20-Year Prospective Follow-Up Study of Specialized Treatment for Adolescents Who Offended Sexually

James R. Worling, Ph.D., C.Psych.*, Ariel Litteljohn, M.A., and David Bookalam, M.A.

Most follow-up investigations of the effectiveness of specialized treatment for adolescents who have offended sexually have not included a comparison group. Furthermore, the average length of most previous studies is approximately 5 years. This investigation is a 10-year extension of our prospective, 10-year follow-up study of specialized treatment (Worling & Curwen, 2000). Recidivism data (criminal charges) were collected from a national database for 148 adolescents who had offended sexually. Adolescents were between 12 and 19 years of age ($M = 15.5$; $SD = 1.5$) at assessment, and the follow-up interval spanned from 12 to 20 years ($M = 16.23$; $SD = 2.02$). Relative to the comparison group ($n = 90$), adolescents who participated in specialized treatment ($n = 58$) were significantly less likely to receive subsequent charges for sexual, nonsexual violent, and nonviolent crimes. These data add to the growing body of research supporting the effectiveness of specialized treatment for individuals who have offended sexually.

Copyright © 2009 John Wiley & Sons, Ltd.

In 1987, we began a prospective follow-up study of the effectiveness of specialized treatment at the Sexual Abuse: Family Education and Treatment (SAFE-T) Program in Toronto, Canada for adolescents who had sexually offended. We (Worling & Curwen, 2000) published our initial results after a follow-up period that ranged from 2 to 10 years ($M = 6.23$; $SD = 2.02$). Data were collected for both a treatment group and a comparison group, and we established that the groups were not significantly different with respect to a wide number of variables that could have influenced outcome. Using criminal charges as an estimate of recidivism, we found that those who participated in treatment were significantly less likely to reoffend both sexually and nonsexually up to 10 years following the initial assessment. At the time, our investigation was the only published study in which a mean follow-up period beyond 4 years was employed, and it

*Correspondence to: James R. Worling, Ph.D., C.Psych., Consulting Clinical & Forensic Psychologist, Sexual Abuse: Family Education & Treatment Program (SAFE-T), Ontario Ministry of Children and Youth Services, 51 Panorama Court Toronto, Ontario, Canada M9V 4L8. E-mail: jworling@can.net

We thank Barbara Rodgers for her unfailing support of research at the SAFE-T Program and for her encouragement to continue our prospective follow-up investigation. We are also greatly indebted to the adolescents who participated in this investigation, and to Youth Justice Services, Central Region, for their financial support and encouragement. The views expressed herein are those of the authors and do not necessarily reflect those of the Thistletown Regional Centre or the Ontario Ministry of Children and Youth Services.
was the only study to employ a national database of criminal charges as opposed to only state and/or local records.

Despite the proliferation of specialized programs since the early 1980s to provide treatment, there are still relatively few studies of the effectiveness of treatment for this population. In their comprehensive, narrative review, Fortune and Lambie (2006) summarized 28 published studies of specialized treatment. They found that only seven of the studies included a comparison group, and only five investigations employed a mean follow-up period beyond 5 years. Fortune and Lambie concluded that, although recidivism rates for treated youth are typically lower than recidivism rates for those who did not receive treatment, methodological problems make it difficult to draw conclusions regarding the outcome of specialized treatment. Fortune and Lambie stressed that researchers should improve research designs by means such as using a comparison group and collecting long-term outcome data.

In an attempt to quantify a treatment effect, Walker, McGovern, Poey, and Otis (2004) conducted a meta-analysis of 10 published and unpublished studies of treatment for adolescents who had offended sexually. They reported an average weighted effect size ($r$) of .37. Although this result is somewhat encouraging, one cannot conclude that treatment necessarily reduces the risk of recidivism, as only 3 of the 10 studies in this meta-analysis used recidivism as an outcome variable. Rather, effect size calculations were based on a blend of dependent variables including psychological test scores, measurements of sexual arousal, and recidivism rates. Furthermore, only 2 of the 10 studies in this meta-analysis employed a comparison group.

