Plethysmography in the Assessment and Treatment of Sexual Deviance: An Overview

Walter T. Simon, Ph.D.¹ and Peter G. W. Schouten, M.A.²

Theoretical and practical implications of previous research on plethysmography in the assessment and modification of sexual preferences has been limited by insufficient attention to definitional and procedural issues and sample heterogeneity. This paper reviews current theory and research on the use of the penile plethysmograph to diagnose or predict sexual deviance and to monitor therapeutic change in the modification of sexual preferences. Phallometric tests are discussed in terms of the basic components of the measurement technique, evidence of technical adequacy, and general principles underlying their use. The review points to a need to guard against false expectations and erroneous conclusions about sexual deviance based on this measurement technique. A case is made for closer attention to previously untested working assumptions and issues of definition, etiology, and measurement. Directions for future research are proposed.

KEY WORDS: sexual deviance; assessment; penile plethysmograph.

INTRODUCTION

The physiological aspects of the human sexual response have been the focus of research on normal sexuality, sexual deviation, and sexual dysfunction. The physiological measurement of arousal was an important component in the pioneering work of Masters and Johnson (1966, 1970) and subsequent research on male sexuality (Quinsey and Marshall, 1983; Rosen and Keefe, 1978). The parameters of sexual arousal are of interest in the

¹Darrow Clinic, 777 Grant Street, Suite 605, Denver, Colorado 80203.
²Focus Research, 3960 Vallejo Street, Denver, Colorado 80211.
scientific study of human sexuality, and also in clinical and legal responses to alleged or known sexual deviations.

The penile plethysmograph provides a measure of tumescence in response to fantasy or laboratory stimuli. Early research with this instrument focused on sexual preferences as to the age or sex of partner and specific sexual activities (Freund, 1963). Quinsey (1973) emphasized its potential in discharge decisions regarding institutionalized child molesters. Freund et al. (1979) later argued its value in the diagnosis of "non-admitters." Although routinely used for clinical purposes, particularly with forensic populations, basic theoretical premises underlying this measurement paradigm have not been clarified and its technical adequacy remains in question. Drawing upon recent research, this paper reviews major conceptual and methodological problems in plethysmography for the evaluation and treatment of sexual deviation.

THEORETICAL AND CONCEPTUAL BACKGROUND

Use of the plethysmograph in the diagnosis and treatment of sexual deviance reflects an implicit, but not necessarily accepted, theoretical explanation of sexual deviance. The individual is thought to act upon a sexual aim (a pattern of behavior) or chooses a sexual object which then becomes a target. Overt sexual acts presumably reflect sexual arousal that is elicited by the individual's fantasy of the act or object (Freud, et al., 1979). Certain preferences are thought to have specific behavioral consequences. Further, distinct erotic interests can presumably be inferred from laboratory measures of penile responses to normal and deviant stimuli. These premises warrant closer examination.

Working Assumptions of Phallometric Testing

Plethysmography in the assessment of sexual deviance assumes a natural operational linkage between stimulus-specific penile arousal patterns and overt sexual acts. This view is the presupposition of the use of phallometric tests to diagnose sexual deviance and predict dangerousness. The assumption of arousal-behavior relationship is also central to a behavior modification doctrine which emphasizes direct attempts to modify the penile arousal patterns that are thought to lead to deviant behavior (Abel and Blanchard, 1976; Earls and Marshall, 1983; Laws and Osborn, 1983; Quinsey, 1973; Quinsey and Marshall, 1983; Rosen and Kopel, 1978). The goal of this treatment approach is reduced arousal to deviant sexual
stimuli and increased arousal to appropriate stimuli (Bancroft, 1970, 1974). Given this therapeutic goal, the plethysmograph might appear to be a logical choice as a measure of treatment effectiveness.

Available evidence does not provide unequivocal support for the hypothesized relationship between sexual arousal in the laboratory and overt sexual acts. Furthermore, to explain sexual interests and acts strictly in terms of categories of stimulus specific arousal is to ignore the role of individual differences variables in mediating relationships between these distinct, though related, aspects of the human sexual response. Examples would include an impotent pedophile whose deviant acts involve fondling rather than actual or attempted intercourse or exhibitionists who expose themselves without accompanying erection. Rosen and Kopel (1978) pointed out that "In such cases it appears that the inappropriate behavior is functionally independent of genital response" (p. 1520). In these and other cases, phallicometric measures would not be relevant to either assessment or treatment. Individual differences variables that may mediate the relationship between physiological measures to overt behavior must be considered in evaluating the validity of plethysmography in the assessment of sexual preferences.

