior problems. The behavior problems that resulted in treatment referrals were not necessarily due to involvement with the legal system, and youth were not accepted for MST treatment if the primary referral reason was for criminal sexual offending (this caveat applies to all MST programs until ongoing research indicates the effectiveness of MST as adapted for use with juveniles who sexually offend; see Letourneau, Borduin, & Schaeffer, in press).

In the original study, SBP were measured using the Sex Problems scale of the parent-reported Child Behavior Checklist (CBCL; Achenbach, 1991). Scale scores were trichotomized to form groups of youth with NSBP (i.e., scale score = 0), youth with limited SBP (i.e., scale score = 1 to 2), and youth with SBP (i.e., scale score = 3 to 12). These groups differed significantly with regards to several personal characteristics (e.g., youth with SBP were younger and more likely to be female than youth with no or limited SBP). There were, however, no significant between-group differences with regards to treatment outcome for the SBP and limited SBP groups. Both groups improved significantly from baseline but, at immediate posttreatment follow-up, were still characterized by more behavior problems than youth with NSBP.

The current study extends the examination of treatment outcome through 12-months posttreatment follow-up and includes examination of criminal charges through an average 48-month posttreatment follow-up. Given the lack of group differences in treatment outcome reported in the initial study for the two groups with any SBP, youth were dichotomized into NSBP and any SBP for the current study. Specific hypotheses regarding patterns of treatment outcome and criminal charges include the following:

**Treatment Outcome**

Based on previous findings that youth with SBP responded well to treatment but continued to exhibit significantly more behavior problems than youth without SBP, it was hypothesized that

1. Parent-reported behavior problems would continue to decline for youth with and without SBP through 12-month posttreatment follow-up.
2. Parent-reported behavior problems would remain higher for youth with SBP than youth without SBP through the 12-month posttreatment follow-up.

**Criminal Charges (Nonsexual)**

Nonsexual offenses typically occur with much greater frequency than sexual offenses, even in samples characterized by previous sexual offending (e.g., Caldwell, 2002; Carpentier et al., 2006). Carpentier and colleagues (2006) reported no between-group difference for nonsexual offenses but did find that higher pretreatment CBCL Externalizing T scores, male gender, and older baseline age were associated with greater rates of general offending. In our previous study (Letourneau et al., 2004), youth with SBP were found to have higher externalizing and internalizing scale scores than youth with NSBP but were younger and more likely to be girls. Based partly on these findings, it was hypothesized that...
3. Youth with SBP would have general offense rates comparable to youth with NSBP through approximately 4-year (48-month) follow-up.

**Criminal Charges (Sexual)**

Based on findings of comparable sexual offense rates at 10-year follow-up among effectively treated youth with and without SBP (Carpentier et al., 2006) it was hypothesized that

4. Youth with SBP would have sexual offense rates comparable to youth without SBP through approximately 4-year (48-month) follow-up.

**METHOD**

**Participants**

Participants were originally recruited for a study examining the transportability of MST to community service settings (for description of the original study, see Schoenwald, Sheidow, Letourneau, & Liao, 2003). A total of 1,979 (98%) caregivers provided usable treatment outcome data on their children. The current study includes only those participants for whom baseline Sex Problems scale scores could be computed. At baseline, caregiver CBCL measures were missing one or more sex problems items for 98 youths. Thus, the sample for the current study is 1,881 or 95% of the sample for whom usable treatment outcome data existed.

**Youth characteristics.** Referral sources included juvenile justice agencies (43%), social services agencies (23%), mental health agencies (18%), educational facilities (3%), and other agencies (13%). Therapists indicated up to 10 reasons for referral and these included criminal (46%) or status (48%) offenses, substance use problems (32%), school problems (30%), threat of harm to self or others (29%), aftercare following out-of-home placement (18%), domestic violence (17%), abuse or neglect by caregivers (10%), serious emotional disturbance (10%), or an unspecified reason (10%). (The primary transportability study did not focus on SBP; and, consequently, this was not a referral reason captured in the set of questionnaires.) Therapists indicated multiple referral reasons for most youth, $M = 2.4$ ($SD = 1.4$, range $= 0$ to 9), and most youth ($81\%$) were referred for behavior problems including delinquent or status offenses, substance use problems, and school suspensions or expulsion. Underscoring the fact that this was a sample of youths with serious behavior problems, most ($69.7\%$) had experienced one or more outof-home placements (e.g., incarceration, treatment at a residential facility, psychiatric hospitalization). Most youths were male ($65\%$) and White ($59\%$), African American ($19\%$), Asian or Pacific Islander ($7\%$), or Other ($15\%$; includes biracial and youth who indicated only Hispanic as race). Slightly more than $7\%$ of youth indicated Hispanic ethnicity. The majority of youth resided with one ($56\%$) or both ($15\%$) biological or adoptive parents. At intake, participants had a mean age of 14 years ($SD = 2.4$). At the 12-month posttreatment follow-up, participants had a mean age of 16.02 years ($SD = 2.34$) and at the criminal records follow-up (which varied from 29 months to 68 months posttreatment), participants had a mean age of 19.60 years ($SD = 1.87$).

