CHAPTER 11

PEDOPHILIA
Assessment and Treatment

JOSEPH A. CAMILLERI
VERNON L. QUINSEY

The literature on the assessment and treatment of pedophilia is sufficiently large that it cannot be reviewed in a single chapter. More comprehensive and detailed reviews of assessment, treatment, and etiology issues can be found in Quinsey and Lalumiére (2001), the Safer Society series (www.saferociety.org), and Quinsey (2003), respectively. A more general overview is provided by Seto (2008). This chapter focuses on diagnostic and methodological issues specific to assessing sexual preferences for children, the modification of pedophilic preferences, and the outcome literature on the treatment of child molesters.

ASSESSMENT

Diagnosis

One of the earliest set of guidelines used by clinicians to diagnose pedophilia was provided by Richard von Krafft-Ebing (1886/1998) in his book Psychopathia Sexualis. He described “paedophilia erotica” as a psychopathology characterized by a primary sexual interest in children that manifests itself in sexual behavior directed toward them. This basic characterization has been retained in all subsequent clinical schemes for diagnosing pedophilia.

According to the Diagnostic and Statistical Manual of Mental Disorders, fourth edition, text revision (DSM-IV-TR; American Psychiatric Association, 2000) and the International Statistical Classification of Diseases and Related Health Problems, 10th revision (ICD-10; World Health Organization, 1994), a diagnosis of pedophilia requires a person to have recurrent, intense, and sexually arousing fantasies, urges, or behaviors directed toward a prepubescent child over a period of at least 6 months; to have acted on these
urges or to be distressed by them; and to be at least 16 years old and at least 5 years older than the child victim. The DSM-IV-TR also requires a clinician to specify whether the person is sexually attracted to males, females, or both; whether the acts or thoughts are limited to incest; and whether the person is exclusively or nonexclusively pedophilic (i.e., attracted only to children, or to children and adults). An individual who is in late adolescence and has an ongoing relationship with a 12- or 13-year-old is to be excluded.

O’Donohue, Regev, and Hagstrom (2000) have criticized the original (1994) DSM-IV approach to diagnosing pedophilia. Their criticisms are that, unlike the criteria for other disorders in DSM-IV, the criteria for pedophilia were not tested for interdiagnostician reliability; there are problems of temporal inconsistency in making the diagnosis; the procedures used to obtain internal consistency in DSM-III were problematic (no such data were collected for DSM-IV); and the validity of the DSM-IV criteria is compromised because they are vague, are arbitrary, and lack adequate operational definitions. It is therefore no surprise that these criteria are rarely used by researchers and clinicians (Marshall, 1997; O’Donohue et al., 2000). However, O’Donohue and colleagues’ solution of focusing strictly on behavior, so that a single sexual act with a child would be sufficient evidence for a diagnosis of “pedophilia response disorder,” is problematic: It removes the important distinction between individuals who sexually prefer children but have never molested a child, and those who have committed a sexual offense against a child but sexually prefer adults (Cohen & Galyonker, 2002).

Not only do the DSM and ICD fail to distinguish between behavior and preference; the attendant confusion between pedophile and child molester categories is found throughout the literature, as noted by Barbaree and Seto (1997). Because pedophilic behavior is used as a sufficient diagnostic criterion, cause and effect are obfuscated. Pedophilia is best understood as a sexual preference for children that may or may not lead to child molestation. Child molestation, on the other hand, involves a sexual offense against a child that may or may not be due to pedophilia.

Pedophiles, child molesters, and pedophilic child molesters should be understood as separate groups, because a diagnosis of pedophilia is a strong predictor of sexual recidivism (Hanson & Bussière, 1998). Using such groups in research and clinical practice allows us to streamline treatments toward targeting those at high risk (see the “Treatment Efficacy and Research Priorities” section of this chapter); to suggest new studies for understanding the link between pedophilia and child molestation by using pedophiles who have not molested children as a control group; and to assess treatment outcomes properly.

The DSM approach to diagnosing pedophilia has also been criticized on the grounds that the victim age criterion is arbitrary (O’Donohue et al., 2000). A solution to this problem comes from a simple yet practical method to understanding psychological disorders: the notion of “harmful dysfunction” (Wakefield, 1992). According to Wakefield, in order for a phenomenon to be classified as a disorder, it must meet two criteria: (1) The mechanism is not functioning in the way it was naturally designed to function; and (2) due to this dysfunction, the person is harmed or is deprived of benefit. This unifying evolutionary approach subsumes the use of biologically relevant diagnostic criteria, such as a preference for sexually immature features, instead of using an arbitrarily selected age (Seto, 1999, 2002). In addition to improvements in diagnostic criteria, approaching sexual behavior as reproductive behavior may help us understand the etiology of pedophilia by focusing on how pedophilic behavior is similar to and different from typical reproductive behaviors (Quinsey, 2002; Quinsey & Lalumière, 1995). A detailed discussion on
whether pedophilia is a mental disorder can be found in the *Archives of Sexual Behavior*
December 2002 issue.

**Comorbidity**

Pedophiles and child molesters exhibit psychiatric comorbidity for mood disorders, substance abuse, and impulse control disorders (Galli et al., 1999; McElroy et al., 1999; Raymond, Coleman, Ohlerking, Christenson, & Miner, 1999); other paraphilias (Abel, Becker, Cunningham-Rathner, Mittelman, & Rouleau, 1988; Galli et al., 1999; Raymond et al., 1999); antisociality and psychopathy (Dorr, 1998; Raymond et al., 1999; Seto, Harris, Rice, & Barbaree, 2004; Virkkuenen, 1976); and assertiveness deficits (as described in Cohen & Galynker, 2002). Such comorbidity is more likely in adjudicated or hospitalized child molesters than in child molesters assessed from community samples (Quinsey, 1986). A substantial proportion of pedophiles have been found to engage at least once in other types of sexual offenses, such as voyeurism, exhibitionism, and public masturbation (Abel et al., 1988). One of the most important possible concurrent conditions is that of psychopathy. For reasons of public safety, practitioners should assess pedophiles for psychopathy because of its dangerous relationship with sexual deviance: It has been demonstrated that both adult and adolescent psychopathic sex offenders are at a very high risk of both sexual and violent recidivism (e.g., Gretton, McBride, Hare, O'Shaughnessy, & Kumka, 2001; Harris et al., 2003). In addition, actuarial assessments to identify men who are at high risk of sexual and violent recidivism should be routinely used (see Quinsey, Harris, Rice, & Cormier, 2006)—not only to determine the appropriate amount of supervision, but also because treatments appear to be less effective with low-risk men (see “Treatment Efficacy and Research Priorities”; see also Andrews et al., 1990).

**Assessment Tools**

Generally, the goal of an assessment for pedophilia is to identify the presence or absence of a sexual preference for children. In most cases, these assessments are conducted on men who were either charged or convicted of child molestation. Some referrals, however, come from men who report pedophilic urges and are distressed by them. This section provides a synopsis of ways to assess men in either situation.

**Penile Plethysmography**

One of the most commonly investigated procedures of assessing pedophilic preferences is the measurement via penile plethysmography (PPG) of changes in penile circumference or volume occasioned by images of persons who vary in age and sex, or audiotaped stories concerning sexual interactions with persons who vary in age and sex. An increase in either penile circumference or volume is assumed to indicate sexual arousal, thereby indicating sexual desire (Kalmus & Beech, 2005; Rempel & Serafini, 1995). Such assumptions have been validated. For example, Harris, Rice, Quinsey, and Chaplin (1996) found a correspondence among phallometric assessment, viewing time (VT), and self-reported attraction to stimulus persons varying in sex and age among nonoffenders; however, and more importantly, they obtained better discrimination between persons who did and did not molest children with phallometric assessment than with the other measures. It has
also been demonstrated that PPG can be used to assess pedophilia among adolescent child molesters (Robinson, Rouleau, & Madrigano, 1997; Seto, Lalumière, & Blanchard, 2000). As we discuss in the “Effective Assessment” section, PPG is the best-validated tool for assessing pedophilia.

