A 25-Year Follow-Up of Cognitive/Behavioral Therapy with 7,275 Sexual Offenders
Barry M. Maletzky and Cynthia Steinhauser
*Behav Modif* 2002; 26: 123
DOI: 10.1177/0145445502262001

The online version of this article can be found at:
http://bmo.sagepub.com/cgi/content/abstract/26/2/123
Outcome data are presented, grouped into 5-year cohorts, for 7,275 sexual offenders entering a cognitive/behavioral treatment program. Assessment variables included treatment completion, self-admission of covert and/or overt deviant behaviors, the presence of deviant sexual arousal, or being recharged for any sexual crime (regardless of plea or conviction). It proved possible to follow 62% for the cohort at 5 years after initiating treatment, but follow-up completion rates decreased with time. Outcomes were significantly different based on offender subtype, with child molesters and exhibitionists achieving better overall success than pedophiles or rapists. Prematurely terminating treatment was a strong indicator of committing a new sexual offense. Of interest was the general improvement of success rates over each successive 5-year period for many types of offenders. Unfortunately, failure rates remained comparatively high for rapists (20%) and homosexual pedophiles (16%), regardless of when they were treated over the 25-year period. Implications for clinical practice and future research are drawn.

A 25-Year Follow-Up of Cognitive/Behavioral Therapy With 7,275 Sexual Offenders

BARRY M. MALETZKY
Oregon Health Sciences University
CYNTHIA STEINHAUSER
The Sexual Abuse Clinic

Since 1989, when Furby, Weinrott, and Blackshaw published their review of treatment outcomes for the sexual offender, pessimism has been expressed not only about the efficacy of treatment (McConaghy, 1998) but about whether efficacy can ever be demonstrated (Barbaree, 1997; Quinsey, Harris, Rice, & Lalumiere, 1993). Recidivism studies have been shown to be insensitive to treatment effects due to low base rates and hence the need for large numbers of participants. Moreover, most studies cited as providing evidence for positive treatment effect suffered one or more methodologic defects, including lack of use of...
comparable control groups, lack of operational definitions of treatment procedures, and lack of standardized assessment instruments.

However, a more sanguine view has been offered by several students of the literature who are also experienced clinicians. Marshall and Barbaree (1990) reviewed four institutionally based programs and an equivalent number of outpatient programs and concluded that although tentative, recommendations for sexual offenders can be made with an anticipation of a positive treatment effect that can be both scientifically demonstrated and clinically relevant. Hall (1995) has conducted a meta-analysis of the treatment outcome literature since the Furby et al. (1989) review and concluded that cognitive/behavioral treatments were significantly effective, with community-based treatment showing better effects than institutionally based treatment (confounded by seriousness of offense history). Higher recidivism rates were found in the majority of untreated, as opposed to treated, samples. One finding was especially optimistic: Longer follow-ups led to more significant treatment effects when comparisons to control groups were available.

In a recent meta-analysis, Alexander (1999) reviewed 79 treatment outcome studies encompassing almost 11,000 participants. She concluded that although treatment appeared effective, in that fewer than 11% of treated participants reoffended, the external validity of such analyses remains suspect because data collection and reporting procedures have not been standardized. Of even greater concern was the lack of exact definitions of sexual offender subtypes and treatment modalities.

Indeed, whereas the results of the most recent meta-analyses seem favorable, even this methodology has been called into question by Quinsey, Khanna, and Malcolm (1996) who, reviewing the same data, concluded that institutional treatment has yet to demonstrate efficacy. These authors pointed to three recent studies finding no treatment effect in institutionally based programs (Hanson, Steffy, & Gauthier, 1993; Marques, Day, Nelson, & West, 1994; Rice, Harris, & Quinsey, 1991). It is possible, however, that these negative findings reflect a Type II error, in which the null hypothesis that treatment lacks efficacy is due to insensitivity of routine statistical tests to detect such an effect, largely due to low Ns and low base rates of reoffending in certain
offender populations. Barbaree (1997) has shown that when base rates are low, significant chi-squares are not obtainable at any level of treatment effect and, hence, small differences may easily be missed. Similarly, large Ns are required to demonstrate significant results if moderate or small treatment effects occur. An average study would need an N of 495 to begin to achieve statistical significance.

To increase the sensitivity of treatment outcome research, studies are needed that address the problem of low base rates. Such studies would ideally involve large Ns over longer time periods to increase the chance that treatment effects would become apparent. Efforts in these directions are being made by several researchers. First, Day and Marques (1998) are carrying out an ongoing treatment evaluation project for institutionally based offenders that it is hoped, will be able to demonstrate a treatment effect (or lack of it) over a longer time period. However, relapse prevention in a group therapy format is the major treatment modality employed in that program. Some evidence indicates that relapse prevention alone may not be an optimum treatment strategy (Maletzky, 1998; McConaghy, 1998). Second, Dwyer (1997) continued to collect data on treatment provided to sexual offenders enrolled in a community-based treatment program. Although no control groups were possible, data appeared optimistic in this recent report, with total recidivism averaging 9%. Although the accumulating N (108) has enhanced confidence in this work, it still falls short of the Ns needed to more definitively assess treatment efficacy.