More recently, Reitzel and Carbonell (2006) conducted a meta-analysis of the effectiveness of treatment for adolescents who had sexually offended. They examined nine published and unpublished studies in which researchers compared a treated group to a control group (including our 2000 investigation). With a combined sample of 2,986 adolescents, it was found that only 7.37% of the treated youth were subsequently charged and/or convicted for sexual reoffenses compared with 18.93% of the youth in the comparison groups.

Although the results of this meta-analysis are promising, it is important to note that the average follow-up period was only 59 months. It has been shown that longer follow-up periods result in higher recidivism rates for adults who offend sexually (see, e.g., Hanson, Morton, & Harris, 2003; Marshall, Jones, Ward, Johnson, & Barbaree, 1991). It has also been demonstrated that recidivism rates for adolescents who have offended sexually are positively correlated with the length of the follow-up interval (Fortune & Lambie, 2006; Gerhold, Browne, & Beckett, 2007; McCann & Lussier, 2008; Worling & Långström, 2006); however, there are very few studies that go beyond a 10-year period. The goal of the present study, therefore, was to examine the effectiveness of specialized treatment for adolescents who had offended sexually for a follow-up period of up to 20 years.

**METHOD**

**Treatment Program**

The Sexual Abuse: Family Education and Treatment (SAFE-T) Program is a specialized, community-based program that provides sexual abuse-specific assessment,
treatment, consultation, and long-term support to (1) children and/or adolescents who have been sexually abused within their families, and their families, (2) children with sexual behaviour problems, and their families, and (3) adolescents who have sexually offended, and their families. Following comprehensive clinical and psychometric assessments, treatment plans are individually tailored for each youth and his/her family, and treatment goals are reviewed approximately every six months (see Worling, 1998, for a brief description of treatment at the SAFE-T Program). Treatment at the SAFE-T Program is individualized to meet each adolescent’s unique strengths and needs; however, common themes include increasing insight and accountability for past offending, developing offense-prevention plans, enhancing awareness of victim impact, enhancing social relationships, reducing the impact of traumatic past events, enhancing family communication and relationships, and enhancing prosocial sexual attitudes and knowledge.

The SAFE-T Program was originally developed to treat child and adolescent survivors of sexual assault, and it evolved from a mental health service focused on adolescents and their families. In contrast, many of the programs developed in the mid-1980s to treat adolescents who offended sexually were based on programs designed for adults who offended sexually. As a result of the family focus at the SAFE-T Program, treatment providers have always made efforts to work collaboratively with youth and their parents to simultaneously address the risk to reoffend sexually and other important clinical needs that might be present (e.g., traumatic distress, antisocial attitudes, depression, and social problems) while building on individual and family strengths.

The SAFE-T Program continues to provide specialized services to adolescents who have offended sexually, and their families. Although the goals of treatment have remained fairly constant since the program began, treatment has changed somewhat over the past 22 years. For example, the average length of treatment is now approximately 16 months (rather than 24 months as noted in our previous study), as the intensity of treatment is more closely matched to the estimated level of risk of reoffending. Also, discussions of an individual’s sexual interests, and the details of past sexual crimes, are now typically limited to individual treatment sessions; group interventions are focused more on the development of the skills and attitudes necessary for healthy interpersonal and sexual relationships.

Participants

Study participants were 148 adolescents (139 males and 9 females) who were convicted of and/or acknowledged a sexual offense, as defined by the Criminal Code of Canada. The adolescents were assessed at the SAFE-T Program between October 1987 and October 1995, and they were between 12 and 19 years of age ($M = 15.5; SD = 1.5$) at the time of assessment. The adolescents had sexually assaulted victims who were intrafamilial (28%), extrafamilial (55%), or both (17%); female (61%), male (16%), or both (23%); children (55%), peers/adults (35%), or both (10%). Most adolescents (98%) were referred for offenses involving direct physical contact with their victims; three adolescents were referred for exhibitionism. At the time of our assessment, the youth were living either at home (47%), in custody facilities (25%), in group homes (19%), in foster homes (6%), or with friends or extended family (3%). It should also be pointed out that none of the adolescents were below borderline intellectual functioning.
The treatment group consisted of 58 adolescents (53 males and 5 females) who participated in at least 10 months of specialized treatment at the SAFE-T Program. We classified adolescents who dropped out after 12 months as members of the treatment group. Note that almost one-third (18 out of 58) of the adolescents in the treatment group dropped out before completing treatment but subsequent to 12 months participation. All adolescents received individual therapy, and most (71%) also participated in both group and family therapy. The average length of treatment, overall, was 24.43 months ($SD = 10.72$). Given that treatment is tailored to meet the unique strengths, risks, and needs of each adolescent, there is no prescribed length of treatment.