**Caregiver characteristics.** Primary caregivers of youths completed the baseline and treatment outcome measures and were likely to be female ($88\%$) and White ($65\%$), African American ($19\%$), Asian or Pacific Islander ($7\%$), or Other ($9\%$). Seven percent of caregivers indicated Hispanic ethnicity. At intake, caregivers had a mean age of 41 years ($SD = 8.6$). One fourth of the caregivers never completed high school, $39\%$ completed high school or obtained a General Equivalency Diploma (GED), and the remainder had vocational or scholastic education beyond high school ($36\%$). Family incomes ranged from less than US$10,000 ($26\%$) to more than US$40,000 ($21\%$), with $36\%$ of families reporting some sort of financial assistance.

**Therapist characteristics.** Of the 1,881 families included in the current SBP study, primary therapists could be identified for 1,805. Primary therapist signifies the therapist who treated the family for the entire treatment episode or, for families treated by more than one therapist, the therapist who provided treatment for the majority of the family’s treatment episode. A primary therapist could not be identified for 76 families, each of which was treated by more than one therapist for approximately equal lengths of time. Differences between families with and without a primary therapist have not been found (Schoenwald & Chapman, 2006). The 1,805 families for whom a primary therapist was identified were treated by 419 therapists at 45 MST treatment sites in 12 states and Canada. The majority of therapists was female ($64\%$), and most held master’s degrees ($69\%$) or bachelor degrees ($23\%$). The most common degree fields included social work and counseling. At the start of treatment, nearly one half of the MST therapists were new to this treatment model (i.e., $46\%$ had been employed by the MST program for 3 or fewer months), although more than one fourth ($27\%$) had one or more years experience with the MST model.
**Measures**

According to family needs and strengths, although duration and frequency of sessions varies, MST is typically delivered for 4 to 6 months, (e.g., schools) to improve generalization, engage settings in the social ecology of the youth and family. Treatment is based in the home and other relevant caregivers to effectively parent their children. The goal of MST is to empower chronic juvenile offenders and juveniles with sub-

**Procedure**

Study procedures have been thoroughly described in previous publications (please see Schoenwald, Letourneau, & Halliday-Boykin, 2005; Schoenwald, Sheidow, & Letourneau, 2004; Schoenwald et al., 2003) and are briefly described here.

**Youth and families.** All youth referred for MST treatment at the selected study sites were eligible for the transportability study except youth who were autistic or suffered severe mental retardation. Families were recruited by clinical supervisors or therapists at the provider organizations, and the consent rate was 82%. Research assistants administered pre- and posttreatment assessment measures to youth, caregivers, and therapists by telephone, and caregivers were reimbursed for completed assessments. Participation in the original study was voluntary, although treatment participation in an MST program could have been court ordered (based on review of archival juvenile justice records, it is known that court-ordered treatment occurred for some youths; however, this variable was not systematically tracked).

**Clinical intervention.** As with study procedures, the clinical intervention has been described previously (for review, see Sheidow & Henggeler, in press). Briefly, MST is a well-validated intervention for chronic juvenile offenders and juveniles with substance use disorders. The goal of MST is to empower caregivers to effectively parent their children. Treatment is based in the home and other relevant settings in the social ecology of the youth and family (e.g., schools) to improve generalization, engage caregivers, and overcome barriers to treatment access. MST is typically delivered for 4 to 6 months, although duration and frequency of sessions varies according to family needs and strengths.

**Measures**

**Descriptive information.** A comprehensive measure of youth and caregiver demographic information was created specifically for the transportability study and administered by researchers to caregivers at baseline. Items queried about youth and caregiver demographics (e.g., age, race, ethnicity, gender), caregiver marital status, income, work status, and information about additional people residing in the home. A separate measure collected a truncated set of demographic variables from therapists, including gender, age, race, ethnicity, salary, and professional experience (e.g., months employed by the MST provider site). Therapists also indicated primary referral source (limited to a single primary source) and primary referral reasons (up to 10 possible referral reasons) for each youth at baseline.

**SBP.** SBP were identified by the caregiver-reported Child Behavior Checklist (CBCL) Sex Problems scale (Achenbach, 1991). As with all CBCL problem-behavior items, items on the Sex Problems scale are rated 0 (not true), 1 (sometimes or somewhat true), or 2 (often true). The Sex Problems scale has six items, with possible scale scores ranging from 0 to 12. Table 1 presents the rates at which caregivers endorsed the individual Sex Problems scale items. The two most frequently endorsed items were “thinks about sex too much” and “sexual problems.” CBCL instructions require elaboration when caregivers endorse the rather vague sexual problems item. However, because of Institutional Review Board concerns about confidentiality (all caregiver reports were collected via telephone where responses might be over-heard), study personnel were prohibited from requesting elaboration when the sexual problems item was endorsed.