The present study extends the original database (Maletzky, 1993) and length of follow-up of a group of sexual offenders treated in a community setting. Whereas this setting provides comprehensive treatment reflecting the use of cognitive/behavioral techniques as typically practiced in North America today (Freeman-Longo, Bird, Stevenson, & Fiske, 1995), reports on large numbers of participants, and reports treatment response by offender diagnostic category, it does not satisfy all the requirements of a controlled study. Nonetheless, such clinically based studies may identify relevant factors to consider in the future evaluation of treatment programs and may single out treatment techniques worthy of further, more intensive research.

The following hypotheses were tested in the present study:
1. That treatment is effective in reducing relapse and preventing recidivism when assessed over longer time periods than previously reported.

2. That treatment results differ based on category of offender, with situational offenders achieving better outcomes than predatory or preferential offenders.

3. That failure rates increase with duration of time after treatment.

4. That recidivism constitutes one among a number of measurable factors related to treatment failure.

5. That participants terminating treatment prematurely are at high risk to reoffend.

6. That treatment techniques have been refined so that treatment success rates improved over the years of the study.

**METHOD**

A retrospective chart review was undertaken of all clients entering treatment in a community-based sexual offender clinic located in a metropolitan area in the northwestern United States. The clinic began enrolling clients in 1973. Data collection continued for the purposes of this study through 1997. A total of 7,275 clients met the study criteria of being male, 18 years of age or older, and a client whose chief mode of offending and type of victim selected could be sufficiently defined as belonging to one of the following categories:

2. Child molesters, male victims: Men who molested at most one male child.
3. Heterosexual pedophiles: Men who molested more than one female child and showed a preference for female children or a predatory style of offending.
4. Homosexual pedophiles: Men who molested more than one male child and showed a preference for male children or a predatory style of offending.
5. Exhibitionists: Men who exposed themselves and did not molest children or rape.
6. Rapists: Men who raped and did not molest children or expose.

In cases in which a participant committed more than one type of sexual offense, his major mode of offending was designated as his category. This did not always concur with his index offense. For example,
an offender who molested young girls occasionally exposed to them without physically molesting them. His most recent offense, which led to the referral, was technically for indecent exposure. However, he was designated a heterosexual pedophile rather than an exhibitionist for the purpose of data collection. For further definitions, particularly for situational or predatory offending, see Maletzky (1993). Categorization data are presented in Table 1.

Although treatment was provided in an outpatient setting, a number of dangerous offenders were treated, especially men on parole for predatory offenses against children or for rape. More than 90% of participants were under judicial supervision for their offenses throughout their treatment course. The average duration of treatment was 1¾ years.

Treatment techniques have been extensively described in the literature (Maletzky, 1991) and included aversive conditioning (primarily using olfactory stimuli), plethysmographic biofeedback, aversive behavior rehearsal, masturbatory reconditioning, vicarious sensitization, sexual impulse control training, relapse prevention, cognitive therapy, and empathy training. Because of the length of time over which data were collected, unavoidable variations in some techniques occurred. In addition, therapists were not constant, and differing and subtle shifts in emphasis may have occurred through the years. For example, clinic staff shifted focus away from earlier aversive conditioning techniques to aversive behavior rehearsal, relapse prevention, and cognitive therapy in recent years. Moreover, these data must be

<table>
<thead>
<tr>
<th>Offender</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child molesters, female victim</td>
<td>2,196</td>
<td>30.2</td>
</tr>
<tr>
<td>Child molesters, male victim</td>
<td>765</td>
<td>10.5</td>
</tr>
<tr>
<td>Heterosexual pedophiles</td>
<td>1,011</td>
<td>13.9</td>
</tr>
<tr>
<td>Homosexual pedophiles</td>
<td>1,251</td>
<td>17.2</td>
</tr>
<tr>
<td>Exhibitionists</td>
<td>1,604</td>
<td>22.0</td>
</tr>
<tr>
<td>Rapists</td>
<td>448</td>
<td>6.2</td>
</tr>
<tr>
<td>Totals</td>
<td>7,275</td>
<td>100.0</td>
</tr>
</tbody>
</table>

© 2002 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.
interpreted with caution in comparison to other programs because almost all participants were treated in individual therapy. Ongoing treatment for more than 95% of this sample was provided in one-to-one therapy. However, a majority of participants also took part in group therapy, usually a 12-week orientation group (Maletzky & Steinhauser, 1998).

Treatment techniques were individually applied to each offender based on perceived need. Whereas most offenders received all techniques mentioned, some received more of one than another, and a few failed to receive one or more approaches. Data were unavailable to determine which offenders received how much of each technique. Hence, the data cannot be used to decide if any particular technique was, by itself, effective, or to what extent it may have contributed to the overall outcome.

Unfortunately, these data were not amenable to an ANOVA or other statistical analyses due to differing methods of data collection through the years, lack of individualized data on participants during the early years of the project, and lack of a control group.

Independent variables included category of offender and length of follow-up. The dependent variable was treatment failure, as defined by the following:

1. Self-reporting of any instance of covert or overt deviant sexual behavior (see below).
2. Deviant sexual arousal greater than 20% to plethysmographic test stimuli at the end of treatment or at any follow-up session.
3. Deceit on any sexually related question on a polygraph at the end of treatment or at any session during follow-up.
4. Any sexual crime charged at any time during treatment or follow-up.