Phallometric assessment, despite its demonstrable utility in the assessment of deviant sexual preferences, is not without limitations, and these have inhibited its more widespread use. First, the procedure requires psychophysiological equipment (even though this is modest in amount and expense), as well as training of the assessors in gathering and interpreting phallometric data. The procedure can also be viewed as intrusive by clients and lawyers. Lastly, for a variety of reasons, the field has not converged upon a standard procedure with published norms. These limitations have led to various attempts to develop other tools to assess sexual age preferences.

Viewing Time

VT has been an attractive alternative to measuring penile tumescence because it is easier to administer. A number of studies have assessed the validity of VT by comparing it to PPG and by testing how well it can identify offender types. Variability in results has been related to offenders’ preference for a particular sex of child victim. Harris and colleagues (1996) found similar patterns between PPG and VT among child molesters who offended against girls. Using VT, Abel, Huffman, Warberg, and Holland (1998) correctly classified 66% of men with an interest in girls and 91% of men with an interest in boys. Letourneau (2002) aimed to replicate Abel’s findings and found that both VT and PPG were able to identify offenders against boys, but not offenders against girls. Most recently, Abel and colleagues (2004) significantly discriminated adolescent child molesters from adolescents who presented with other sexual problems with a VT measure.

These mixed findings, though encouraging, are not conclusive enough to warrant the use of VT in clinical practice. We recommend that future studies on VT should incorporate standardized procedures for data manipulation (e.g., z-scores and use of a deviance differential). It should be noted that VT is often used in combination with self-report measures, as, for example, in the popular Abel Assessment for Sexual Interest (AASI; Abel, Jordan, Hand, Holland, & Phipps, 2001). Some researchers have criticized the AASI (Fischer & Smith, 1999; Smith & Fischer, 1999) by outlining concerns about its development, reliability, and validity (see Abel, 2000, for a response).

Behavior

A diagnosis of pedophilia can be inferred from past behavior. For example, Seto and Lalumière (2001) developed the Screening Scale for Pedophilic Interests (SSPI), a four-item scale that assesses characteristics of past child molestation. These characteristics include having at least one male victim, committing offenses against multiple victims, having at least one prepubescent victim, and having an unrelated victim. The SSPI was found to correlate with phallometrically measured pedophilia in both adults and adolescents, and it was also found to predict sexual recidivism among adult sex offenders (Seto et al., 2004; Seto, Murphy, Page, & Ennis, 2003). Seto, Cantor, and Blanchard (2006) provided additional validity for the use of behavior to diagnose pedophilia: They found that men who were charged for possession of child pornography were more likely to have phallometrically measured pedophilic sexual interests. These results suggest that the SSPI is a useful tool to assess pedophilia.

Implicit Ass

Cognitive tests are a useful way to measure the implicit feelings and beliefs of a person who does not actively admit to pedophilic interests. One such test is the Implicit Association Test (IAT, Greenwald, McGhee, & Schwartz, 1998). The IAT measures the strength of associations between words related to pedophilia and words related to children. In the IAT, participants are shown pairs of words and are asked to press one of two buttons as quickly as possible. The speed at which participants respond to each word-pair is recorded, and the results are used to infer the strength of the association between the words. The higher the score, the stronger the association.

Scales and Tests

Scales and tests are important tools for forensic psychologists and researchers in the field of pedophilia. They allow for the assessment of individuals who may or may not be pedophiles, and can be used to guide treatment decisions. For example, self-report measures such as the AASI and SSPI can be used to assess the presence of pedophilic interests and the severity of the behavior.

Neuroimaging

Neuroimaging studies have focused on the brain structures and functions associated with pedophilic behavior. For example, studies have found differences in brain activity when people with pedophilic interests are shown images of children compared to when they are shown images of adults. These findings suggest that pedophilic behavior is associated with changes in brain function and structure.

If such findings are confirmed, they could have important implications for the treatment and prevention of pedophilic behavior. For example, targeted interventions could be developed to address the underlying brain changes associated with pedophilia. Additionally, these findings could inform policies and interventions aimed at preventing child sexual abuse.

Behavioral measures have also been used to assess pedophilic behavior. These measures include self-report measures such as the AASI and SSPI, as well as behavioral measures such as the IAT. These measures can be used to assess the presence of pedophilic interests and the severity of the behavior.

Conclusion

Pedophilia is a complex and challenging behavior to assess. However, a number of tools and methods have been developed to help assess and understand pedophilic behavior. These tools and methods include phallometric assessment, self-report measures, and behavioral measures. The development and validation of these tools and methods is ongoing, and further research and development is needed to improve our understanding of pedophilia and to develop effective treatments for this behavior.
among adolescent girls, & Blanchard, best-validated tool of deviant behavior in gathering data by interviewing upon a standard protocol.

because it is easier to use in results has been arris and colleagues to correctly classify in interest in boys. Both VT and PPG tests girls. Most recent child molesters VT measure. VT should incorporate use of a deviation with self-report sexual Interest (AASI) have criticized these concerns about its development.

for example, Sero and others (SSPI), a unique characteristic in multiple victims, have been used. The SSPI was found to be useful and offenders (Seto et al., 2006) provided. They found that men more likely to have high scores. Suggest that the SSPI is a useful alternative for researchers or clinicians who do not have more sophisticated tools to assess pedophilia.

**Implicit Association Test**

Cognitive tests have recently been used to assess pedophilia. The popular Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) is based on the assumption that a person who holds a favorable view of a topic (e.g., sex with a child) will associate words related to this topic (e.g., “sex” words and “child” words) faster than a person who does not hold such a favorable view, and faster than he will associate words unrelated to this topic (e.g., “sex” words and “adult” words). The potential benefit of this approach is that it can assess the presence of a preference that a person may not want to disclose. Two research groups have used the IAT paradigm to assess sexual preference for children, and found that child molesters provided more positive associations between children and sex than other offenders or nonoffender controls did (Gray, Brown, MacCulloch, Smith, & Snowden, 2003; Mihailides, Devilly, & Ward, 2004). With further validation, such as identifying the correspondence between IAT scores and phallicometric responding or sexual recidivism, this approach may be a very useful alternative for assessing pedophilia.

**Scales and Card Sorts**

Scales and card sort methods have been developed to identify sexual preference in both forensic and nonforensic populations. Assessing pedophilia in this way is useful for researchers and clinicians who do not have access to a phallicometric laboratory. Although such methods should not replace PPG, there is some indication that combining PPG and self-report procedures improves discrimination (Laws, Hanson, Osborn, & Greenbaum, 2000). A list of self-report tools and scales to assess pedophilia, including their psychometric properties, can be found in Quinsey and Lalumière (2001).

**Neuroimaging**

Neuroimaging and neurophysiological methods have been used to compare child molesters with nonforensic controls on measures of neural activity and composition. Differences between these groups have been demonstrated with various technologies, including computerized axial tomography (Hucker et al., 1986; Wright, Nobrega, Langevin, & Woztman, 1990), electroencephalography (Flor-Henry, Lang, Koles, & Frenzel, 1991), and recently magnetic resonance imaging (Cantor et al., 2006). Some studies focusing on pedophiles indicate anomalies in the temporal lobe (Cohen & Galynker, 2002). Bassarath (2001) reviewed the literature on neuroimaging studies of antisocial behavior and found converging evidence for atypical circuitry in the prefrontal cortex. If such findings indicate pathological differences instead of simply differences, they are consistent with pedophilia defined as harmful dysfunction: Atypical circuitry provides some evidence for a dysfunctional psychological mechanism. Neuropsychological measures have the potential for being quite useful, but at this time, due to limited access to equipment and little research on treating anomalies identified from neuroimages, this technology is still in the developmental stage. It should be noted that the neural structural correlates of homosexual preference in men are too small to be identified with
current methods of brain imagery used with living participants, the same is likely to be true with pedophilia.

**Other Physiological Measures**

Other methods used to assess pedophilia include measuring penile temperature, penile surface blood volume, skin conductance responses, and pupillometry. Their ability to discriminate pedophiles from nonpedophiles is poorer than that of PPG, however. Because there is an abundance of research supporting both the discriminative and predictive validity of PPG, and because very little research using these alternative methods have been employed in the last 20 years, PPG remains the most trusted assessment tool (for a detailed review, see Kalmus & Beech, 2005). The following section describes procedures that can be used to maximize PPG capabilities.