There were 90 adolescents in the comparison group (86 males and 4 females). The largest subgroup ($n = 46$) received only an assessment by staff from the SAFE-T Program. The majority (30/46) of this group were recruited specifically to serve as comparison participants, and they were receiving treatment elsewhere. The remaining adolescents who received an assessment only (16) were referred to the SAFE-T Program only for an assessment, typically because the youth were going to be receiving treatment elsewhere. The comparison group also included 17 adolescents who refused treatment (treatment refuser), and 27 adolescents who dropped out of treatment before a 12-month period (treatment dropout). Overall, 67% of the adolescents in the comparison group received some form of treatment outside of the SAFE-T Program. Unfortunately, information regarding the nature and duration of treatment for this group was unavailable.

**Prospective Recidivism Data**

As in the original study, recidivism data for the 20-year follow-up period (October 1987 to October 2007) were gathered from the Canadian Police Information Centre, which provides a national registry of charges and convictions maintained by the Royal Canadian Mounted Police. Access to these data was granted through an order from a youth court judge. In the present study, the follow-up period ranged from a minimum of 12 years to a maximum of 20 years ($M = 16.23; SD = 2.02$). Criminal charges (rather than convictions) were once again used as our measure of recidivism, as this less conservative measure more closely approximates true reoffense rates (Gerhold et al., 2007; Worling & Långström, 2006).

Recidivism data were categorized as follows: sexual offenses (any Canadian Criminal Code offense of a sexual nature); violent nonsexual offenses (any criminal charges involving actual or threatened violence towards a person such as assault, assault with a weapon, robbery, or uttering death threats); and nonviolent offenses (nonviolent criminal offenses such as theft, break and enter, weapon possession, trafficking in narcotics, or breach of probation, for example). In addition to observing simply the presence or absence of new charges, as was done in our original study, the number of different charges that recidivists accumulated over the 20-year period was also tabulated for the present investigation. Furthermore, we examined the number of different reoffense episodes—i.e., discrete days on which a recidivist reoffended. Finally, we also examined the number of recidivists who accrued new sexual offense charges as adults (i.e., age 18 and over).
Materials

A number of psychological tests were used at the time of assessment to identify individual strengths, risks, and needs. Several scores from these tests were also used to examine pretreatment differences between groups. Although most of these measures are no longer used at the SAFE-T Program (with the exception of the revised Youth Self-Report), the scales that were used are outlined below (see Worling & Curwen, 2000, for a detailed description of these measures and their psychometric properties):

- three scales from the Assessing Environments (III) Scale (AEIII; Berger, Knutson, Mehm, & Perkins, 1988): Feelings of Parental Rejection, Negative Family Atmosphere, and Physical Punishment
- two scales from the Tennessee Self-Concept Scale (TSCS; Roid & Fitts, 1988): Self-esteem and Self-criticism (an estimate of socially desirable responding)
- three scales from the Youth Self-Report (YSR; Achenbach, 1991): Social Problems, Aggressive Behavior, and Delinquent Behavior
- the total score from the Beck Depression Inventory (BDI; Beck & Steer, 1987) as a measure of depression
- the total score from the Buss–Durkee Hostility Inventory (BDHI; Buss & Durkee, 1957) to assess hostility
- the Socialization (So) scale from the California Psychological Inventory (CPI; Gough, 1987) as a measure of impulsive and antisocial traits commonly ascribed to psychopathy

Procedure

Participants completed the self-report measures during the initial assessment. If an individual experienced difficulty reading, the questionnaires were read aloud by either one of the investigators or an assistant. Although most of the above-listed tests were used at the inception of the study, others were available only at a later date. Furthermore, there were many occasions when it was not deemed clinically appropriate to give the entire battery of tests to a participant. As a result, not all adolescents completed every test. Demographic and offense-related data were collected from clinicians shortly after the adolescents were assessed.