Note that a charge by itself, here defined as recidivism, constituted a reason for failure. Arrests and convictions were not required. This policy took into account the fact that many sexual allegations are dropped due to technical reasons or are altered to nonsexual charges for ease of prosecution. This policy also ensured a bias against successful treatment outcome, necessary for community safety. Note also that nonsexual charges, or even convictions, did not constitute grounds for failure in a sexual offender treatment program.
As part of clinic policy, attempts were made to continually assess all offenders following treatment, including those offenders who had terminated prematurely. Telephone contact was established within 1 month following the end of treatment and continued throughout the course of the study period. The following questions were asked during telephone interviews:

1. Have you had any fantasies, urges, or dreams about your sexual offending?
2. Have you committed any sexual acts or masturbated to fantasies about your sexual offending?

An answer of yes to Question 1 was termed a lapse (covert behavior) and to Question 2, a relapse (overt behavior). The above questions were asked at 6 months, 1 year, and then at annual intervals, and each participant was invited at that time to participate in a free interview, plethysmograph, and polygraph, with a guarantee of confidentiality. However, participants were informed that if a sexually dangerous situation was discovered, a report might be made to the proper authorities. During the follow-up period, a legal records search for each participant, with all known aliases, was conducted. Thus, a number of assessment techniques were available for all offenders during each time period.

Because many offenders had not established stable lifestyles or were in a transient state following treatment, follow-up was marked by difficulties, especially for those who discontinued treatment prematurely. Nonetheless, with persistence, contact with a reasonable percentage of men has continued. Table 2 presents percentages of participants available by telephone across all categories. (Later tables will show numbers available for the other assessment techniques.) Under these conditions, and without a complete guarantee of confidentiality, it is surprising that between 54% and 71% of participants were available at any 5-year period, and that, overall, data on 39% could be collected.

However, as expected, availability declined with duration of time following treatment, so that during any 1 year, contact was maintained with approximately 75% of offenders; by 5 years, this figure was approximately 59%, and at 25 years, just 32%. Moreover, those partic-
Participants making themselves available for study might have been pre-selected for positive outcomes. Countering this potential bias, however, were the strict criteria governing definitions of treatment success. Note also that attempts were made to follow participants dropping out of active treatment. Men unavailable for follow-up were simply not included in data analysis. A retrospective analysis did not reveal a differential number of noncontact rates among categories of offender. For example, nearly equivalent percentages of men who molested girls were available as men who exposed, or men who raped.

RESULTS

Table 3 lists the percentages of participants meeting criteria for treatment failure by offending category, including percentages of participants self-reporting deviant behavior, those charged with a new sexual crime, or those who left treatment prematurely, at each of the 5-year follow-up periods (1973 to 1977, 1978 to 1982, etc.), whereas Table 4 shows these failure rates by category over the entire 25-year span. Overall, most participants met criteria for treatment success at 5- and 25-year follow-ups, with child molesters and exhibitionists enjoying better outcomes than pedophiles and rapists. Men unavailable for follow-up were not included in the data analysis. Still, among all offenders entering a treatment program, a substantial number were

<table>
<thead>
<tr>
<th>Period (inclusive)</th>
<th>Number Beginning Treatment</th>
<th>Number Available at 5 Years (%)</th>
<th>Number Available in 1997 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1973-1977</td>
<td>1,137</td>
<td>683 (60.1)</td>
<td>360 (31.7)</td>
</tr>
<tr>
<td>1978-1982</td>
<td>1,742</td>
<td>1,014 (58.2)</td>
<td>542 (31.1)</td>
</tr>
<tr>
<td>1983-1997</td>
<td>1,801</td>
<td>975 (54.1)</td>
<td>559 (31.0)</td>
</tr>
<tr>
<td>1988-1992</td>
<td>980</td>
<td>652 (66.6)</td>
<td>452 (46.1)</td>
</tr>
<tr>
<td>1993-1997</td>
<td>1,615</td>
<td>1,154 (71.5)</td>
<td>924 (57.2)</td>
</tr>
<tr>
<td>Totals (1973-1997)</td>
<td>7,275</td>
<td>4,479 (61.6)</td>
<td>2,837 (39.0)</td>
</tr>
</tbody>
</table>
available for follow-up in most offense categories (see Table 3), strengthening reliance on the data sample. As expected, failure rates were higher for predatory and preferential offenders: Across the 25 years, rates were 6.3% for men who molested girls, 9.4% for men who molested boys, 9.7% for heterosexual pedophiles, 16.3% for homosexual pedophiles, 13.5% for exhibitionists, and 21.2% for rapists.

The contrast between Tables 3 and 4 represents the extent to which some men failed to meet criteria for success after 5 years. The higher percentages in Table 4 reflect the fact that a number of men in all categories seemed successful in treatment as much as 5 years following treatment termination, but at some point thereafter again met criteria for failure. This was especially true in the rapist category, in which, after any 5-year period, the failure rate was 15.6%; yet, overall, between 1973 and 1997, this rate actually reached 21.2%. The corresponding figures for homosexual pedophiles, the group with the next highest failure rate, were 15.8% and 16.3%, respectively.