**Effective Assessment**

If we adhere to a definition of pedophilia as a “sexual preference for children,” then, after surveying the different methods of assessment, we must conclude that phallometric assessment has emerged as the most reliable and valid procedure to assess this preference. Over the past 30 years, research has addressed concerns over equipment, sensitivity and specificity, and optimal procedures and statistics to optimize discrimination.

**Volumetric versus Strain Gauge**

Two types of equipment are used to assess sexual preferences phallometrically: an air chamber that measures change in penile volume, and a mercury-in-rubber strain gauge that measures change in penile circumference. An important question is this: Which type of phallometric equipment can better discriminate between pedophiles and nonpedophiles? There has been much discussion in the literature on the equivalence of these apparatuses. In a well-controlled study, where both tools were used simultaneously on a nonforensic sample, Kuban, Barbaree, and Blanchard (1999) found a strong positive correlation between the two methods (r = .80) on standardized scores across stimulus categories when they included participants with responses that were 10% greater than a full erection. This relationship disappeared among low responders (mean r = .15). Other researchers found that when they employed deviance differential scores (i.e., difference between highest response to deviant and nondeviant stimuli) and a circumferential device, discriminant validity was not improved by excluding low responders (Harris, Rice, Quinsey, Chaplin, & Earls, 1992). The important questions of which technique more accurately differentiates child molesters from normal individuals, and which more closely correlates with history of victim choice, have yet to be addressed directly.

**Sensitivity and Specificity**

Three studies have assessed phallometric sensitivity and specificity for diagnosing pedophilia (Blanchard, Klassen, Dickey, Kuban, & Blak, 2001; Freund & Blanchard, 1989; Freund & Watson, 1991). “Sensitivity” is calculated by dividing the number of men identified as pedophiles by the phallometric assessment out of the total number of pedophiles in the sample. For child molesters who had multiple child victims, sensi-
tivity ranged from 61% to 88.6%. Better sensitivity was found among offenders with male victims. "Specificity" is calculated by dividing the number of men identified as gynephiles (i.e., men who prefer adult women) by the phallicometric assessment out of the total number of true gynephiles in the sample. Research has shown specificity to range from 80% to 96.9%. Similar specificity was found among adolescents who sexually offended against children (92%), while sensitivity was lower but still acceptable at 42% (Sato et al., 2000). Of course, these values only apply to the specific stimuli and methods used in these particular demonstrations. In the absence of standardized procedures, clinicians need to determine the psychometric properties of those that they use.

Because sensitivity and specificity are influenced by base rates, we recommend the use of the "receiver operating characteristic" (ROC) statistic, or the "common language effect size" (CLES). Not only are ROC and CLES indices robust to variability in base rates of pedophilia in samples,(15,179),(976,993)

Procedures and Statistics

Lalumière and Harris (1998) reviewed the literature to address some common questions on best practices for optimal discrimination using phallicometric testing. Some of their recommendations are as follows:

- Use of images best discriminates age and gender preference (e.g., Harris et al., 1996).
- Use of graphic and violent narratives best discriminates preference for coercive sex.
- Use of more than one stimulus per category is optimal. Validity has been established with two to five stimuli per category (e.g., Lalumière & Quinsey, 1994).
- The same duration of stimulus presentation should be used as was used in the validation of the stimulus set (e.g., Quinsey, Chaplin, & Varney, 1981).
- Computation of a "deviance differential"—an index of relative preference between deviant (child or coercive sex) and nondeviant (adult or consenting sex) stimuli—is recommended. The deviance differential is calculated by subtracting the largest mean response of the deviant category from the largest mean response of the nondeviant category (e.g., Harris et al., 1992).
- Use of $z$-score transformations addresses individual differences in responding (i.e., high vs. low responders), thus improving discriminant validity.
- When auditory stimuli are used, faking can be reduced by including a semantic tracking task to ensure that attention is being paid to the stimuli. For example, subjects can be required to press one button when a description of sex is presented.
and another when violence is presented (Proulx, Côté, & Achille, 1993; Quinsey & Chaplin, 1988).

**Pedophilia Typologies**

Much of the research on pedophilia subcategories has focused on differences in overt behavior and not on differences in sexual preferences per se. We can still draw from this literature to understand possible categories of pedophilic sexual interest. The simplest typologies that account for much of the variability in pedophilia are those of victim sex (i.e., heterosexual–homosexual) and relationship to victim (i.e., intrafamilial–extrafamilial). In order to be useful, other typologies must be able to outperform these distinctions in terms of dealing with treatment issues and predicting recidivism (Quinsey, 1986). Though other typologies have been proposed, such as the one developed by Knight, Carter, and Prentky (1989), none have yet met these criteria.

**Heterosexual–Homosexual Typology**

A common and important distinction among pedophiles is not just their age preference, but their sexual orientation or gender preference. Although most sex offenses are directed toward females (approximately 67%), a substantial proportion are committed against same-sex victims (estimates range from 12% to 33%) and victims of both sexes (20%; Abel et al., 1988). There is a generally good relationship between a victim’s sex and the sex preference found phallometrically (Freund, 1967; Harris et al., 1996). Identifying homosexual pedophiles is important, because homosexual pedophiles are about twice as likely to recidivate as heterosexual child molesters (Quinsey, 1986).

**Intrafamilial–Extrafamilial Typology**

There is some evidence to support the traditional view that exclusively intrafamilial (i.e., incest) and extrafamilial child molesters are different groups. For example, although victims of incest offending tend to experience a greater frequency of assaults and a greater likelihood of completed intercourse than victims of nonincest offending (Gebhard, Gagnon, Pomeroy, & Christenson, 1965), intrafamilial child molesters are not as sexually deviant, are at lower risk for sexual recidivism, are less psychopathic, are less likely to have intercourse with the victims, are less likely to have male victims, cause less injury, and have lower sexual and violent recidivism rates (Rice & Harris, 2002).

In terms of sexual preference, a clear difference between intrafamilial and extrafamilial sex offenders on phallometric measures of preference has not emerged, even though at least 12 studies have addressed this issue (Abel, Becker, Murphy, & Flanagan, 1981; Barsetti, Earls, Lalumiè re, & Belanger, 1998; Blanchard et al., 2006; Frenzel & Lang, 1989; Freund, Watson, & Dickey, 1991; Lang, Black, Frenzel, & Checkley, 1988; Langevin & Watson, 1991; Malcolm, Andrews, & Quinsey, 1993; Marshall, Barbaree, & Christophe, 1986; Murphy, Haynes, Stalgaitis, & Flanagan, 1986; Rice & Harris, 2002; Seto, Lalumiè re, & Kuban, 1999). A consistent finding across these studies is that incest offenders were never more deviant than extrafamilial child molesters, and at least one study found that both intrafamilial and extrafamilial child molesters showed a stronger sexual preference for children than rapists and nonforensic controls did (Seto et al., 1999). In order for these results to be interpretable, however, offenders must be matched...
on differences in overt can still draw from this interest. The simplest are those of victim sex afamilial–extrafamilial). rm these distinctions in Quinsey, 1986). Though by Knight, Carter, and ust their age preference, sex offenses are directed s of both sexes (20%; in a victim’s sex and the l., 1996). Identifying ho- nies are about twice as 86). usively intrafamilial (i.e., or example, although vic- of assaults and a greater est offending (Gebhard, esters are not as sexually opathic, are less likely to victims, cause less injury, arris, 2002). t intrafamilial and extra- e has not emerged, even- ter, Murphy, & Flanagan, d et al., 2006; Frenzel & renzel, & Checkley, 1988; 93; Marshall, Barbaree, & 86; Rice & Harris, 2002; these studies is that incest esters, and at least one esters showed a stronger controls did (Seto et al., offenders must be matched on their own age at time of offense, and the age and sex of their victims (see Quinsey, Chaplin, & Carrigan, 1979).