Although limited in scope, the CBCL scale and derivatives of that scale have frequently been used to assess SBP in children (e.g., Cohen & Mannarino, 1988; Einbender & Friedrich, 1989; Friedrich, 1993; Längström, Grann, & Lichtenstein, 2002; Silovsky & Niec, 2002). More comprehensive measures of SBP in children and adolescents were developed by Friedrich and his colleagues, including the Child Sexual Behavior Inventory (CSBI; Friedrich, 1993; Friedrich, Grambsch, Broughton, Kuiper, & Beilke, 1991) and the Adolescent Clinical Sexual Behaviors Inventory (ASBI; Friedrich, Lysne, Sim, & Shamos, 2004). More comprehensive measures of SBP were not included in the original transportability study, however, because the study was not originally designed to address problem sexual behaviors and outcome measures were

| Table 1: Caregiver Reported Sexual Behavior Problems at Baseline |
|--------------|-----------------|-----------------|-----------------|
| CBCL Sex Problems Items | Score of 0 (not true) | Score of 1 (somewhat true) | Score of 2 (very true) |
| Behaves like the opposite sex | 1708 (91%) | 119 (6%) | 54 (3%) |
| Plays with own sex parts in public | 1808 (96%) | 49 (3%) | 24 (1%) |
| Plays with own sex parts too much | 1769 (94%) | 66 (4%) | 46 (2%) |
| Sexual problems | 1652 (88%) | 117 (6%) | 112 (6%) |
| Thinks about sex too much | 1392 (74%) | 269 (14%) | 220 (12%) |
| Wishes to be of opposite sex | 1834 (98%) | 32 (2%) | 15 (< 1%) |
necessarily limited to those that would best captured the outcomes of primary interest. It is important to note that the content, concurrent, and discriminant validity of the CBCL Sex Problems scale is supported by several studies (Bonner et al., 1999; Friedrich et al., 1992; Pithers et al., 1998).

**Treatment outcomes.** Two measures of caregiver-rated child behavior problems were collected at four time points: pretreatment (baseline), immediately posttreatment, and at 6- and 12-months posttreatment. These two measures were the CBCL and the Vanderbilt Functioning Inventory (VFI; Bickman, Lambert, Karver, & Andrade, 1998). The CBCL is one of the best validated measures of child behavioral functioning (Achenbach, 1991). The 113 behavior problem items applicable to children aged 2 to 18 years were administered, and two broadband behavior problem scales (Internalizing and Externalizing) were derived from these items.

The VFI was developed to evaluate youth behavioral functioning (Bickman et al., 1998), with items assessing school problems, aggression, self-protection, theft, and substance use. VFI proportion scores are computed by summing the responses to all completed items and dividing by the number of completed items, resulting in one proportion score per youth. Adequate internal consistency and support for concurrent, discriminant, and predictive validity have been reported (Bickman et al., 1998).

**Criminal charges.** Of the 1,881 participants included in the current study, criminal charge data were obtained for 1,716 (91.2%), across a mean posttreatment follow-up period of 48.79 months ($SD = 8.7$), with a range of 29- to 68-month follow-up. Youth charge data were obtained from county and state juvenile justice agencies and courts. For participants who had reached adulthood at the time of the follow-up request, charge data were obtained via public record searches available through the Internet, or from agencies housing adult criminal records. Raw data were obtained on the dates, types, and severity of lifetime pretreatment charges and charges accrued throughout the follow-up period. These data were coded by research staff to reflect five charge types (i.e., person, property, drug, public order, status or other offense) and charge severity (e.g., person offenses were rated the most severe, and within person offenses, murder was the rated as the most severe; status offenses were rated the least severe and within status offenses “incorrigible/ungovernable behavior” was rated the least severe). The coding scheme was based on coding systems used in previous randomized trials of MST (C. L. Hanson, Henggeler, Haefele, & Rodick, 1984); these systems, in turn, were based on the Uniform Crime Reports standards used by the FBI. In an ongoing study involving this same coding scheme, data on 433 charges were entered separately by two “blind” raters. There was 98.6% agreement regarding individual charges (i.e., both coders entered 427 of the 433 records but differed on whether they considered the remaining entries to be independent charges vs. modifiers of previously entered charges). These raters obtained similarly high agreement regarding dates of charges (97.9% agreement) and literal charge descriptions (97.7% agreement). Interrater agreement regarding coding of the 427 charges that both raters entered was 96.4%, indicating excellent specification for coding rules (Letourneau, 2006b).

Of the 1,716 youth for whom juvenile or adult justice records were obtained, 1,254 had at least one known charge (which could have occurred pre-, during, or posttreatment), and 462 had no known charges at final follow-up. Information on charges could not be obtained for 165 (8.8%) of the 1,881 participants included in the current sample. Most of these participants (159) were treated in jurisdictions that ultimately were unable to provide any juvenile justice data, despite initial agreements to do so.