Figure 1 depicts failure rates over time for each category of offender. As expected, these rates gradually rose following treatment termination, reaching relative asymptotes between 10 and 15 years. The relative rates of increase did not appear to differ among offender categories, except among homosexual pedophiles and rapists, in which failure (mainly reoffenses) occurred earlier. These data may

<table>
<thead>
<tr>
<th>Offender (N)</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child molesters, female victim (1,264)</td>
<td>53</td>
<td>4.2</td>
</tr>
<tr>
<td>Child molesters, male victim (485)</td>
<td>37</td>
<td>7.6</td>
</tr>
<tr>
<td>Heterosexual pedophiles (612)</td>
<td>57</td>
<td>9.3</td>
</tr>
<tr>
<td>Homosexual pedophiles (758)</td>
<td>120</td>
<td>15.8</td>
</tr>
<tr>
<td>Exhibitionists (957)</td>
<td>120</td>
<td>12.8</td>
</tr>
<tr>
<td>Rapists (403)</td>
<td>63</td>
<td>15.6</td>
</tr>
<tr>
<td>Totals (4,479)</td>
<td>450</td>
<td>10.1</td>
</tr>
</tbody>
</table>

TABLE 3
Failure Rate of Each Category of Offender Available at 5-Year Follow-up (1973-1992 cohorts only)
### TABLE 4
Failure Rates and Reasons for Failure for Each Category of Offender Available for Follow-up in 1997

<table>
<thead>
<tr>
<th>Offender (N)</th>
<th>Overall Failure Rate</th>
<th>Of Failures, Self-Reported Lapses and Relapses</th>
<th>PLETHYSMOGRAPH Failures</th>
<th>POLYGRAPH Failures</th>
<th>Recidivists</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>CM, FV (833)</td>
<td>53</td>
<td>6.4</td>
<td>21</td>
<td>39.6</td>
<td>9</td>
</tr>
<tr>
<td>CM, MV (279)</td>
<td>26</td>
<td>9.4</td>
<td>8</td>
<td>30.8</td>
<td>6</td>
</tr>
<tr>
<td>Het ped (302)</td>
<td>29</td>
<td>9.7</td>
<td>3</td>
<td>10.3</td>
<td>8</td>
</tr>
<tr>
<td>Homo ped (419)</td>
<td>68</td>
<td>16.3</td>
<td>7</td>
<td>10.3</td>
<td>7</td>
</tr>
<tr>
<td>Exhib (570)</td>
<td>77</td>
<td>13.5</td>
<td>11</td>
<td>14.3</td>
<td>6</td>
</tr>
<tr>
<td>Rapists (434)</td>
<td>92</td>
<td>21.2</td>
<td>3</td>
<td>3.3</td>
<td>8</td>
</tr>
<tr>
<td>Totals (2,837)</td>
<td>345</td>
<td>12.2</td>
<td>53</td>
<td>15.4</td>
<td>44</td>
</tr>
</tbody>
</table>

NOTE: CM, FV = child molesters, female victim; CM, MV = child molesters, male victim; Het ped = heterosexual pedophiles; Homo ped = homosexual pedophiles; Exhib = exhibitionists.
partly reflect the declining propensity for men to sexually aggress as they age (Knight & Prentky, 1990).

Table 4 also reports the reasons for failure in each category. These figures demonstrate that for nonpredatory offenders, reasons for treatment failure were not accounted for by actual legal charges. Instead, it appeared that lapses, such as deviant urges, or relapses, such as masturbation to deviant fantasies, constituted the majority of failures during follow-up. However, in the case of homosexual pedophiles and rapists, new charges accounted for the majority (75% and 84%, respectively) of all failures. In contrast, new charges for child molesters were a less frequent cause of treatment failure (just 13% for molesters of girls and 31% for molesters of boys).

In contrast to new charges, dropouts seemed to correlate negatively with severity of offense category. Whereas rapists and pedophiles had relatively low rates of terminating treatment prematurely (11% and 15%, respectively), child molesters had higher rates (between 24% and 33%) as did exhibitionists (31%). This might have been due to the nature of treatment mandated in each case: Those with more serious offenses were usually under stricter guidelines to participate in, and complete, a treatment program.
Plethysmograph and polygraph failures are reported in Table 4, although it must be remembered that not all participants available by telephone agreed to such tests and, even among those who did, tests were not completed consistently across time frames. Table 4 presents the number taking these tests at any point during follow-up. Unfortunately, data collection methods did not allow for an analysis of test results by length of follow-up period.