Blanchard and colleagues (2006) found that paternal–nonpaternal groups are better discriminated by PPG scores than intrafamilial–extrafamilial groups. In their study, nonmolesters had the lowest sexual preference for children; nonpaternal molesters (i.e., nonpaternal incest offenders and molesters against unrelated children) had the highest sexual preference for children; and paternal molesters (i.e., biological fathers and stepfather) were in between. In other words, offenders against extended family members are similar to extrafamilial offenders, and the belief that incest offenders are less likely to be pedophilic appears to be true only for fathers and stepfathers. Implications of these mixed findings suggest that a simple categorization based on relationship to the victim has limited clinical utility (Bickley & Beech, 2001). Seto and colleagues (1999) recommended that incest offenders should not be excluded from phallometric assessments because of the strong relationship between a sexual preference for children and sexual recidivism.

Massachusetts Treatment Center: Child Molester Typology, Version 3

Some researchers have constructed child molester typologies that are guided by traditional psychological theories. A popular typology for child molesters has been advanced by Knight and colleagues (1989). Their typology, the Massachusetts Treatment Center: Child Molester Typology, Version 3 (MTC:CMT3), consists of two axes. Axis I includes fixation (i.e., degree of pedophilic interest—low or high) and social competence (low or high); Axis II includes amount of contact (i.e., actual physical contact with children—low or high), meaning of high contact (i.e., either interpersonal or narcissistic), level of physical injury for low contact, and whether injuries were sadistic or nonSadistic. Child molesters can reliably be assigned to these groups (Knight et al., 1989; Looman, Gauthier, & Boer, 2001), but further research on the MTC:CMT3’s predictive validity and treatment relevance is still required.

Assessment Research Priorities

Our evaluation of the research on assessing pedophilia has led us to several methodological and theoretical suggestions. With regard to methodology, there appears to be some evidence that using multiple assessment methods improves the accuracy of identifying pedophilia. Laws and colleagues (2000) found that combining self-report (i.e., card sort), PPG audio, and PPG slides had better diagnostic accuracy for pedophilia than any method alone. Additional research should verify these findings, and clinicians should, if feasible, use such multiple methods to assess a sexual preference for children. When it comes to actuarial assessments of risk, however, it appears as though combining tests is no better than using the best test (Seto, 2005).

Although researchers have identified numerous ways to assign pedophiles into subgroups, we think that more research separating pedophiles, child molesters, and pedophilic child molesters will greatly improve our understanding of the relationship between pedophilic interests and overt behavior. We would also like to see greater use of standardized procedures and stimuli, especially those outlined by Lalumière and Harris (1998). Such standardization not only optimizes discrimination of phallometric assessments, but allows for higher-quality meta-analytic research.
Lastly, we need to find methodologies to assess pedophilia as a dysfunctional mechanism. For example, Quinsey and Lalumière (1995) suggested that pedophilia results from malfunctioning body shape detectors. In the only test of this hypothesis, Rice and Harris (2003b) found that child molesters responded more to visual images of older women than of younger women—a finding that was predicted on the grounds that older women have a waist–hip ratio more like that of prepubertal children than that of sexually mature young women. Additional research using not just VT, but eye-tracking methodology would help us understand which physical aspects of sexual stimuli pedophiles and nonpedophiles attend to. Results would lend support to understanding the etiology of pedophilia, which may be a necessary step in developing treatment programs that properly address pedophilia. We should take cues from research on homosexuality (an area moving much faster on the etiological front), such as using the fraternal birth order effect at the neurohormonal and neural structure level for homosexual pedophilia, and conducting genetic studies of the heritability of pedophilia (see Quinsey, 2003, for a discussion of these issues).

TREATMENT

Treatments for pedophilia emerged as a means to prevent child molestation or as a way to reduce sexual recidivism. That is, pedophilia is treated either when a person wants to remedy his pedophilic tendencies, or when a person has committed a sexual offense against a child and his pedophilic interests are addressed to reduce sexual recidivism. This latter approach combines preference-based treatment with programs designed to manage other criminogenic needs. Both approaches are discussed below.

Before we begin our review of the treatment literature, it may be helpful to describe the optimum state of affairs. In the best of all possible worlds, treatment would change factors known to be related to reoffending. It would be best if these factors were etiologically relevant, although this is not necessary. Improvement in these factors would correlate highly and negatively with recidivism. Progress in treatment technology would be reflected in greater reductions in recidivism over time. Treatment outcome studies using random assignment would show large effect sizes, and better-implemented programs would be more successful. Unfortunately, as we document below, only one of these desired criteria has been met.

Treatment Based on Sexual Preference

Conditioning

We have established that “pedophilia” is a term that describes a sexual preference for prepubescent children, and that PPG is the most effective method for assessing such a preference. PPG is also useful for tracking changes in response patterns across interventions. Several studies have used PPG with treatments that are guided by the principles of conditioning in order to reduce sexual responding to children and to increase sexual responding to adults. These treatments are based on pairing an aversive stimulus with stimuli depicting an individual’s sexual preference. Variations on this principle include classical aversive conditioning, classical discriminative conditioning, covert sensitization, satiation, and signaled punishment.
In classical aversive conditioning, an unconditional stimulus (UCS), such as shock, is preceded by a conditional stimulus (CS+; a description of a deviant fantasy or image of a child) in order to reduce interest in the deviant stimulus (Marshall, 1973), whereas classical discriminative conditioning includes a conditional stimulus (CS--; a description of consenting sex with an adult or an image of an adult) that signals the absence of the UCS (Quinsey, Berghersen, & Steinman, 1976). Covert sensitization requires individuals to imagine a situation in which they approach a child but feel nauseous (sometimes they are asked to think of aversive items such as vomit or feces; Barbaree & Seto, 1997) and feel relief when leaving that situation (Barlow, Leitenberg, & Agras, 1969). Satiation can either be verbal or masturbatory; the latter involves a person masturbating to ejaculation while thinking about appropriate sexual acts, then continuing to masturbate after ejaculation while thinking about his inappropriate preference (Marquis, 1970). Signaled punishment involves the use of an indicator such as a red light that goes on when the individual is over a preset penile tumescence criterion in the presence of a deviant stimulus; the red light is associated with a probability per unit time of electric shock to the arm as long as it is on. A green light goes on when the person is over the criterion in the presence of an adult stimulus. An alternative nonconditioning procedure, biofeedback, can be used to determine whether aversive conditioning is actually required for an individual to gain control of his sexual responses in the laboratory. In this procedure, the individual is instructed not to respond in the presence of a deviant stimulus and to respond sexually in the presence of an appropriate stimulus. Penile tumescence changes are signaled to the individual with red and green lights, as in the signaled punishment procedure (Quinsey, Chaplin, & Carrigan, 1980). Several studies have used these methodologies to demonstrate reductions in phallometrically measured arousal to children (e.g., Quinsey et al., 1976, 1980) and in other sexual paraphilias (Abel, Levis, & Clancy, 1970).

More generally, other studies have shown that sexual arousal can be manipulated via conditioning paradigms not just in humans (O’Donohue & Plaud, 1994), but in nonhuman animals as well (Pfaus, Kippin, & Centeno, 2001). Some conditioning procedures appear to be more effective than others. Barbaree and Seto (1997) reported limited evidence for thematic shift, where deviant thoughts while masturbating change to appropriate thoughts during orgasm; limited support for fantasy alternation, where contents of sexual thoughts are changed daily or weekly; and some support for satiation and directed masturbation (i.e., masturbating exclusively to appropriate stimuli). Early conditioning procedures for treating paraphilias used mild electric shocks as the aversive stimuli (e.g., Quinsey et al., 1980). Use of shock has waned in favor of olfactory, visual, and cognitively aversive stimuli.