**Statistical Methods**

An essential feature of the current data is the nesting of repeated measurements of youth behavior problems (Level-1), within youths (Level-2), within therapists (Level-3), and the nesting of youth criminal charge or discharge data (Level-1) within therapists (Level-2). Random-effects regression models (RRMs) model the dependency in outcome variance attributable to the nested data structure and accommodate continuous, dichotomous, and count-distributed outcomes (Raudenbush & Bryk, 2002). Hierarchical linear modeling (HLM) software (Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2004) was used to perform all RRMs, with full maximum likelihood estimation for continuous outcomes, Laplace approximation of maximum likelihood for dichotomous outcomes (i.e., Bernoulli models), and penalized quasi-likelihood for count-distributed outcomes (i.e., Poisson models). Specification of random effects was based on the likelihood ratio test when possible and otherwise was based on the Wald test for variance components (Singer & Willett, 2003). Test statistics for behavior problem outcomes were computed using robust standard errors (Maas & Hox, 2004), and population average results were interpreted for count-distributed outcomes, as recommended (Raudenbush & Bryk, 2002; Raudenbush et al., 2004).
### RESULTS

#### Group Characteristics

In the original report (Letourneau et al., 2004), it was noted that SBP groups differed significantly with regards to youth age and gender, referral source, and reason for referral. In the current study (which includes a larger sample than did our original study), youth with any SBP (*n* = 696) did not significantly differ from youth with NSBP (*n* = 1,185) with regard to age, race, or Hispanic ethnicity. The SBP group did have significantly more girls (46%) than the NSBP group (28%), *χ²(1) = 9.1, p < .01, φ = .18*. Consequently, gender was included as a covariate in subsequent analyses.

Neither type of referral agency nor mean number of referral reasons differed between groups. Separate chi-square analyses compared groups on each of the 10 referral reasons. To account for the increased risk of Type I error due to multiple analyses, alpha was adjusted to .005. Based on this adjusted significance level, groups differed on two referral reasons. Planned follow-up analyses on the two referral reasons were conducted to control for the effect of gender. Results for the victimization referral reason indicated gender, *b* (1) = .77 (*SE = .10), *p < .001*. Wald = 56.9 and victimization, *b* (1) = 9.1, *p < .005, φ = .07*. Specifically, there were significantly more SBP than NSBP youth with victimization and harm-related referral reasons. Planned follow-up analyses on the two referral reasons were conducted to control for the effect of gender. Results for the victimization referral reason indicated gender, *b* (1) = .77 (*SE = .10), *p < .001*. Wald = 56.9 and victimization, *b* (1) = 9.1, *SE = .45*, *p = .01*. Wald = 6.5 terms, but not a gender by victimization interaction term contributed significantly to the final equation. The same pattern of results occurred for the harm referral reason, where both gender, *b* (1) = .87 (*SE = .11), *p < .001*, Wald = 60.0 and harm, *b* (1) = .51 (*SE = .13), *p = .001*. Wald = 12.0 terms, but not a gender by harm interaction term contributed significantly to the final equation. Thus, between-group differences in referral reasons are not accounted for solely by the higher percentage of girls in the SBP group.

#### Parent-Reported Treatment Outcome Measures (Hypotheses 1 and 2)

Repeated reports by caregivers of youth behavior problems were modeled according to linear and quadratic polynomial terms (i.e., the number of months since treatment start) and SBP group membership in a three-level RRM as described above. In these models, the linear effect represents the instantaneous rate of change, and the quadratic effect represents the acceleration of change over time. The results are presented in Table 2 and depicted in Figures 1, 2, and 3. For CBCL Externalizing and Internalizing scales, and for the VFI psychosocial functioning problems, youths in the SBP group and females had significantly higher scores at the start of treatment. Holding constant the effect of gender, the significant negative linear effect indicates that the scores on each of these three outcome measures for youths in the NSBP group decreased from the start of treatment. The linear effect was negative and significantly stronger for youths in the SBP group. This indicates that SBP youths experienced even greater reductions in behavior problems immediately following the start of treatment, relative to their NSBP counterparts. The quadratic effect was positive and significant, indicating that for youths in the NSBP group, change occurred more rapidly initially and then gradually slowed over time. The quadratic effect was positive and significantly stronger for youths in the SBP group. The combination of the negative linear and positive quadratic effects indicates that, for