RESULTS

As we noted in our original investigation, the treatment group and the three comparison groups were not significantly different with respect to many variables that may have impacted on subsequent recidivism (see Worling & Curwen, 2000, Tables 2 and 3). In particular, groups were not significantly different with respect to personal characteristics (age at assessment, gender, socioeconomic status, family composition, place of residence), offense characteristics (presence of previous charges, victim gender,
relationship to victim, number of sexual offense victims), or any of the test scores (listed in Materials above) that were examined (i.e., test scores reflecting delinquency, aggression, social difficulties, hostility, depression, self-esteem, socially desirable responding, rape-supportive attitudes and interests, attitudes and interests supportive of child sexual assault, physical abuse, parental rejection, and a hostile family environment). As such, we combined the assessment only, treatment dropout, and treatment refuser participants to form the comparison group, and it was not necessary to control for these variables when comparing treatment and comparison participants.

The 20-year recidivism rates for the treatment and comparison groups are presented in Table 1, along with the results from our original, 10-year follow-up study (Worling & Curwen, 2000). Over the 20-year follow-up period, the overall rate (for all 148 participants) for any, nonviolent, nonsexual violent, and sexual recidivism was 49.32% (73 of 148), 42.57% (63 of 148), 32.4% (48 of 148), and 16.22% (24 of 148), respectively. Kaplan–Meier survival functions were calculated to compare the 20-year recidivism rates between treatment and comparison groups (see Figures 1–4). Tests for differences between survival functions are reported as $\chi^2$ values based on the log rank statistic. There were significant differences between the treatment and comparison groups after 20 years for all categories of reoffending. Specifically, adolescents who received treatment were significantly less likely to be charged for a sexual reoffense, $\chi^2(1, N=148) = 4.41, p < .05$, a nonsexual violent reoffense, $\chi^2(1, N=148) = 4.35, p < .05$, a nonviolent reoffense, $\chi^2(1, N=148) = 10.57, p < .001$, or any criminal reoffense, $\chi^2(1, N=148) = 6.37, p < .05$. It is also readily apparent from Table 1 that there were only very small (2–7%) increases in recidivism rates across the various offense categories for both the treatment and comparison groups during the second 10-year follow-up interval.

In addition to the simple count of the number of adolescents who reoffended, we also examined the recidivism data from a harm-reduction standpoint: i.e., the number of new criminal charges and the number of discrete episodes (days) involving new charges. The data in Table 2 show that there was a general trend for the recidivists from the comparison group to accumulate a higher average number of charges and to have reoffended on a higher average number of different occasions; however, these differences were not significant. Given that the homogeneity-of-variance assumption was violated for most of these comparisons, Mann–Whitney $U$ tests were computed. No group differences were significant, all $p > .05$.

Table 1. Differences between 10- and 20-year recidivism rates for treatment and comparison groups

<table>
<thead>
<tr>
<th>Recidivism (charges)</th>
<th>(10 year follow-up, Worling &amp; Curwen, 2000)</th>
<th>(20 year follow-up, present investigation)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatment</td>
<td>Comparison</td>
</tr>
<tr>
<td>Any</td>
<td>35% (20/58)</td>
<td>54% (49/90)</td>
</tr>
<tr>
<td>Nonviolent*a</td>
<td>21% (12/58)</td>
<td>50% (45/90)</td>
</tr>
<tr>
<td>Violent nonsexual*b</td>
<td>19% (11/58)</td>
<td>32% (29/90)</td>
</tr>
<tr>
<td>Sexual</td>
<td>5% (3/58)</td>
<td>18% (16/90)</td>
</tr>
</tbody>
</table>

All differences between treatment and comparison groups are significant, all $\chi^2$ (log rank) $> 4.40$, all $p < .05$.

*aIncludes offenses such as theft, break-and-enter, breach of probation, escape lawful custody, possession of a weapon, trafficking in narcotics, and driving while impaired.