The greatest problems with conditioning approaches (and other approaches described later) are (1) that changing actual preferences (as opposed to indices that reflect them) has been quite difficult (Furby, Weinrott, & Blackshaw, 1989; Lalumière & Quinsey, 1998); and (2) no studies have shown long-term changes in sexual preference or behavior after treatment (Laws, 2001; Quinsey & Earls, 1990). From a conditioning view, this latter outcome might be explained by extinction: Removal of the contingencies results in reduced frequency of the conditional response. A problematic outcome is the expectation produced by a “honeymoon effect,” where patients believe that their deviant preferences are being fixed because of the quick effects of treatment and then fail to acknowledge the possibility of relapsing. Thus practitioners have been advised to inform patients that conditioning procedures do not make permanent changes (Laws, 2001). The difficulty in changing preference and the lack of permanent changes suggest that con-
conditioning procedures do not work by conditioning mechanisms. A plausible explanation for how conditioning works is by training the person to identify and control his responses in order to ease the suppression of sexual responding (also known as “cognitive mediation”; for further discussions of this topic, see Barbaree, Bogaert, & Seto, 1995; Quinsey & Earls, 1990). Training people in this way is easy with signaled punishment, for instance, but the effects of this procedure do not persist or reduce recidivism. The issue of why conditioning and other therapies have had limited success in changing arousal patterns is discussed in the “Treatment Efficacy and Research Priorities” section.

Treatments Designed to Reduce Sexual Recidivism among Child Molesters

As mentioned earlier, many treatment programs have been designed for convicted child molesters in order to address their possible pedophilic tendencies and other criminogenic needs. The ultimate goal of these programs is to reduce sexual recidivism by focusing on different causal factors thought to be related to child molestation. Humanistic and psychodynamic mechanisms are addressed in general psychotherapy; sexual preference, cognitive distortions, and procriminal attitudes are targeted in cognitive-behavioral therapy (CBT); maintaining therapeutic change is the goal of relapse prevention (RP); personal and environmental risk and protective factors are addressed in multisystemic therapy (MST); and inhibiting general sexual arousal is the aim of medical treatments.

General Psychotherapy

General psychotherapies are programs derived from either humanistic or psychodynamic traditions (some of these are atheoretical or eclectic). These programs are not well documented, appear to be unstructured, and have only a few well-controlled outcome evaluations (Lalumière, Harris, Quinsey, & Rice, 2005). Of these studies, none have demonstrated reductions in recidivism, and some evidence even suggests a possible increase in recidivism (e.g., Furby et al., 1989; Romero & Williams, 1983). For all these reasons, we do not describe these programs in any detail.

Cognitive-Behavioral Therapy and Relapse Prevention

CBT and RP are two of the most popular psychotherapeutic treatment approaches used with child molesters. CBT combines behavioral interventions (such as the conditioning techniques described earlier) with cognitive therapies. The behavioral portion focuses on sexual preference, whereas the cognitive portion focuses on the cognitive distortions, attitudes, and thinking errors related to sexual aggression, which are assumed to be the products of social learning (Murphy, 1990). As expected, child molesters appear to have more cognitive distortions associated with child molesting than men who sexually assault adult women have (e.g., Abel et al., 1989; Blumenthal, Gudjonsson, & Burns, 1999). Standard cognitive therapy works to change these aberrant thoughts by (1) teaching offenders how cognitions influence sexually aggressive behaviors, (2) informing them how offending injures victims, (3) training them how to identify their own cognitive distortions, and (4) using various pedagogical tools to help these men understand and work through their distortions (Murphy, 1990).
RP is a separate program module focusing on factors that maintain changes brought by treatment (Rice, Harris, Quinsey, & Cyr, 1990). RP targets three factors in particular: self-efficacy (confidence in coping with high-risk situations), coping skills (skills to cope with high-risk situations), and motivation (increasing desire not to relapse) (Bauml & Quinsey, 1992). Both CBT and RP utilize group and individual modalities, are implemented in inpatient and outpatient settings, and can last anywhere from 4 months to 30 months, with some programs providing booster sessions upon release (Marshall & Barbaree, 1990). Many programs include other modules, such as sex education, anger management, general counseling, empathy training, and social/heterosexual skills training (Rice, Harris, & Quinsey, 2001).

Treatment success has been assessed in terms of both reducing recidivism and achieving more short-term treatment goals. Though rigorous process evaluations of CBT for sex offenders have yet to be established, two studies highlight some early findings. Pithers (1994) found evidence to suggest that CBT programs are capable of reducing cognitive distortions among child molesters, and Eastman (2004) found increases in prosocial attitudes toward sexual behavior among adolescent child molesters after CBT. Unfortunately, the persistence of these improvements is unknown. In a related area, a meta-analysis examining changing rape-supportive attitudes among nonforensic samples, Brecklin and Forde (2001) found that the effects of these interventions did not last; in fact, effects of treatment disappeared just a few months after the intervention. According to Marshall and Laws (2003), lasting change will occur only if the schemas underlying attitudes supportive of sexual aggression are changed. Though there is much research on variables that influence attitude persistence (Petty & Wegener, 1998), we could not locate any study that looked at variables moderating attitude persistence in correctional programs. Also, to our knowledge, no reported studies have demonstrated that a change in either cognitions or attitudes is related to a change in recidivism among sex offenders; one study found no relationship between either therapist ratings of treatment progress or pre–post improvements in self-report attitudinal measures and recidivism (Quinsey, Khanna, & Malcolm, 1998).

Some studies have used alternative measures of treatment success to predict variability in recidivism. Marques, Wiederanders, Day, Nelson, and van Ommeren (2005) measured how well participants learned RP concepts, and found that men at high risk who learned the concepts recidivated less than high-risk men who did not (no differences were found for low- and medium-risk men); however, the program did not reduce recidivism overall. When Rice, Quinsey, and Harris (1991) looked at pre- and posttreatment phallometric scores, the former predicted recidivism better than the latter. Although it is difficult to draw conclusions about the success of CBT from the few properly designed studies conducted to date, the results so far are not encouraging.

The best-designed treatment program using CBT and RP methods was developed by Marques and colleagues (2005). Rice and colleagues (2001) outlined three major advantages of Marques and colleagues’ design: clinical sophistication (i.e., the study targeted criminogenic needs and changes throughout treatment); random assignment; and the use of “hard” outcome data (i.e., official records of violent, sexual, and general recidivism). Such a design allowed researchers to properly assess the relationship among specific treatment components, therapeutic change, and recidivism. Unfortunately, after much investment in planning and implementation, and after these men were followed up for at least 5 years, there was no effect of treatment on sexual or violent recidivism. Men who were
treated with this program had a sexual recidivism rate of 22%, whereas 20% of the volunteer controls (i.e., men who volunteered for treatment but were randomly selected for no treatment) and 19% of the nonvolunteer controls (i.e., those who refused treatment) had sexually recidivated.

Meta-analyses that included incidental control groups showed modest effects of treatment when the data were collapsed across all sex offenders (Hanson et al., 2002; Hanson & Bussière, 1998) and when the data for child molesters were examined more specifically (R. K. Hanson, personal communication, August 3, 2006), but Hanson and colleagues’ conclusions are tempered by the quality of the studies included in their meta-analyses. At the very least, an interesting question is whether treatment programs have shown improvement in their ability to treat offenders over time. In order to answer this question, we looked at the relationship between effect sizes of treatment on recidivism and time of study publication, with the assumption that a correlation should exist if improvements have been made in treatment programs over time. The general finding is that no such relationship exists (Box 11.1). A similar finding has been made within the same program over time. Quinsey, Khanna, and Malcolm (1998) compared sex offenders treated in a single program before and after 1986. Survival curve analysis—a statistic that compares change in recidivism rates over time—yielded no differences between the two groups, though a trend emerged where people treated later in the program reoffended more frequently than men who participated earlier in the program.

Many of these points can be understood as concerns about effectiveness and efficacy (Rice & Harris, 2003a; Streiner, 2002). “Efficacy” refers to whether treatment can work under ideal circumstances (i.e., a randomized control design), whereas “effectiveness” refers to whether treatment does work in the real world. Across all types of interventions, including psychotherapies and medical interventions, most studies are best classified as effectiveness studies. Streiner (2002) has stated that if treatment is not shown to work under ideal circumstances, there is no chance that it can work under suboptimal circumstances. Based on this line of reasoning, Rice and Harris (2003a) found that references to outcome studies had easily accepted positive treatment effects without considering their methodological limitations. They continued to say that any conclusion must be based on studies using at least minimally informative criteria. In order for a study to be minimally informative, it needs to (1) use official records of recidivism; and (2) use random assignment to treatment and control groups, or match comparison groups on jurisdiction, volunteering for treatment, and established predictors of recidivism (Quinsey et al., 1993; Rice & Harris, 2003a). Studies that fail to meet these criteria have no probative value.