### Table 2: Random Regression Models for Parent-Reported Treatment Outcome Measures

<table>
<thead>
<tr>
<th>Parameter</th>
<th>CBCL Externalizing</th>
<th></th>
<th></th>
<th>CBCL Internalizing</th>
<th></th>
<th></th>
<th>VFI (Functioning)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>γ</td>
<td>SE</td>
<td>df</td>
<td>γ</td>
<td>SE</td>
<td>df</td>
<td>γ</td>
<td>SE</td>
<td>df</td>
</tr>
<tr>
<td>Intercept (γ_{00})</td>
<td>64.86***</td>
<td>.448</td>
<td>417</td>
<td>58.87***</td>
<td>.434</td>
<td>417</td>
<td>.3625***</td>
<td>.0073</td>
<td>417</td>
</tr>
<tr>
<td>SBP (γ_{10}; 0 = No)</td>
<td>9.22***</td>
<td>.515</td>
<td>1,798</td>
<td>8.54***</td>
<td>.519</td>
<td>1,798</td>
<td>.1022***</td>
<td>.0090</td>
<td>1,798</td>
</tr>
<tr>
<td>Gender (γ_{20}; 0 = Male)</td>
<td>1.19*</td>
<td>.509</td>
<td>1,798</td>
<td>-.87</td>
<td>.518</td>
<td>1,798</td>
<td>-.0386***</td>
<td>.0069</td>
<td>1,798</td>
</tr>
<tr>
<td>Linear (γ_{30})</td>
<td>-.64***</td>
<td>.066</td>
<td>417</td>
<td>-.57***</td>
<td>.072</td>
<td>1,799</td>
<td>-.0180***</td>
<td>.0013</td>
<td>417</td>
</tr>
<tr>
<td>SBP x Linear (γ_{40})</td>
<td>-.63***</td>
<td>.105</td>
<td>417</td>
<td>-.62***</td>
<td>.106</td>
<td>1,799</td>
<td>-.0102***</td>
<td>.0020</td>
<td>1,799</td>
</tr>
<tr>
<td>Quadratic (γ_{50})</td>
<td>.02***</td>
<td>.003</td>
<td>1,799</td>
<td>.01**</td>
<td>.004</td>
<td>1,799</td>
<td>.0005***</td>
<td>.0001</td>
<td>417</td>
</tr>
<tr>
<td>SBP x Quadratic (γ_{60})</td>
<td>.02***</td>
<td>.006</td>
<td>417</td>
<td>.02***</td>
<td>.005</td>
<td>1,799</td>
<td>.0004***</td>
<td>.0001</td>
<td>1,799</td>
</tr>
</tbody>
</table>

**NOTE.** CBCL = Child Behavior Checklist; VFI = Vanderbilt Functioning Inventory; SBP = sexual behavior problems. The *T* ratio test statistic for each parameter (omitted) was computed as β/SE. *p < .05, **p < .01, ***p < .001.*
the SBP group, early change was even more rapid and that change slowed more over time for these youths relative to their NSBP counterparts. Of note, given the between-group difference at baseline, the level of each outcome remained higher across time for the SBP group as compared to the NSBP group despite the stronger linear and quadratic effects for the SBP group.

Groups were not compared with respect to post-treatment changes on the CBCL Sex Problems scale, given the floor effect for the NSBP group. Descriptively, it is interesting to note that mean Sex Problems scale scores for the SBP group declined from 2.27 ($SD = 1.62$) at baseline to 1.35 ($SD = 1.77$) at immediate posttreatment, then to 1.07 ($SD = 1.60$) at 6-months posttreatment, then remained virtually unchanged at 1.09 ($SD = 1.77$) for the final 12-month posttreatment follow-up. In this same time period, scores for the NSBP group increased only marginally, from a mean of 0 at baseline to a mean of .24 ($SD = .79$) at final follow-up.

**General Offense Charge Rates (Hypothesis 3)**

As noted previously, charge data were available for 1,716 youths. Identification of a primary therapist was missing for 71 of these youths, precluding their inclusion in the subsequent RRM s that utilized therapist as one of the nested levels. A total of 1,645 youths, therefore, had available charge data, an identifiable primary therapist, and valid SBP data. Of these, 437 (27%; $n = 275$ and 162 in the NSBP and SBP groups, respectively) had zero lifetime pre-, during-, and posttreatment charges. Because these cases provide no information directly pertaining to the charge hypotheses, they were removed from subsequent analyses. Of the remaining 1,208 youths, 274 (23%) had zero posttreatment charges, resulting in a significantly nonnormal outcome distribution. As a result, these data were modeled according to a two-step approach where the outcome was first modeled according to a dichotomous distribution (i.e., presence vs. absence of at least one posttreatment charge) and then, among cases with nonzero posttreatment charges, modeled according to the number of charges (Min & Agresti, 2005). The results of each model are presented in Table 3.

For the first model (including all youths with at least one pre-, during-, or posttreatment charge), the results of the two-level Bernoulli RRM with Laplace approximation (holding constant the effects of age, gender, race, and number of lifetime pretreatment charges—all of which significantly influenced the likelihood of criminal charges) indicated that the average probability of at least one posttreatment
charge was .80 for youths in the NSBP group and .81 for youths in the SBP, a nonsignificant difference. Similarly, for the second model (including only those youths with at least one posttreatment charge), results of a two-level Poisson RRM indicated that the rate of posttreatment charges did not differ significantly between those in the SBP and NSBP groups. Holding constant the effects of the model covariates, the rate of posttreatment general offenses for youths in the NSBP and SBP groups was approximately .13 charges per month for each group.