*bIncludes offenses such as assault, robbery, forcible confinement, and uttering death threats.
We also examined the number of participants who reoffended sexually as adults (i.e., aged 18 and over). Recall that the average age of participants at the time of this follow-up period was 31.5 years ($SD = 1.5$). Overall, only 11.49% (17 of 148) of the participants were charged for sexual offenses as adults. Interestingly, of the 17 participants who were charged with a sexual crime as an adult, most (15 of 17) were adults when they were first charged with a sexual reoffense. Just 2 of the 17 who were charged with a sexual reoffense as an adult were also charged with a sexual reoffense as a youth. Of the 24 sexual recidivists in this investigation, therefore, 29.17% (7 of 24) were charged for sexual reoffenses only as adolescents, whereas 70.83% (17 of 24) were charged for new sexual offenses as young adults.

**DISCUSSION**

The results of the present investigation indicate that specialized treatment for adolescents who offended sexually led to significant reductions in both sexual and nonssexual recidivism after a follow-up period that ranged from 12 to 20 years ($M = 16.23$ years; $SD = 2.02$). Only 9% of those adolescents who participated in at least 10 months of specialized treatment were charged with a new sexual offense during this follow-up period, whereas 21% of those adolescents who did not receive specialized
treatment were charged for a subsequent sexual offense. Furthermore, given that treatment at the SAFE-T Program was not focused solely on sexual recidivism risk, it was encouraging to find that those adolescents who participated in treatment were also significantly less likely to be charged for nonsexual crimes relative to the comparison group.

From a harm-reduction standpoint, although there was a general trend for recidivists from the comparison group to accumulate more charges when compared to recidivists from the treatment group, and reoffend on more discrete occasions, these differences were not statistically significant. This may be partially attributable to the tremendous variability in these data and the relatively small number of recidivists. For example, although the average recidivist in this study received 7 criminal charges for nonviolent offenses on approximately 3 different occasions, there were two recidivists in the comparison group who each accumulated over 40 nonsexual criminal charges on over 19 different occasions. It will be important to determine whether other researchers can demonstrate that those adolescents who receive treatment but subsequently recidivate do so at a lower frequency relative to untreated youth.

In their meta-analysis of treatment outcome studies for adolescents who had committed sexual offenses, Reitzel and Carbonell (2006) reported a 5-year sexual recidivism rate of 12.53%—collapsed across treatment and comparison groups. This is quite similar to the 5-year sexual reoffense rate of 14% found for adults (Hanson et al., 2009).
With our mean follow-up period of 16 years, however, we found an overall sexual recidivism rate (collapsed across treatment and comparison groups) of approximately 16%. This is markedly lower than the 15-year sexual recidivism rate of 24% reported for adults who offend sexually (Hanson et al., 2003). From an inspection of Figures 1–4, it appears that most sexual and nonsexual recidivism occurs in the first few years after adolescents are initially assessed. Indeed, there appears to be a significant flattening of the slope of the survival curves at about the 10-year mark for both treatment and comparison participants. This change occurs when the participants would be, on average, about 25 years of age. This observation is certainly consistent with research suggesting that the brain processes that inhibit risky teenage behaviors continue to develop well past adolescence and into the early and mid-20s (e.g., Casey, Getz, & Galvan, 2008; Steinberg, 2005; Yurgelen-Todd, 2007). It is not surprising, therefore, that we observed only very small (2–7%) increases in recidivism rates for both treatment and comparison participants across offense categories during the second decade of this follow-up investigation.

Public policy in the United States regarding adolescents who offend sexually seems to be predicated on the notion that these youth are very likely to commit sexual offenses as adults (Zimring, Piquero, & Jennings, 2007). In the present study, however, only 11.5% of the participants were charged as adults for subsequent sexual offenses—up to an average age of approximately 31 years. In a prospective follow-up study from Australia, Nisbet, Wilson, and Smallbone (2004) found that, by an average age of 23.7
years, only 9% of adolescents who had offended sexually were charged as adults for subsequent sexual offenses. In the United States, Zimring et al. (2007) examined retrospective data for three cohorts of adolescents who had offended sexually. They reported that only 8.5% of this sample received adult criminal charges for sexual crimes up to age 32. It would appear, therefore, that most adolescents who offend sexually are not charged for subsequent sexual crimes by the time they are in their mid-20s to early 30s. Individuals who are first charged for sexual crimes in adulthood seem to have a different recidivism trajectory over time.