TEST AND MEASUREMENT PARADIGM

Empirically derived assessments of deviance have a statistical basis for confidence in diagnosis and prediction (Gabor, 1986). In plethysmography, evidence of the validity of conclusions based on test results must be evaluated in light of measurement issues. In this paradigm, sexual deviance has been operationally defined as a penile response pattern elicited in the laboratory. Philosophical issues aside, a basic question remains: Does the plethysmograph accurately assess male erectile responses? (Farkas, 1978; Farkas et al., 1979; Freund et al., 1988). To answer this question, the basic components of the measurement technique and testing and scoring procedures must be evaluated.

TRANSDUCERS AND RECORDING EQUIPMENT

Phallometric measures are commonly obtained using a strain gauge (Barlow et al., 1970) or mercury-in-rubber gauge (Bancroft et al., 1966). Circumferential measures are contrasted with penile volume measures. The latter are designed to assess complete volume changes in terms of both penis length and diameter. Volumetric devices and strain gauges have been
compared (Wheeler and Rubin, 1987) and have been found to provide conflicting results (McConaghy, 1974; Freund et al., 1974). The relationship between volumetric and circumferential measures is unclear due to differences in metrics. However, different types of circumferential measures may also yield conflicting results (Laws, 1977).

Increased sensitivity and precision in measurement may be possible with a better understanding of the operating characteristics of the penis, particularly at low levels of responding. Earls and Marshall (1983) found that in the initial stages of penile responding, the magnitude of increases in length (28%) is substantially greater than that of diameter change (2.5%). They noted that “at 10% of diameter-assessed arousal, almost half of the total length change had occurred” (p. 349). Earls and Marshall (1983) concluded that circumferential measures are inadequate because they fail to assess the lengthening of the penis during erection. Current phallometric techniques may also fail to tap other potentially important, but as yet unidentified, physiological dimensions of the human sexual response.

**Laboratory Stimulus Materials**

Plethysmography utilizes various media to provide erotic cues to produce sexual arousal in the laboratory. These include pictures of nudes, movies, erotic stories, and audiotaped narratives. Certain test stimuli presumably constitute “relevant inappropriate material” (Laws and Osborn, 1983; Quinsey et al., 1981) suitable for eliciting deviant arousal. The importance of stimulus materials in phallometric assessment is apparent from evidence showing that subjects’ ability to bring arousal under voluntary control is influenced by stimulus content (Malcolm et al., 1985). However, the critical role of these materials in the laboratory assessment of sexual deviance has been largely overlooked. Few studies have directly compared the effectiveness of modalities or the impact of erotic cues within a single class of stimuli (Abel et al., 1975; Abel, Blanchard, and Barlow, 1981). In fact, studies in this area have typically provided little descriptive detail regarding the stimulus sets used.

Recent research has underscored a lack of precision in the use of laboratory assessment of pedophilia (Fuller et al., 1988). To confirm a diagnosis of pedophilia, a subject may be exposed to pictures of nude prepuberal children since this diagnostic category is defined by an erotic preference for children at this stage of development (American Psychiatric Association, 1987). Pedophilic stimuli have traditionally been selected on the basis of age. However, age categories are imprecise because persons
within a given age range are characterized by idiosyncratic patterns of
pubertal change and varying degrees of physical development (Fuller et
al., 1988). The distinction between developmental and age-graded
categories has major implications for the assessment of sexual deviance.
The distinction means that some pedophiles may be aroused only by child
sexual stimuli which are defined by highly specific developmental aspects;
others may be aroused by materials within much broader classes of child
stimuli. It is not always possible to determine what type of stimuli are
most effective.

The impact of laboratory stimuli is dependent not only on aspects of
physical development or sexual maturity. For example, in a study reported
by Wormith (1986), a blonde model elicited significantly greater arousal
than a redhead. Interactions between model attributes (bodily shapes, facial
features, skin color, extent of undress, posture, etc.), setting, the exact na-
ture of sexual activity, or the quality of pictures may also contribute to
variations in sexual arousal.

Still other aspects of the stimulus materials must be considered,
namely, elements of consent and violence. Quinsey et al. (1984) found no
differences in the phallometric responses of rapists and non-sex-offenders
to audiotaped sadomasochistic stories. The appropriateness of
sadomasochistic themes in the assessment of rapists is unclear.