In contrast to past data analyses (Maletzky, 1993), dropouts were not counted automatically as treatment failures. Clinical experience indicates that there are multiple reasons for ending treatment prematurely aside from treatment resistance, including legal and financial problems. Nonetheless, leaving treatment early predicted treatment failure, particularly recidivism. Table 5 presents data that were collected on dropouts. It shows that even among men with situational offenses, recidivism was highly correlated with ending treatment prematurely. Among all participants, dropping out could predict with

<table>
<thead>
<tr>
<th>Offender (N)</th>
<th>Completed Treatment</th>
<th>Recidivated</th>
<th>Terminated Treatment</th>
<th>Recidivated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>CM, FV (2,196)</td>
<td>1,675</td>
<td>76.3</td>
<td>4</td>
<td>0.24</td>
</tr>
<tr>
<td>CM, MV (765)</td>
<td>514</td>
<td>67.2</td>
<td>3</td>
<td>0.58</td>
</tr>
<tr>
<td>Het ped (1,011)</td>
<td>851</td>
<td>84.2</td>
<td>4</td>
<td>0.47</td>
</tr>
<tr>
<td>Homo ped (1,251)</td>
<td>1,067</td>
<td>85.3</td>
<td>19</td>
<td>1.8</td>
</tr>
<tr>
<td>Exhib (1,604)</td>
<td>1,102</td>
<td>68.7</td>
<td>8</td>
<td>0.73</td>
</tr>
<tr>
<td>Rapists (448)</td>
<td>397</td>
<td>88.6</td>
<td>39</td>
<td>9.8</td>
</tr>
<tr>
<td>Totals (7,275)</td>
<td>5,606</td>
<td>72.1</td>
<td>77</td>
<td>1.4</td>
</tr>
</tbody>
</table>

NOTE: CM, FV = child molesters, female victim; CM, MV = child molesters, male victim; Het ped = heterosexual pedophiles; Homo ped = homosexual pedophiles; Exhib = exhibitionists.

a. Recidivism means being charged with a new sexual crime. These percentages represent the entire sample and are thus lower than those in Table 4, in which figures are given for only those participants able to be followed by telephone. Note, however, that total participants recidivating in each category were the same, as recidivism data were presumably available for all participants.
some accuracy that an offender would eventually be charged with a new sexual crime. Of 1,669 dropouts, 136 (8.1%) were eventually charged with a new sexual crime, compared to 69 of 5,607 (1.4%) of offenders who completed treatment. These data were compromised by the difficulty of obtaining follow-up in men who dropped out, the low N consequently generated, and the likelihood that men available for follow-up were preselected toward positive outcomes.

Table 6 presents data on the progress of treatment outcomes through the years of study. This table provides failure rates in 1997 for each type of offender by type of cohort. For example, molesters of girls had a failure rate of 7.6% for the 1972 to 1977 cohort; this rate increased to 7.9% for the 1978 to 1983 cohort, then declined to 6.5%, 4.9%, and then 4.3% for the next three 5-year cohorts. The data indicate that the techniques employed may have improved in efficacy over the 25-year span. The same can be noted for each type of offender through the years, although for homosexual pedophiles and rapists the rate still remained unacceptably high for the 1993 to 1997 cohort. Although the data imply that techniques are becoming more effective, they might also be explained by patient selection factors or altered patterns of referral.

**DISCUSSION**

The following discussion addresses the hypotheses generated in the introduction.

1. **Treatment was effective in reducing relapse and preventing recidivism over long periods of time.**

When assessed at 5-year intervals, and over a 25-year span, cognitive/behavioral treatment for most offenders appeared effective when provided in individual and group therapy, as measured by self-reports, criminal records reviews, and, when available, by plethysmograph and polygraph assessments. Although deficiencies in the present study design, enumerated later, preclude definitive statements, the criteria for success were relatively stringent. In addition, although many
## TABLE 6
Failure Rates in 1997 for Each Category of Offender as a Function of Time of Cohort

<table>
<thead>
<tr>
<th>Cohort Period (N)</th>
<th>CM, FV</th>
<th>CM, MV</th>
<th>Het Ped</th>
<th>Homo Ped</th>
<th>Exhib</th>
<th>Rapists</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>1973-1977 (1,137)</td>
<td>26 of 340</td>
<td>7.6</td>
<td>15 of 130</td>
<td>11.5</td>
<td>18 of 180</td>
<td>10.0</td>
</tr>
<tr>
<td>1978-1982 (1,742)</td>
<td>41 of 520</td>
<td>7.9</td>
<td>18 of 163</td>
<td>11.0</td>
<td>14 of 165</td>
<td>8.5</td>
</tr>
<tr>
<td>1983-1987 (1,801)</td>
<td>35 of 540</td>
<td>6.5</td>
<td>14 of 189</td>
<td>7.4</td>
<td>22 of 257</td>
<td>8.6</td>
</tr>
<tr>
<td>1988-1992 (980)</td>
<td>15 of 301</td>
<td>4.9</td>
<td>7 of 100</td>
<td>7.0</td>
<td>12 of 163</td>
<td>7.4</td>
</tr>
<tr>
<td>1993-1997 (1,615)</td>
<td>21 of 490</td>
<td>4.3</td>
<td>12 of 181</td>
<td>6.6</td>
<td>16 of 246</td>
<td>6.5</td>
</tr>
<tr>
<td>Totals (7,275)</td>
<td>138 of 2,191</td>
<td>6.3</td>
<td>66 of 829</td>
<td>8.0</td>
<td>82 of 1,011</td>
<td>8.1</td>
</tr>
</tbody>
</table>

NOTE: CM, FV = child molesters, female victim; CM, MV = child molesters, male victim; Het ped = heterosexual pedophiles; Homo ped = homosexual pedophiles; Exhib = exhibitionists. These figures do not exactly match those in Table 3 due to slight alterations in categorization rates of offense over time.
of these participants who terminated treatment prematurely had new sexual charges, between one half and three quarters, depending on category of offender, did not.