Table 2. Mean (and SD) number of reoffense charges and discrete reoffense episodes

<table>
<thead>
<tr>
<th>Type of charge/episode</th>
<th>Treatment recidivists</th>
<th>Comparison recidivists</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Sexual reoffense charges</td>
<td>5</td>
<td>3.00 (2.74)</td>
</tr>
<tr>
<td>Sexual reoffense episodes</td>
<td>5</td>
<td>1.00 (0.71)</td>
</tr>
<tr>
<td>Nonsexual violence reoffense charges</td>
<td>13</td>
<td>2.92 (3.32)</td>
</tr>
<tr>
<td>Nonsexual violence reoffense episodes</td>
<td>13</td>
<td>1.92 (2.18)</td>
</tr>
<tr>
<td>Nonviolent reoffense charges</td>
<td>16</td>
<td>3.44 (6.25)</td>
</tr>
<tr>
<td>Nonviolent reoffense episodes</td>
<td>16</td>
<td>1.18 (2.59)</td>
</tr>
<tr>
<td>Any reoffense charges</td>
<td>22</td>
<td>5.27 (7.74)</td>
</tr>
<tr>
<td>Any reoffense episodes</td>
<td>22</td>
<td>2.95 (4.06)</td>
</tr>
</tbody>
</table>

None of the treatment vs. comparison differences were significant.
Unlike most treatment outcome studies that have been published with adolescents who have offended sexually, where the mean follow-up period is approximately 5 years, the follow-up period in this investigation ranged from 12 to 20 years. Additional methodological advantages of the present study include the fact that we employed a national registry of criminal charges rather than state/provincial or local records, and we established that treatment and comparison participants were not significantly different on a large number of pretreatment variables that may have influenced the outcome. Furthermore, the present investigation was prospective rather than the more typical retrospective investigation that is based on file review. As such, the participants herein provided assessment data for this study 12–20 years ago when they were originally assessed. Finally, it is also important to point out that we used criminal charges as our estimate of recidivism rather than a more conservative estimate such as convictions.

Despite these strengths, however, the most significant threat to the validity of these findings regarding the long-term effectiveness of specialized treatment is the lack of random assignment to treatment. Although treatment and comparison participants were not significantly different with respect to a large number of variables that could have impacted on recidivism, it is possible that the groups varied on some untapped variables that were related to subsequent recidivism. Another limitation is that we relied on official reoffense statistics. Of course, this particular limitation impacts virtually all recidivism research in this field, as most victims of a sexual crime never report their victimization to authorities (see, e.g., Brennan & Taylor-Butts, 2008). As we pointed out previously (Worling & Curwen, 2000), official recidivism statistics are obviously a considerable underestimate of actual reoffending, given that researchers who utilize official data are dependent on victimized individuals coming forward with a disclosure, the disclosure being passed along to the criminal justice system, an investigation being conducted by police, sexual offense charges being laid, the timely entry of the charge into a database, and the accurate retrieval of this information by the researcher. If criminal convictions are being used to estimate recidivism, then researchers are additionally dependent on the charges not being dropped, the charge not being altered to a nonsexual charge through plea bargaining, and/or the outcome of the court process. Finally, it is also important to be mindful of the sample size in this investigation. Larger samples will likely provide for more generalizable results.

**CONCLUSION**

The results of this investigation suggest that specialized treatment for adolescents who offend sexually leads to significant reductions in both sexual and nonsexual reoffending—even up to 20 years following the initial assessment. These data add to the growing body of research supporting the effectiveness of specialized treatment for individuals who have offended sexually (Lösel & Schmucker, 2005; Reitzel & Carbonell, 2006). It will be important for researchers to continue to investigate not only the overall impact of specialized treatment for adolescents who offend sexually, but also the specific mechanisms of treatment that lead to significant changes for these youth and their families. The results of this investigation also support the finding that only a minority of adolescents who offend sexually are likely to be charged for sexual crimes by their late 20s or early 30s (Nisbet et al., 2004; Zimring et al., 2007).
REFERENCES