Among pedophiles elements of consent have been linked to greater
arousal to child stimuli (Hall et al., 1988). Clearly, violence and consent
are critical dimensions of laboratory stimuli used in the assessment of al-
leged or known sexual offenders and their treatment. In this context, it
should be noted that standardized laboratory stimuli may not be useful for
treatment applications of the plethysmograph. As Rosen and Kopel (1978)
suggested, stimulus materials used in the modification of arousal patterns
should be specifically developed to approximate an individual's unique
erotic interest. Such an approach reflects a realistic appraisal of the
idiosyncratic nature of sexual preference.

**Duration of Stimulus Presentation**

Duration of stimulus presentation may be an important factor in the
accurate assessment of sexual preferences. The strength of penile responses
may be influenced by how long a subject is exposed to stimulus materials.
Examination of recent research reveals that duration of presentation may
vary widely with stimulus modality, ranging from relatively brief exposures
for individual slides to extended videotape or audiotape sequences
Some subjects may not achieve levels of arousal indicative of their true interests if given insufficient exposure. Longer exposure times may facilitate stronger responses, but may also give subjects more time to modify their responses. These issues are complicated by individual differences in responsiveness, with some subjects achieving full erection much sooner, and some much later, than others in the presence of certain stimuli. Attempts to determine optimum duration of stimulus presentation must take into account differences between modalities, the impact of specific content within modalities, and their interaction with individual differences in responsiveness.

**Scoring Procedures**

The penile response can be measured in several ways (Abel, Blanchard, Murphy, et al., 1981; Earls et al., 1987). Penile circumference change is the most common, and is assessed as amount of change in millimeters. In view of individual differences in penis size, the raw data may be transformed to percentage of full erection. Alternatively, because many subjects do not achieve full erection in the laboratory, the raw data may be transformed to reflect magnitude of change from baseline using z-score transformations (e.g., Wormith, 1986). For a sample of paid volunteers, Earls et al. (1987) found that z scores were more sensitive to differences in stimulus categories than raw score and percentage change indices. In addition, these researchers pointed out that unlike the percentage change index, the z score method does not require an upper limit (i.e., full erection). Data reported by Earls et al. (1987) indicate that the amount of variance accounted for within a sample is influenced by scoring method. Thus the validity and generalizability of individual and group data are qualified by the type of scoring method used.

A commonly used ratio index of deviant arousal is calculated by dividing arousal to inappropriate material (e.g., depictions of rape) by arousal to appropriate material (e.g., depictions of consenting sex). Quinsey and Marshall (1983) summarized the rationale for this ratio index: “What differentiates child molesters and rapists from non-sex offenders are their penile responses to the relevant inappropriate material in relation to their responses to appropriate material” (p. 268). As noted in our foregoing discussion of test stimuli, issues of definition relating to the appropriateness of stimulus materials are difficult to resolve. In addition, it is not clear how the ratio index would function in the assessment of individuals with mixed, rather than pure, erotic interests. For example, a nonexclusive pedophile may be equally aroused to child and adult laboratory stimuli. The individual
may be interested in both children and adults, but no less deviant than an exclusive pedophile.

Controlling for Faked Responses

Erectile responses can be faked or suppressed to simulate a normal sexual response profile (Freund, 1963; Hall et al., 1988). Voluntary control of penile tumescence appears to be mediated by both cognitive and mechanical processes. Mechanical distortions may involve muscle contractions or "pumping" behaviors (Freund et al., 1988). The cognitive aspects may include elaborations of laboratory stimuli or inattention to those stimuli (Abel and Blanchard, 1976; Laws and Rubin, 1969). These processes in turn may be regulated by factors such as vividness of imagery (Smith and Over, 1987) and other individual differences in proneness to fantasy or subjective appraisal of sexual stimuli which are currently not well understood.

Concerns about the effect of voluntary control on test results have led to systematic efforts to identify key signs of faking and to devise means of offsetting it (Freund et al., 1988; Malcolm et al., 1985; Quinsey and Bergersen, 1976; Wydra et al., 1983). This work has yielded a number of potentially useful and theoretically interesting findings about the conditions under which faking is more likely to occur, and has provided direction for determining the parameters of voluntary control and the study of individual differences in faking (Freund et al., 1988; Malcolm et al., 1985). For example, Freund et al. (1988) found evidence that age preference is easier to feign than gender preference. Malcolm et al. (1985) found that a forensic sample consisting primarily of rapists exhibited greater control of penile responses in the presence of nonpreferred sexual stimuli and less control in response to preferred stimuli.