However, when a participant could not be followed at any point after treatment, he was considered neither a success nor a failure beyond that point. Thus, failure rates over time may reflect a bias toward success, in that participants making themselves available for follow-up might have been those most likely to succeed (Studer & Reddon, 1998). This is not as reasonable an assumption as it appears, however. Reasons for unavailability might include many not related to treatment outcome, such as a move for legitimate purposes (a new job), a desire for privacy, or a lack of financial resources.

If the reason for treatment is to prevent future sexual crimes, recidivism (new sexual charges) might be the most reasonable determinant of success (Hunter & Figuered, 1999; Rasmussen, 1999; Studer & Reddon, 1998). However, other criteria were included because the data indicate that offenders admitting to deviant sexual behavior (Maletzky, 1993) or failing plethysmograph or polygraph tests (Campbell, 1995; Lalumiere & Quinsey, 1994) are at higher risk to reoffend. Including multiple assessment techniques attempts to counterbalance positive outcome bias and can help identify those offenders most likely to be at risk.

Unfortunately, the percentage of participants undergoing each technique was variable. The numbers of offenders taking plethysmographs and polygraphs were considerably less than those available by telephone. Therefore, the plethysmograph and polygraph failures reported in Table 4 comprise a smaller percentage of overall failures and cannot be taken as representative because those men agreeing to such tests might have been preselected for positive outcomes. Nonetheless, it was reassuring to employ these instruments with some offenders, especially in the more dangerous categories. Data collection methods, however, prevented an analysis of which offenders failed more than one criterion. For example, we do not know if, among the 13 recidivist molesters of girls, any of the same men are represented in the 21 reporting lapses or relapses or in the 9 failing the plethysmograph test. Data within similar studies indicate some correlation and overlap among these variables (Dwyer, 1997; Hall, 1995).
2. Treatment results differed, based on category of offender.

Offenders categorized as predatory or preferential, including pedophiles and rapists, had 2 to 2 1/2 times the overall failure rate of child molesters. This matches the findings of a recent meta-analysis in which homosexual pedophiles and men who raped had the highest relapse rates (Alexander, 1999). Moreover, among pedophiles and rapists, sexual charges accounted for a disproportionate number of failures. Whereas recidivism was moderate for men who molested girls (24.5% of all failures) and boys (30.8%), it accounted for 55.2% of failures among heterosexual pedophiles, 75% among homosexual pedophiles, and 83.7% among rapists. At the same time, dropouts were greater among child molesters, perhaps because they were not as closely supervised or rigidly mandated to continue treatment.

Among exhibitionists, rates seemed intermediate, although very few failed a plethysmograph test, perhaps due to its lack of sensitivity in distinguishing this group (Maletzky, 2000). This is especially pertinent because exhibitionists failed polygraph tests in greater numbers and at higher percentages than almost any other group. In assessing men who expose, this instrument might be more sensitive than penile measures.

Self-reporting of covert or overt deviant behavior appeared to be inversely proportional to severity of crime. Among pedophiles and rapists, self-reports accounted for approximately 3% to 14% of all failures. Among child molesters, the corresponding rates were between 31% and 40%. It is possible that men convicted of more serious crimes tended to conceal deviant urges, fantasies, dreams, or behaviors, whereas child molesters might have had less fear of the consequences subsequent to revealing those activities.

It appeared that overall failure rates varied proportionately to severity of typical crimes along a continuum from situational child molesters to predatory pedophiles and rapists. The method of outcome assessment, however, may also be important in indicating which group achieved success, depending on definition of treatment goals (elimination of deviant arousal, reduction in recidivism), with drop-outs and self-reports more sensitive indicators among child molesters and the occurrence of new sexual charges more sensitive among pedophiles and rapists.
3. Failure rates increased with the duration of time after termination of treatment.

Although failure rates remained generally low in the first 6 to 12 months after treatment ended, they rose over time, doubling in 5 years but then remaining rather steady. Men who raped remained an exception. Their rates continued to rise until the 8th or 10th year; thereafter, they also remained stable. The slopes of ascent in Figure 1 parallel each other except for that one exception. If, as initially feared (Hanson et al., 1993; Miner, 1997), men continue to offend at increasing rates as time progresses, this trend remains hidden in the present data set, perhaps because of methodological limitations.

One such flaw of particular concern is that, as time advances, fewer offenders are available for follow-up, and those who are may be preselected for treatment success. However, availability for follow-up may be influenced by a variety of factors not linked to outcome. Moreover, the data included a legal records search, which would have identified men unavailable for other types of assessment at any time. Perhaps a larger factor influencing the decline in reoffense rates over time was age. The average age of offenders in this program from 1973 to 1975 was 37. The average age 10 years following treatment, when rates for most offending sharply decline (Hudson & Ward, 1997), was 47, and at the latest data collection point in 1997, 59.