Sexual Offense Charge Rates (Hypothesis 4)

There were too few posttreatment sexual offenses (n = 29 youth with posttreatment sexual offenses) to support RRM analysis. Descriptively, of the 1,032 NSBP and 612 SBP youths with valid charge data, an identifiable primary therapist, and valid SBP data, 20 (1.9%) NSBP and 9 (1.5%) SBP group members had at least one posttreatment sexual offense charge. When examining only those youths with at least one posttreatment charge for any type of offense, the percentage of NSBP and SBP youths with at least one posttreatment sexual offense charge was 2.6% and 2.0%, respectively. These differences were not statistically significant. The mean SBP score for the 29 youths with posttreatment sexual offenses was .52 (SD = .95, range = 0 to 4), somewhat lower than the entire sample, which had a mean SBP score of .85 (SD = 1.47, range = 0 to 12). Neither SBP group membership nor SBP scores was, therefore, associated with higher risk for future sexual offending.

Person offense charge probability. Unplanned additional analyses were conducted on “person offenses,” given that there were too few sexual offenses with which to conduct RRM. Person offenses, such as assault and battery, most closely approximate sexual offenses, and sexual offenses are often pled down to other person offenses (Letourneau, 2006a). Results of the two-level Bernoulli RRM with Laplace approximation, holding constant the effects of age, gender, race, and number of lifetime pretreatment criminal charges, indicated that the average log odds of at least one posttreatment person offense charge did not differ significantly between youths in the NSBP and SBP groups, \( \gamma_{SBP} = .01, SE = .135, T(1,193) = .07, p = .95 \), odds ratio (OR) = 1.01, 95% Confidence Interval (CI) = .78 to 1.32. Specifically, the average probability of at least one posttreatment person offense charge was identical (.34) for youths in the NSBP and SBP groups.

DISCUSSION

Youth in the current study were referred for treatment because of serious behavior problems, most often involving delinquency. In addition to their primary referral problems, a substantial portion of youths (37% of the sample) was identified by their caregivers as having one or more SBP. Results of the current study suggest that SBPs as indexed by the CBCL (e.g., sexual behaviors directed primarily at oneself vs. other-directed behaviors) are effectively addressed by MST, a relatively short-term, empirically supported treatment that focuses considerable attention on improving caregiver skills and resources. Findings also indicate that youths identified as having SBP and treated by an effective intervention are not at increased risk for future sexual offending. As such,

<table>
<thead>
<tr>
<th>Parameter</th>
<th>( \gamma )</th>
<th>SE</th>
<th>df</th>
<th>OR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dichotomous model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (( \gamma_{00} ))</td>
<td>1.28***</td>
<td>.143</td>
<td>328</td>
<td>3.61</td>
<td>2.73-4.78</td>
</tr>
<tr>
<td>Age (( \gamma_{10} ))</td>
<td>–.08</td>
<td>.045</td>
<td>1, 193</td>
<td>.92</td>
<td>.85-1.01</td>
</tr>
<tr>
<td>Gender (( \gamma_{20}; 0 = \text{male} ))</td>
<td>–.81***</td>
<td>.170</td>
<td>1, 193</td>
<td>.44</td>
<td>.32-1.62</td>
</tr>
<tr>
<td>Race (( \gamma_{30}; 0 = \text{White} ))</td>
<td>.46**</td>
<td>.161</td>
<td>1, 193</td>
<td>1.59</td>
<td>1.16-2.18</td>
</tr>
<tr>
<td>Pretreatment charges (( \gamma_{40} ))</td>
<td>.03***</td>
<td>.011</td>
<td>1, 193</td>
<td>1.03</td>
<td>1.01-1.06</td>
</tr>
<tr>
<td>SBP (( \gamma_{50}; 0 = \text{No} ))</td>
<td>.04</td>
<td>.147</td>
<td>1, 193</td>
<td>1.04</td>
<td>.78-1.39</td>
</tr>
<tr>
<td>Count model (&gt; 0 Posttreatment charges)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept (( \gamma_{00} ))</td>
<td>–2.02***</td>
<td>.057</td>
<td>305</td>
<td>.13</td>
<td>.12-15</td>
</tr>
<tr>
<td>Age (( \gamma_{10} ))</td>
<td>.04</td>
<td>.023</td>
<td>920</td>
<td>1.04</td>
<td>.99-1.09</td>
</tr>
<tr>
<td>Gender (( \gamma_{20}; 0 = \text{male} ))</td>
<td>–.40***</td>
<td>.085</td>
<td>920</td>
<td>.67</td>
<td>.57-79</td>
</tr>
<tr>
<td>Race (( \gamma_{30}; 0 = \text{White} ))</td>
<td>.08</td>
<td>.074</td>
<td>920</td>
<td>1.08</td>
<td>.95-1.25</td>
</tr>
<tr>
<td>Pretreatment charges (( \gamma_{40} ))</td>
<td>.02***</td>
<td>.004</td>
<td>920</td>
<td>1.02</td>
<td>1.01-1.03</td>
</tr>
<tr>
<td>SBP (( \gamma_{50}; 0 = \text{no} ))</td>
<td>–.06</td>
<td>.082</td>
<td>305</td>
<td>.94</td>
<td>.80-1.11</td>
</tr>
</tbody>
</table>

NOTE. SBP = sexual behavior problems. The T ratio test statistic for each parameter (omitted) was computed as \( \gamma_j / SE \).