However, there are at present no generally accepted procedures for estimating the probability of faking, detecting it, or controlling for it (Malcolm et al., 1985). Available methods of minimizing or compensating for it are inadequate (Earls and Marshall, 1983). Rather than directly improving the accuracy of measurement, these strategies may only provide a basis for concluding that test results do not reflect an individual's true sexual preferences. Even what is considered the most obvious clue to faking is an ambiguous indicator. That is, if a subject shows no tumescence, this could mean either suppression of arousal or lack of interest (Earls and Marshall, 1983).
TECHNICAL ADEQUACY

The concept of internal validity relates to methodological soundness and the extent to which findings are free from the confounding effects of extraneous factors (Campbell, 1957). Farkas (1978) noted two major threats to internal validity, namely, subjects' ability to exert voluntary control over their penile responses and the weak relationship between self-reports of arousal and phallicometric indices. Additional challenges to the internal validity of these indices are apparent with respect to reliability, the discriminability of sexual preference, classificational accuracy, sample selection, and procedural and contextual confounds. Threats to external validity are also noted in this review of recent research findings.

Reliability

Campbell (1969) pointed out that "instability" is a threat to internal validity. This is an issue given the common practice of repeated phallicometric assessment to monitor treatment efficacy. It is also an issue in view of the state–trait distinction (Zuckerman, 1971). Sexual arousal as measured by the plethysmograph may reflect a temporary state of sexual tension rather than a stable, trait-like tendency to become aroused to a special class of sexual stimuli. An individual's arousal pattern must be sufficiently stable across time to serve as a basis for determining sexual preferences or for predicting future activity.

As Eccles et al. (1988) have pointed out, although reliability is central to the notion of sexual preference, few studies have directly addressed this issue. Evidence of temporal stability for penile response patterns has been mixed and is qualified by the length of the test–retest interval (e.g., Freund et al., 1979) and by selection bias (e.g., normal volunteers vs. offenders) or inadequately defined selection criteria. Further, reliability is influenced by stimulus content. For example, Wormith (1986) reported a wide range of reliability coefficients for a relatively large sample of sex offenders and non-sex-offenders presented with male and female adult and child stimuli.

In previously published research, reliability coefficients for phallicometric scores have ranged between .38 and .94 (Earls and Marshall, 1983; Wormith, 1986). Uncontrolled-for differences in test procedures and environmental conditions (e.g., laboratory stimulus materials, instructional set, examiner characteristics and demand characteristics), may account for variations in reliability.

In addition, reliability coefficients may vary considerably with different scoring methods (e.g., Wormith, 1986). Test–retest reliabilities of
penile responses calculated using standard scores have been found to be lower than those of self-reports of sexual interest (Wormith, 1986). This finding is interesting in view of the controversy regarding the relative merits of phallometric and self-report measures in the assessment of sexual preferences (Barbaree and Marshall, 1989; Quinsey, 1973; Zuckerman, 1971).

An issue of statistical conclusion validity (Cook and Campbell, 1979) should be noted here. Specifically, estimates of the temporal stability of an individual diagnosis should be expressed in terms of classificational accuracy rather than correlation coefficients. It is necessary to determine the proportion of subjects consistently classified at different times.

Classificational accuracy is also relevant in determining agreement consistency between independent judges in the assessment of sexual deviance. Interjudge agreement studies are needed to justify confidence in diagnostic classifications that are based on phallometric findings. But classification decisions must be both accurate and consistent. Meaningful use of an agreement index presupposes optimum cutting scores (Rorer et al., 1966). However, there are presently no generally agreed-upon guidelines as to normal and deviant phallometric response ranges.

The clinical assessment of sexual arousal to monitor treatment progress in terms of deviant and nondeviant response levels may call for repeated measurement. Treatment effects may easily be confounded with habituation. Researchers and clinicians are therefore concerned about the impact of repeated measurement on reliability. In a sample of college students, Eccles et al. (1988) found no evidence of habituation for college subjects who were tested on the plethysmograph up to nine times in a 3-week period. Interestingly, in a previous study with college volunteers, O'Donohue and Greer (1985) found evidence not only of habituation to both constant and variable stimuli but successfully unconfounded it from physiological fatigue.