4. Recidivism constituted one among a number of factors related to treatment failure.

Although a new sexual offense represents a clear treatment failure, other outcome criteria may also be important to consider when determining a program’s success rate. Many sexual offenses are not reported. Of those that are, the sexual component is sometimes altered or dropped. Moreover, because all offenders cannot be followed accurately over longer time periods and because legal records searches are sometimes incomplete, indicators other than recidivism can be used to determine treatment outcome. This is especially true for those factors shown in prior research to indicate a likelihood of repeat offending, including prematurely terminating treatment (Studer & Reddon,
1998), continuing to have lapses and relapses (Maletzky, 1993), and the presence of deviant sexual arousal (Lalumiere & Quinsey, 1994).

Among all offenders treated, a difference emerged in reasons for treatment failure, based on type of offense: Predatory and preferential offenders, such as homosexual pedophiles and rapists, failed due to reoffending at a higher rate than child molesters, with heterosexual pedophiles and exhibitionists in between. In contrast, dropouts were far more frequent among offenders with situational, less severe offenses. This trend may be explained by the importance of two factors: More severe offenders were at higher risk to reoffend and yet under greater supervisory pressure to attend and complete a treatment program. Indeed, it is not unusual in clinical experience for corrections officers to not require a child molester to return to treatment for a variety of reasons, including lack of funds or geographic relocation.

The importance of lapses and relapses and of plethysmograph and polygraph results is less clear from the data. Child molesters reported a greater amount of deviant sexual behavior than pedophiles, exhibitionists, and rapists but might have had fewer legal repercussions to fear in doing so. Although deviant arousal accounted for a greater percentage of failure among child molesters with male victims and in heterosexual pedophiles, numbers of participants involved in these tests, and in polygraphs, were too small to allow general conclusions.

Nonetheless, the data indicate that failure can be taken broadly to include not only actual reoffenses but also factors known to predict offenses in the future, thus preserving community safety as the highest priority of treatment programs.

5. Participants terminating treatment prematurely were at higher risk to reoffend.

Although fewer than one quarter of all participants dropped out of treatment, those who did represented a population at higher risk to reoffend. This risk again varied with type of offense. Men who preferentially molested boys, for example, were at a higher risk than child molesters, exhibitionists, or men who preferentially molested girls. However, a striking difference was noted in men who rape. Almost three quarters of such men terminating treatment prematurely raped again.
Fortunately, as stated above, most offenders did not drop out of treatment. However, 24% to 33% of child molesters and exhibitionists did. Clearly, methods need to be developed that will enable offenders to remain in treatment. This is unquestionably the responsibility of treatment providers and supervising officers alike. Moreover, for those who rape, better treatment methods are needed, even for those who do stay: Almost 10% of rapists completing a treatment program raped again.

6. Treatment techniques improved throughout the years of the study.

Over time, two trends may have combined to improve treatment success: New techniques have been developed, and therapists have become more skilled in applying them. Although the retrospective and amalgamated data presented here cannot disentangle individual treatment efficacies, they imply that one or both factors may have played a role in the declining failure rates. For most offender categories, such rates dropped by 35% to 40% between cohort years 1973 to 1977 and 1993 to 1997. During the late 1980s to mid-1990s, several powerful techniques were added to the repertoire of clinical approaches, including aversive behavior rehearsal (Maletzky & Steinhauser, 1998) and vicarious sensitization (Weinrott & Riggan, 1997). In addition, therapist turnover in the clinic declined, thus increasing the number of experienced therapists available to treat offenders. However, the amount that such factors contributed to improving success rates is inestimable, given the limitations of the present dataset.

For one group, however, failure rates remained high: men who raped. More than 20% of these men raped again following treatment in the 1993 to 1997 cohort, compared with 24% of those in the 1973 to 1977 cohort. These figures underscore the need to further refine treatment techniques for this difficult and dangerous group of offenders.

An alternative possibility for declining failure rates over time must be explored: It is possible that referral patterns changed over the years, such that fewer preferential and predatory offenders were referred during the later years of the study. A retrospective analysis of the categorization of offenders admitted to the clinic, however, does not support this view. The possibility that some other broad sociological change occurred in offender characteristics throughout the years
seems improbable. It appears more likely that techniques, and therapists employing them, have become increasingly sophisticated over time, although not to the extent needed to eliminate risk, especially among more dangerous offenders.

**STUDY LIMITATIONS**

These data were generated by retrospective review and, hence, were limited to the weaknesses inherent in searching past records, including loss of data, inadequate data, or missing records. Even an ongoing legal records search is prone to missing information if a sexual charge is altered to a nonsexual one or if an offender successfully uses an alias. Also, a control group of untreated offenders was not available, partly due to ethical concerns (Barbaree, 1997). Moreover, the data are uneven, in that follow-up was available for only a certain percentage of offenders, although that percentage remained fairly constant across offender categories. Hence, severity of offense should not measurably have skewed the results in a favorable direction. Nonetheless, those available for follow-up could have been preselected for positive outcomes as being more stable and cooperative, another factor weakening reliance on this type of analysis. In addition, data were not subjected to statistical review due to lack of individualized data collection for all participants and lack of adequate control groups.