These unfortunate results also need to be understood not simply as a failure of the treatment programs alone, but as a failure to assess exactly what has led to ineffective psychological interventions. A successful intervention requires not just the right intervention, but proper implementation and design, in order for treatment effectiveness to be properly assessed. The literature on program evaluation refers to this as a “process evaluation,” and it is considered an essential component of any program evaluation (Rossi, Lipsey, & Freeman, 2004). Because CBT is often implemented alongside other programs unrelated to cognitive or behavioral principles, it is difficult to assess separately, as noted by Rice and colleagues (2001). Moreover, though many clinicians record treatment progress on a daily basis, few studies report the relationship of such progress (or lack thereof) to longer-term outcome measures. Of the studies that were able to demonstrate changes in a variety of theoretically relevant measures involving attitudes, knowledge, and sexual
It has sometimes been assumed that considerable advances in the treatment of sexual offenders have occurred over the past 30 years. The unresolved question is whether these advances have made practical improvements in reducing subsequent sexual offenses. If these advances have allowed for better treatment programs, we should expect the effects of treatment on recidivism to improve over time. To test this question, we looked at the relationship between the year of publication and treatment effect size from the 44 studies used in Hanson and colleagues' (2002) meta-analysis on sex offender treatment effectiveness. Year of publication ranged from 1977 to 2000. Although effect sizes increased (as indicated by smaller odds ratios), there was no significant relationship between the year of publication and treatment effect size for sexual recidivism, $r(35) = -0.14, p = .43$, or any recidivism, $r(27) = -0.30, p = .12$.

This analysis includes a number of limitations: (1) There was variability in the time gap between when studies were actually conducted and when they were published (information about this could not be obtained); (2) relationship between improvements may depend on either the study design (i.e., efficacy vs. effectiveness) or the type of intervention. There are too few studies using optimal designs in order to assess their change over time, and so we are unable to separate analyses based on methodology quality.

These results can be understood in two ways. First, all results are in the expected direction, but are not significant. Even though the programs are not getting worse over time, a problem with this conclusion is that no significant relationships were found, and the one result that approached significance was found for any recidivism, not the intended outcome of changing sexual recidivism. A second interpretation is that no obvious improvement in the treatment of sex offenders occurred over the last 30 years. This conclusion is supported by two recent evaluations of CBT programs for sex offenders: the Twin Rivers Sex Offender Treatment Program in Washington State (Barnoski, 2006) and the Western Australian Sex Offender Treatment Unit (SOTU) Program (Greenberg, DaSilva, & Lob, 2002).

The Greenberg and colleagues (2002) report illustrates a result and a conclusion typical of much of the treatment literature. The investigators studied the records of 2,165 convicted male sex offenders referred to the SOTU between 1987 and October 2000. There were no significant differences associated with treatment in survival analyses, including those that controlled for risk. In the Executive Summary, the investigators observe:

Examination of the treated and untreated recidivism rates reveal there to be no significant effect of treatment. Despite this however it cannot be said with certainty that treatment is ineffective. Whilst it is clear that a strong treatment effect would have been apparent in the study results, it must be noted that there are methodological limitations of this evaluation, such as missing and incomplete data sets, which may have rendered it insensitive to smaller treatment effects. (p. viii)
They conclude:

Despite the lack of a strong treatment effect, it is the opinion of the reviewers that sex offender treatment is essential and necessary. The programs detailed in the SOTU manuals are fitting with international standards and with the addition of psycho-pharmacological treatment options and improved community-based post-release follow-up, are likely to have a long-term beneficial effect on sex offender recidivism, thus providing improved safety for our community. (p. ix)

We must conclude that these are indeed docile data.

The ultimate goal of sex offender treatment programs is to reduce sexual recidivism, and so there appears to have been no improvement in the ability to do so since publishing outcome studies began. Unfortunately, either conclusion depends on studies that allow us to assess true effect sizes of treatment. Considering that the bulk of studies used weaker designs, even if there had been an improvement over time, we would not be able to assess it properly. This unfortunate situation again highlights the importance of conducting more efficacy studies (see “Treatment Efficacy and Research Priorities”) or, more likely, admitting that different approaches must be tried.

preference, most have not shown that pre–post shifts were related to recidivism (Quinsey et al., 1998; Rice et al., 1991).

Marshall (2006) has argued that the effects of treating sex offenders are comparable to the effects of treating individuals with physical and mental health disorders. Although interventions with small (but statistically significant) effect sizes can be important, that importance is qualified by the cost of the intervention and the clinical significance of the effect. It is clear that the effects of treatment on recidivism pale in comparison to the effects of static risk factors (i.e., factors that do not change, such as criminal history). Meta-analyses on treating sex offenders show treatment effect sizes (using Cohen’s $d$) that range from 0.10 to 0.47, whereas actuarial assessments provide effect sizes that range from 0.64 to 0.76, with some studies showing effect sizes between 1.00 and 2.40 (Hanson, Morton, & Harris, 2003). Though some researchers have suggested that treatment and other dynamic risk factors add to the accuracy of actuarial assessments, as described earlier, none have demonstrated that changes in dynamic risk factors relate to changes in long-term recidivism among sex offenders (for further discussion of this topic, see Rice, Harris, & Quinsey, 2002). Any dynamic risk factor either must be able to predict recidivism outcome after actuarial scores are controlled, or must significantly add to actuarial models. Thus far, there is no evidence that dynamic risk factors add to the long-term predictive accuracy of actuarial instruments using static risk factors among sex offenders. In other words, we can still predict who will recidivate, regardless of whether they receive treatment or not.

In order to have legitimate expectations of CBT treatment success, we first need to identify whether there is a true link among cognitions, attitudes, and subsequent sexually aggressive behaviors. Second, we need to know whether treatment can reliably change
these cognitions and attitudes. Third, the effects of treatment on sustaining changes in cognitions and attitudes need to be established, especially when child molesters are released to the community. Fourth, researchers need to demonstrate that a change in treatment is related to a change in recidivism. Lastly, we need to show that changes from treatment predict recidivism over and above actuarial assessments. The first two points have received some support, but for the last three the evidence is negative or lacking.

**Multisystemic Therapy**

MST addresses general adolescent antisociality. Developed in the late 1970s, MST has become more popular in recent times for dealing with high-risk juvenile offenders. This approach is unique because it is a community-based treatment that targets multiple risk and protective factors, such as aspects of the individual, family, peers, school, neighborhood, and community. Borduin, Henggeler, Blaske, and Stein (1990) reported a random assignment study of MST with adolescent sex offenders. Though their results were highly encouraging (12.5% recidivism for treated offenders vs. 75% for untreated offenders at a 3-year follow-up), there were only 16 participants. Borduin and Schaeffer (2001) referred to an unpublished paper that reported similar results (12.5% vs. 41.7%) with a larger sample (N = 48). These sexual recidivism rates are far above those reported in other follow-up studies of adolescent sex offenders (e.g., Prentky, Harris, Frizzell, & Righthand, 2000, found 4% over a 12-month period) and children with sexual behavior problems (Carpentier, Silovsky, & Chaffin, 2006), leading us to wonder about the nature of the samples involved in the MST studies. Finally, MST is a less feasible approach to treating adult sex offenders because of difficulties in involving an adult's family, peers, and social network during community reintegration.

**Medical Treatments**

Medical procedures are used to reduce general sexual arousal by targeting sex hormones and regions of the brain related to sexual functioning. It is assumed that reducing general sexual arousal can help the management of any paraphilic thoughts and behavior (Berlin & Meinecke, 1981). Similar to other treatments for pedophilia, most medical interventions target sex offenders—not pedophiles per se—with the intended outcome of reducing recidivism. Historically, this form of treatment involved either surgical castration to reduce testosterone levels or surgical lesioning of brain regions related to sexual activity. These methods have become virtually obsolete because of the emergence of alternative procedures that are similarly effective in adjusting testosterone levels, but are less invasive and not irreversible (Hill, Brik, Kraus, Strohm, & Berner, 2003).