*\( p < .05 \), **\( p < .01 \), ***\( p < .001 \).
the current study complements the findings of Carpentier and colleagues' (2006) 10-year follow-up study that reached similar conclusions.

Specific hypotheses regarding treatment outcome were upheld. CBCL Externalizing and Internalizing scores and VFI proportion scores all dropped significantly for both groups over time, with an initial linear effect subsequently tapering off across the 12-month posttreatment follow-up. It is important to note that both groups evidenced statistically significant improvement that was maintained through final follow-up. As hypothesized, youth in the SBP group, who started treatment with higher rates of behavior problems, continued to exhibit higher rates of behavior problems through follow-up, relative to their counterparts in the NSBP group. It is relevant to note that mean scores for both groups dropped from clinical to nonclinical or subclinical levels across measures. These results suggest that, though the groups’ trajectories remained different over time, these differences might not be clinically meaningful. This conclusion is further supported by the lack of between-group differences in rates of posttreatment offense charges. Groups had nearly identical probabilities of having been charged with any posttreatment offense, posttreatment person offenses, or posttreatment sexual offenses, with sexual offense charges being rare.

The clinical implications of the current study, considered together with the growing body of SBP research, are considerable. As noted at the beginning of this article, there appears to be increasing concern that youth with SBP are resistant to standard treatments and at increased risk for engaging in future sexual crimes. Consequently, such children have been targeted for restrictive treatment placements (e.g., specialized SBP units within residential facilities; see Carpentier et al., 2006). It appears, however, that SBP can be effectively addressed with empirically supported, community-based interventions, including MST and specialized CBT interventions that substantively engage caregivers (Carpentier et al., 2006; Chaffin et al., 2002; Cohen & Mannarino, 1996, 1997; Deblinger & Heflin, 1996; Letourneau et al., 2004; Silovsky et al., 2007). To date, there is no empirical support for restrictive treatment settings, although it has been suggested that such settings might benefit a small portion of youth who exhibit particularly dangerous SBP and/or whose problem behaviors continue beyond outpatient treatment (Carpentier et al., 2006; Chaffin et al., 2002). Future research should endeavor to shed light on whether restrictive treatment settings offer any advantages over empirically validated community-based settings for any youths.

The implications for public policy are also potentially profound. The sexual behavior of minors—even between consenting peers and sometimes between young children—has been criminalized in many states (Caldwell, 2002; Carpentier et al., 2006; Trivits & Reppucci, 2002), and the recent spate of laws specifically targeting sexual offenders often include (or fail to exclude) minors. For example, the newly enacted federal Adam Walsh Child Protection and Safety Act of 2006 requires the public registration of children as young as age 14 years if convicted of certain sexual crimes. The current study does not address sexual offender recidivism, as none of the youths had pretreatment sexual offenses. However, the extant research strongly suggests that youth with SBP, including youth with charges for sexual offenses, are unlikely to commit future sexual crimes, particularly when they have completed empirically validated interventions (Borduin, Henggeler, Blaske, & Stein, 1990; Borduin & Schaeffer, 2001; Caldwell, 2002; Letourneau et al., in press; Nisbet, Wilson, & Smallbone, 2004; Reitzel & Carbonell, 2006; Walker, McGovern, Poey, & Otis, 2004; Worling & Curwen, 2000). Policies that require lengthy incarceration or residential treatment, that sentence children and adolescents as adults, or that impose post-sentence-completion restrictions on children and adolescents (e.g., in the form of indeterminate civil commitment, sex offender registration, and community notification) fail to recognize the tremendous potential of these youths, with the help of their families and evidence-based interventions, to overcome early mistakes and even early but serious criminal acts (Letourneau & Miner, 2005; MacArthur Foundation Research Network, 2006).

Interpretations of study findings should be influenced by study strengths and limitations. Specific strengths of the current study include the large, geographically and ethnically diverse sample, the extended follow-up of treatment outcome and the average 4-year follow-up of criminal charges. Of special relevance is the fact that many participants had reached early adulthood when criminal records were requested. Youth remain at increased risk for criminal behavior through early adulthood, and results of prospective studies are more generalizable (and presumably more valid) when youth are followed through this important developmental period (Farrington, 1986; R. K. Hanson & Bussière, 1998; Loeber & Farrington, 2000). Longer follow-up is also essential when examining low base-rate events such as...
sexual offending. Although the percentage of study participants with posttreatment sexual charges was low, a contribution of the current study to the field is that the absolute number of youths committing new sexual offenses (n = 29) was substantial, relative to many other studies examining this phenomenon. Furthermore, the current study contributes to the small sample of prospective (vs. retrospective) studies of problem sexual behavior in youth and, as such, is less likely to provide inflated estimates of the risk for sexual offending (Carpentier et al., 2006).