Data were also uneven in that not all offenders were available for all types of follow-up. As anticipated, many more participants agreed to a telephone interview than a plethysmographic reevaluation, even when offered without charge. Even fewer agreed to a brief polygraph. Participants interviewed could also be suspected of being biased toward giving a positive self-report for deviant urges and behaviors. Nonetheless, a legal records search was theoretically available for all offenders. These difficulties underscore the importance of including a variety of assessment techniques in evaluating program efficacy.

Those participants not available for follow-up were not included in most of the data analyses. However, again, a legal records review was available for all participants. The fates of participants who could not be located remain obscure and cloud the data. It is reasonable to
assume, however, that the majority of such men did not reoffend, based on overall treatment results and legal records searches, although the more predatory and preferential offenders who did not complete treatment were at markedly greater risk. Determination of the exact risk remains unclear.

Outcome data also did not take into account treatment and recovery, or lack of it, for comorbid conditions that may have contributed to reoffending, such as alcoholism (Marx, Gross, & Adams, 1999), major depressive disorder (Raymond, Coleman, Ohlerking, Christianson, & Miner, 1999), or other concomitant Axis I and II disorders (McElroy et al., 1999).

**STUDY STRENGTHS**

Data were drawn from a typical, clinically based population of offenders, apparently similar to those throughout North America (Dwyer, 1997; Hall, 1995), and included follow-up on a large number of offenders over a longer time period than previously reported. In addition, therapeutic techniques used were composed of those most often employed in treating this population (Freeman-Longo et al., 1995). In addition, the criteria for success were stringent. Those offenders admitting to lapses or relapses were counted as failures, even without the occurrence of recidivism. Moreover, charges, not convictions, were considered adequate grounds for failure. Hence, internal biases were established against treatment success in an attempt to counter the weaknesses inherent in this retrospective review.

**CONCLUSIONS**

Within the limitations of this methodology, it appears that the treatment techniques employed in a cognitive/behavioral program generated long-lasting, positive results by reducing recidivism and risk to the community. Absolute statements with regard to treatment success cannot be made due to the limitations of the data collection methods. Outcomes appeared to be better in situational offenders, such as child
molesters and exhibitionists, than in predatory and preferential offenders, such as homosexual pedophiles and rapists. Failure rates, however, increased with the duration of time after treatment. Recidivism appeared to be the most sensitive indicator among the more dangerous offender categories, such as pedophiles and rapists, whereas self-reports and plethysmograph failures appeared to represent more sensitive measures in child molesters. Participants terminating treatment prematurely had higher failure rates than those who did not but, even among those dropping out, recidivism was uncommon except for rapists. From the cumulative data, it appeared therapy was becoming increasingly successful over the 25-year study span, perhaps a reflection of improved treatment techniques and increasing sophistication in their use.

RECOMMENDATIONS

Clearly, equally large prospective studies ensuring adequate follow-up are needed. One means to build in follow-up is to coordinate efforts with departments of corrections, especially now that mandatory reporting laws have been enacted (Walsh, 1997). Also, smaller studies employing institutionally based control groups and using the plethysmograph and polygraph as assessment techniques would provide additional outcome data in a safe fashion. In addition, smaller clinical studies might be combined to yield larger Ns. Offenders could also serve as their own controls if rates of individuals offending before and after treatment could be measured. In addition, multiple assessment techniques should be incorporated into study design to capture as much potential risk as possible. Recidivism rates alone may underestimate future risks.

The data point to the need to develop stronger techniques for homosexual pedophiles and men who rape. There is some indication that, even among these dangerous offenders, additional techniques, such as aversive behavior rehearsal and vicarious sensitization, may further lower rates of reoffense. In addition, because dropouts showed higher rates of recidivism, programs need to develop means of retaining offenders in treatment. This responsibility should be shared by treatment providers and supervising officials alike and might include a
program of incentives for successful completion of successive treatment phases. Finally, programs will need to continue working in cooperation with departments of corrections to ensure that offenders, particularly those perceived to have committed less severe offenses, remain in treatment. Even among child molesters, a second victim should be considered intolerable.

REFERENCES


Dr. Barry M. Maletzky is a clinical professor of psychiatry at the Oregon Health Sciences University in Portland and director of the Sexual Abuse Clinic in Portland. He has been involved in the treatment of sexual offenders for 30 years and has contributed more than 60 articles and chapters of books to the literature on this subject and other topics in general psychiatry. He is also the immediate past editor-in-chief of Sexual Abuse: A Journal of Research and Treatment, published by Kluwer/Plenum in New York and London.

Dr. Cynthia Steinhauser received her Ph.D. from the University of Chicago and is a licensed clinical social worker in Oregon. She has more than 20 years of experience in the fields of mental health and corrections and is the former director of the Sex Offender Treatment Program, Correctional Services of Canada, Pacific Region. Dr. Steinhauser specializes in the cognitive/behavioral treatment of sexual offenders. She has also spearheaded the development of techniques for the treatment of the disabled offender and has contributed many hours of volunteer effort to assist those with developmental disabilities. At present she is the associate director of the Sexual Abuse Clinic, Portland, Oregon.