Medical treatments currently in use target either neural activity or hormonal regulation related to sexual arousal. One method uses selective serotonin reuptake inhibitors (SSRIs) to block the reuptake of serotonin, a neurotransmitter associated with sexual function. Fluoxetine and sertraline are the most common types of SSRIs used in this manner (Bradford, 2001). Although SSRIs have been successful in reducing sexual arousal (Meston & Frohlich, 2000), and this reduction may be sufficient for reducing sexual interest in children, there is no evidence of serotonin dysfunction in the paraphilias (Hill et al., 2003). The second method regulates sex hormone levels in one of three ways: (1) using medroxyprogesterone acetate (MPA) to inhibit gonadotropic secretions and reducing testosterone; (2) using an antiandrogen, such as cyproterone acetate (CPA), in order to
block androgen receptors throughout the body; or (3) overstimulating the hypothalamus by using a luteinizing hormone-releasing hormone (LHRH) agonist to reduce testosterone to levels usually attained by castration (Bradford, 2001; Briken, Nika, & Berner, 2001). There is evidence supporting the intended effects of medical treatment to reduce sexual arousal and paraphilic desire. For example, Bradford and Pawlak (1993) used a double-blind placebo crossover procedure with CPA to treat a group of paraphilic men. As expected, CPA had the intended effects on hormonal levels, sexual fantasies, and self-reported sexual arousal. Although the relationship between these targets and sexual thoughts and behavior are well established, we could not locate a paper that found aberrant levels of either serotonin or hormones to be diagnostic of pedophilia (or any other paraphilia).

Despite the strong relationship between neurobiological variables in sexual behavior and treatment, reducing general arousal does not alter sexual preference. Researchers found that men with phallometrically measured deviant sexual interests had the same preferences after hormonal treatment (Bancroft, Tennent, Loucas, & Cass, 1974; Cooper, Sandhu, Loszyn, & Cernovsky, 1992). Also, in Bradford and Pawlak’s (1993) study described earlier, treatment using CPA showed no changes in PPG-measured arousal throughout treatment, only changes from self-report arousal; this leads us to suspect that demand characteristics may account for this discrepancy (i.e., child molesters are more likely to self-report a change in treatment even if no change actually occurred). It appears as though treatment for paraphilias works by decreasing sexual interest in general, suggesting that medical treatments do not “cure” the sexual preference but mask it by reducing sexual desire. Still, these interventions could be useful and effective among forensic groups if a change in general sexual interest results in a change in sexual recidivism.

To date, there have been no randomized assignment studies relating medical treatment to recidivism among pedophiles. Hanson and Harris (2000) unexpectedly found that men who started antiandrogen treatment were more likely to recidivate than men who did not start treatment, presumably because only high-risk men received antiandrogen treatment. Despite their finding, a meta-analysis showed a moderately lower recidivism rate among those treated medically than men who did not receive medical treatment (Hall, 1995), but this latter finding must be understood in light of the methodological quality of the relevant studies. Hall noted considerable heterogeneity across samples: Follow-up times, source of recidivism data, and recidivism base rates all differed across studies. These and other medical studies also lacked placebo groups and provided treatment on a voluntary basis, using a treatment-refusers-as-controls design (Rice et al., 2001). Also, almost all medical papers specifically addressing pedophilia have been case studies (e.g., Bourgeois & Martina, 1996; Varela & Black, 2002); medical studies with larger samples have collapsed their data across paraphilia subtypes (e.g., Bradford & Pawlak, 1993; Hall, 1995; McConaghy, Blaszczynski, & Kidson, 1988), making it difficult to draw conclusions about treatment effectiveness specific to pedophilia. In view of the many possible confounding factors, the results of these treatment studies should be interpreted with caution.

Drawbacks unique to medical treatments for pedophilia include side effects and noncompliance. Commonly cited side effects include hypertension, hyperglycemia, feminization, depression, and headaches (Hill et al., 2003; Saleh & Guidry, 2003). These adverse effects are likely to cause the low compliance rates with medical treatments. Hucker, Langevin, and Bain (1988) found that of the 100 men eligible for treatment, only 18
began treatment, and just 11 completed the 3-month intervention. Across the literature, Barbaree and Seto (1997) found noncompliance rates to range from 30% to 100%, though Langevin and colleagues (1979) were able to increase compliance by providing assertiveness training. Because a possible consequence of noncompliance is an elevated risk of recidivism (Hanson et al., 2002), dealing with dropouts should receive greater attention.

**Combining Treatment Approaches**

Most programs include treatment “modules” that are separate interventions targeting known or assumed criminogenic needs. Typical treatment programs for sex offenders include heterosocial skills training, RP, modification of paraphilic arousal, and sex education, with an emphasis on attitudes and values. There has been little research on how such combinations of modules interact to reduce recidivism (Palmer, 1995). Others have provided a rationale for particular combinations of treatments, such as Bradford’s (2000, 2001) algorithm to treat paraphilias. Bradford suggested that treatment should first categorize paraphilias into one of four levels of severity—mild, moderate, severe, or catastrophic—and then use these categorizations to determine the level of treatment. For example, CBT and RP are provided for all levels, but full antiandrogen or hormonal treatments are reserved for all severe and catastrophic cases. Other treatment algorithms for paraphilias have been described in the literature (e.g., Hill et al., 2003). Though some evidence suggests that a combination of medical and CBT treatments will provide optimal outcomes (e.g., Bradford & Greenberg, 1996), the utility of specific treatment algorithms still needs empirical support.

**Treatment Efficacy and Research Priorities**

The research described in this chapter provides a somber view of the effectiveness of treatment for pedophilia. Meta-analyses have shown either weak or no effects of treatment on sexual recidivism (Hanson et al., 2002; R. Lieb, personal communication, August 8, 2006). These meta-analyses are bedeviled by the methodological quality of the studies evaluated—especially the limited use of randomized control designs, and the difficulty of distinguishing treatment dropouts and refusers from completers when control groups include men who have dropped out or refused treatment if they had been given the opportunity (this procedure artificially decreases the recidivism rate of the treatment group; see Rice & Harris, 2003a). With these considerations in mind, along with efficacy studies showing no effect of treatment (Marques et al., 2005), is there anything at this point that clinicians can do to treat pedophilia? Apart from designing new studies based on a better understanding of the theoretical links between treatment targets and behavior, some recommendations for the design of interventions to reduce recidivism do emerge from the correctional “what works” literature.

These recommendations are best understood from Andrews and colleagues’ (1990) principles of risk, need, and responsivity for effective correctional intervention. In their meta-analysis, correctional programs had greater reductions in recidivism when they targeted (1) offenders at high risk of recidivism; (2) criminogenic needs (i.e., aspects of individuals that are empirically linked to criminal behavior); and (3) responsivity to treatment (i.e., adapting programs to learning styles of offenders). Despite such an important
finding, a review of the literature 9 years later identified only 10% of outcome studies reporting recidivism rates for higher- and lower-risk groups (Andrews, Dowden, & Gendreau, 1999). Of these studies, Andrews and Dowden (2006) found moderate support for the risk principle. The correlation between presence (vs. absence) of primary treatment and recidivism was .03 for low-risk and .10 for high-risk offenders. Andrews and Dowden also discovered that the risk principle interacted with the criminogenic needs principle; that is, the risk principle was supported only when programs targeted criminogenic needs. Unfortunately, these quantitative reviews collapsed data across all types of offenders, so we are not sure whether these effects would be different for sexual versus violent offenders. We suggest that if resources are scarce, clinicians should utilize these potentially useful principles of correctional intervention, especially by focusing on high-risk offenders (Quinsey et al., 2006).