Some limitations must also be noted. Sexual offenses are underreported; and, thus, official charge data likely underestimate the rate of sexual offending, although this limitation should apply equally to youths in both SBP groups. As noted by Carpentier and colleagues (2006), the addition of self-reported information would likely improve our understanding regarding the developmental trajectories of children and youths with SBP. Likewise, youths who moved out of state might have committed offenses not captured by relying on local or state juvenile justice agencies; although again, there is no reason to suspect that this limitation applied differentially to the SBP groups. As noted in our previous study (Letourneau et al., 2004), the absence of a treatment control group precludes attributing with certainty symptom reduction to the effects of treatment. Regression to the mean remains an alternative hypothesis to treatment success, although the magnitude of CBCL and VFI problem score decreases in the current study mirror those found in previous randomized clinical trials of MST. A more significant limitation of the current study is that generalizability might be limited by the use of the CBCL. Sex Problems scale to identify SBP. The items on this scale are few in number and relatively limited in scope. This concern is mitigated by the fact that results from the current study are similar to those from studies that utilized more comprehensive measures of SBP (e.g., Carpentier et al., 2006) and given the significant relationships found in previous research between CBCL Sex Problems scale scores and more robust measures of SBP (Bonner et al., 1999; Friedrich et al., 1992; Pithers et al., 1998). Whether MST can successfully treat children or adolescents with more serious SBP, however, cannot be determined by the current study. In addition, though we believe the age range of children and adolescents included in the current study to be a strength, it is important to acknowledge that the expression of sexual behavior, including problem behaviors, changes across developmental stages (DeLamater & Friedrich, 2002; Friedrich et al., 1991).

Different patterns of results might have emerged had more restricted age groups been examined. Last, the current study made use of the same measure, the CBCL, to provide the grouping variable and indicators of outcome. The pretreatment between-group differences in severity might be an artifact of this process, and future research should seek to obtain independent indicators of baseline clinical severity.

In summary, results from the current study should allay concerns that children and adolescents who present with SBP similar to those captured by CBCL are at increased risk for treatment failure or for subsequent sexual offending. Rather, these results demonstrate that though youth with SBP apparently represent a substantial minority of delinquent youth referred for treatment (or at least those referred for MST), these youths appear to respond well to intensive, caregiver-focused treatment and are no more likely to commit future sexual offenses than delinquent youth without SBP when effectively treated.

REFERENCES


Letourneau et al. / TREATMENT OUTCOME


Elizabeth J. Letourneau, PhD, has conducted research on adult and juvenile sexual offending since 1989. She graduated from Northern Illinois University in 1995. She is currently associate professor, Family Services Research Center, Medical University of South Carolina, where she has focused her research on treatment effectiveness with juvenile sex offenders and other juvenile delinquents, and the effects of legal policies such as registration and notification on judicial decision making. She has numerous publications in this field and is principal investigator for several grant-funded projects that aim to examine the intended and unintended effects of sex offender registration and notification policies. She has been active in the Association for the Treatment of Sexual Abusers (ATSA) since 1991.

Sonja K. Schoenwald, PhD, is professor of psychiatry and behavioral sciences at the Medical University of South Carolina and was associate director of the Family Services Research Center there from 1994-2004. She is among the leading clinical services researchers in the United States on issues relating transportability, implementation, and dissemination of effective treatments. She pioneered the development, refinement, and empirical testing of the quality assurance protocols used to transport multisystemic therapy (MST) to usual care settings. She has cultivated National Institutes of Health (NIH) and Foundation-funded collaborations with leading treatment, services, organizational, and economics researchers across geographic distances and institutional affiliations. She is principal investigator of National Institute of Mental Health (NIMH)- and National Institute on Drug Abuse-funded studies on treatment transportability, and coinvestigator on NIMH-funded randomized trials testing the effects of innovative organizational and service delivery interventions on the implementation and outcomes of evidence-based treatments for youth in usual care settings. She has published numerous peer-reviewed papers and book chapters and has coauthored two books and several treatment manuals and monographs for diverse stakeholder groups focused on supporting the implementation of effective treatments in usual care settings.

Jason E. Chapman, PhD, is an assistant professor in the Family Services Research Center in the Department of Psychiatry and Behavioral Sciences at the Medical University of South Carolina. He is coinvestigator and data analyst for multiple National Institute on Drug Abuse- and National Institute of Mental Health-funded clinical and services research projects focusing on adolescent substance abuse and the transportability and dissemination of evidence-based practices to usual care settings. His statistical work focuses on the application of mixed-effects regression models to longitudinal and nested data with linear and nonlinear outcomes, as well as the application of Rasch-based measurement models to the development and refinement of research instruments.