We can also use Andrews and colleagues’ (1990) principles of correctional intervention to help us understand some reasons why treatments for sex offenders are not successful. If pedophilia is a major path to child molesting, a possible explanation for why sexual preferences have been difficult to change might be that treatment programs do not actually address the neurohormonal or psychological mechanisms controlling sexual preferences. In other words, programs have not been able to treat a central criminogenic need. By drawing from the biological literature on sexual behavior, we can get a better understanding of how these mechanisms work. One of the most important periods for sexual preference development occurs prenatally, when hormones have neural organizational effects (Ellis & Ames, 1987; Meyer-Bahlburg, Ehrhardt, Rosen, & Gruen, 1995). As expected, evidence for an environmental influence on homosexuality occurs prenatally. Blanchard and Bogaert (1996; Bogaert, 2006) found that with every older brother a male has, the chances of being homosexual increase by 33%; this is known as the “fraternal birth order effect.” It has been theorized that the fraternal birth order effect occurs because each male fetus, the maternal immune system becomes more sensitized to some aspect of the male fetus, perhaps the male fetus’s H-Y antigens (Y-chromosome-linked histocompatibility antigens found on fetal cells). Quicker and stronger responding of the maternal immune system impedes masculinization of the brain. It has been theorized that among pedophiles, similar neurodevelopmental incidents may alter not just gender preferences, but age preferences as well (Quinsey, 2003). Supporting this view, Lalumière, Harris, Rice, and Quinsey (1998) found that a fraternal birth order effect was related to PPG indices of sexual deviance (i.e., a preference for children or coercive sex), though no such birth order effect was found among heterosexual pedophiles in another study (Blanchard et al., 2000).

Two further comments on the child molester treatment literature are relevant here. The first is that very few sexologists believe that modifying homosexual orientation via current technologies is possible (this is not to argue that such a goal is desirable). It is more than a little surprising that no one applies this central conclusion from the literature on the treatment of homosexuality (now moribund) to the treatment of pedophilia; pedophilia is doubtless a different condition, but nevertheless also an anomaly of sexual preference.

Second, we are a very long way from determining the effective ingredients of treatment programs for child molesters. In much of the psychotherapy literature, treatment gains are easily demonstrated (at least in the short term), and researchers focus their attention on trying to disentangle the effects of specific treatment procedures and related
Outcome studies reported moderate to high percentages of primary and secondary offenders. Andrews and colleagues (1998) noted that the criminogenic programs targeted for this population resulted in minimal data across all offenders. Sexual offenders should utilize resources by focusing on their specific needs, see Quinsey et al. (1996).

Correctional interventions are not successful for why sex offenders do not respond to treatments. The neural criminogenic programs are a theoretical approach to sexual desires, see the neural organization of sex offenders, see Gruen et al. (1995).

Male sexual orientation occurs in early childhood. It has been theorized that the order effect may be stronger in men. The Y-chromosome may influence sexual orientation in men.

Dopamine is another ingredient of treatment-resistant sexual behavior. Treatment approach, treatment for sexual disorders, and related factors from placebo and expectancy effects. In the treatment of pedophilia, investigators remain divided over whether there is any overall treatment effect.

After reviewing each treatment approach, we can highlight the central reasons why they have not worked in treating pedophilia:

1. Conditioning/behavioral: Sexual preferences are neither developed nor maintained by conditioning processes.
2. General psychotherapy: Treatments are unrelated to sexual preferences or criminogenic needs.
3. Cognitive: CBT addresses cognitions and attitudes correlated with antisocial behavior, but these have no theoretical link to causing or maintaining atypical sexual preferences.
4. RP: This approach is based on a theory of maintenance, not etiology.
5. Medical: Hormones are used to reduce general sexual arousal, but this does not address pedophilic interests because in adulthood, hormones have motivational rather than organizational effects. If in fact pedophilia results from neurodevelopmental incidents, a better understanding of the psychological and neural processes affected by such incidents is needed to guide the development of novel treatment programs.

SUMMARY AND FUTURE DIRECTIONS

The etiology of pedophilia remains obscure, although (mostly indirect) evidence points to the importance of perturbations of neurohormones during early brain development. Recent advances in brain imaging (Cantor et al., 2006) promise to provide more direct clues to the etiology of this sexual anomaly, although behavior genetic investigations should also receive high priority. Phallometry remains the best-developed method of identifying abnormalities in male sexual preference, but direct measurements of brain activity are likely to supersede these methods soon.

Treatment methods for child molesters appear to have produced quite limited reductions in recidivism, despite the ability of some programs to effect proximal improvements in variables assumed to cause child molestation. In our view, these dismal results stem from the failure to develop a method that can durably alter the central criminogenic need factor in pedophilia—sexual preference for children. More etiologically oriented research will be required to remedy this deficiency. It is possible that preventative interventions will prove to be more easily developed than cures for the condition once it is established.

On the other hand, there is hope that CBT approaches developed for offenders more generally can be profitably applied to reducing the recidivism of child molesters who are not pedophilic. There remains grave doubt, however, about the applicability of these methods for the most antisocial of these men. Fortunately, these psychopathic individuals are relatively rare.

ACKNOWLEDGMENTS

We would like to thank R. Karl Hanson for reanalyzing his meta-analysis using only child molesters, and Grant Harris, Martin Lalumière, and Michael Seto for their helpful comments.
NOTES

1. Another criticism of the DSM-IV diagnostic criteria is the use of clinical criteria. A result of using clinical criteria, such as distress, is the exclusion of men with pedophilic sexual interest who do not experience affective symptoms (see Spitzer & Wakefield, 1999). Revisions to DSM-IV-TR have corrected this issue by including clinical criteria as a sufficient, not necessary condition for diagnosis.

2. Among child molesters, a correspondence between sexual attractiveness ratings and penile tumescence responses for adolescents and children was not found; self-reports suggested a preference for adult females, whereas phallometric assessment suggested a preference for children (cf. Quinsey, Steinman, Berghersen, & Holmes, 1975). Such a result, however, is probably due to biased reporting (consider the legal implications of divulging a sexual preference for children).

3. Some have even argued that the relationship between attitudes and behavior is weak (Eckes & Six, 1994).

4. Medications used to treat paraphilias are better recognized by their brand names, such as Prozac (fluoxetine), Zoloft (sertraline), and Depo-Provera (MPA). These drugs were initially designed to treat other illnesses, such as cancer and depression; because reduced interest in sex was a side effect of these treatments, secondary uses emerged.

5. Though Aos, Miller, and Drake (2006) initially found an effect of treatment on sexual recidivism, a recent update found nonsignificant results (R. Lieb, personal communication, August 8, 2006).

6. The risk principle applies in the sense that management and intervention priorities should be directed at those at risk of causing greatest harm to society. The claim that treatments targeting high-risk men result in greater reductions in recidivism is challenged by research demonstrating that men at highest risk (i.e., psychopaths, especially sexually deviant psychopaths) still have extremely high recidivism rates even after treatment (Gretton et al., 2001; Hildebrand, de Ruiter, & de Vogel, 2004; Rice & Harris, 1997; Rice, Harris, & Cormier, 1992; Serin, Mailloux, & Malcolm, 2001; Seto, 2005).

REFERENCES


tepdauughters, bio-
ization in
sexual orientation. 774.
tic treatment of ad-
omen of pedophilia 257–258.
ological approach.
tical treatment in Psychiatry, 46(1),
viant sexual behav-
arousal patterns of
s. Violence and Vic-
; hormone-releasing
the Association for
ent for children nd Clinical Psychol-
24(1–2).
, R. D. Da-
). New York:
enders: A cross-over
enschaft: Eine
ion: A theory of ho-

Hildebrand, M., de Ruiter, C., & de Vogel, V. (2004). Psychopathy and sexual deviance in treated rap-


relapse prevention program on sexual recidivism: Final results from California’s Sex Offender Treatment and Evaluation Project (SOTEP). Sexual Abuse: A Journal of Research and Treatment, 17(1), 79–107.


Rice, M. E., & Harris, G. T. (2003b). What’s shape got to do with it?: Age and waist–hip ratio prefer-
ences among child molesters and normal controls. Paper presented at the 22nd Annual Conference of the Association for the Treatment of Sexual Abusers, St. Louis, MO.