Repeated testing has also been linked to increased potential for faking through familiarity with the procedures and stimuli (Quinsey et al., 1983) and practice in controlling arousal (Earls and Marshall, 1983). Faking, rather than habituation, is particularly of concern in pretrial evaluations and the assessment of offenders involved in court-ordered treatment who are motivated to demonstrate progress.

The Discriminability of Sexual Preferences

Clear evidence of differences in the penile response patterns of normals and sexual deviates would support the utility of the plethysmograph
in the assessment of deviance. Evidence along these lines has been mixed. Normal volunteers have been found to respond with arousal to slides of 8- to 11-year-old girls and to slides of the buttocks of prepubescent boys (Freund, 1972). Furthermore, non-sex-offenders also become aroused in response to slides of pubescent females (Barbaree and Marshall, 1989; Freund et al., 1972; Quinsey et al., 1975), and men in the general population may exhibit arousal patterns similar to those of rapists (Barbaree et al., 1979; Malamuth et al., 1986; Quinsey et al., 1981). These findings do not show that most nondeviants are as attracted to deviant stimuli as are deviants. Rather, the findings indicate that arousal to deviant themes does not unequivocally confirm sexual deviance.

Plethysmography presupposes the ability to differentiate sexual deviates and normals as well as deviant subtypes on the basis of arousal. Evidence of discriminant validity has been ambiguous. In a study reported by Freund et al. (1986), rapists could not be distinguished from an entire family of deviant sexual preferences, including exhibitionism, frottage, and voyeurism. Hall et al. (1988) found no differences between rapists and molesters (victims to age 16) in terms of penile responses to four categories of stimuli (consensual sex with an adult, consensual sex with a child, sexual assault of a child, and nonsexual assault of a child).

Classificational Accuracy

Statistical conclusion validity, specifically concerning the use of appropriate statistical methodology, is a special case of internal validity (Cook and Campbell, 1979). In plethysmography, this issue has so far been largely overlooked with respect to the use of correlational techniques to relate phallicometric indices to diagnostic group criteria. Correlational data are inadequate evidence of validity because such data do not indicate the probability that a test will accurately identify individuals in terms of the proportion of subjects accurately classified (Meehl and Rosen, 1955). Evidence of classificational accuracy must be evaluated independently because in some distributions high rates of misclassification may occur even when correlational data suggest substantial concurrence between predictor and criteria. The validity of phallicometric findings with respect to group criteria cannot be evaluated without classificational statistics.

Wormith (1986) reported a discriminant analysis of the phallicometric data of rapists, pedophiles, and convicts incarcerated for nonsexual offenses. The results indicated misclassifications of up to 36%. Inspection of the classificational results reveals that 42% of the pedophiles were classified as having normal sexual preferences and 33% of those with normal sexual
preferences were erroneously assigned to the rapist group. Also of interest in this context are the Barbaree and Marshall (1989) findings. Specifically, 40% of incest offenders and 68% of matched nonoffenders exhibited a normal adult profile. Another 40% of incest offenders exhibited nondiscriminating patterns. A subset of incest offenders and nonfamilial child molesters responded with arousal to both child and adult stimuli. Only 35% of child molesters showed a pure “child” profile (Barbaree and Marshall, 1989).

Plethysmography has proceeded on the assumption that deviant sexual interest can be inferred from higher degrees of arousal to deviant stimuli relative to arousal to appropriate stimuli (Abel et al., 1977). Hall et al. (1988) found that some between-stimulus category correlations either approached or exceeded the correlations within stimulus conditions for a large sample of sex offenders. They concluded that, contrary to the notion of stimulus-specific responding, arousal among sexual offenders occurs in a generalized manner.

External Validity

In plethysmography, evidence of external validity relates to the ability of laboratory findings to predict external criteria, such as history of deviance or recidivism (Earls and Marshall, 1983; Farkas, 1978). This type of validity overlaps with discriminant validity, which is the focus of studies that compare sexually deviant subtypes and normals. The external validity of phallometric data is also relevant in demonstrating change maintenance for interventions that are designed to modify penile response patterns. External validity is essential to justify clinical uses of the plethysmograph.

Studies have reported weak to moderate correlations between phallometric indices and actual sexual behavior. For example, in a study with 155 male volunteers, Malamuth (1986) found a correlation of .43 between physiological arousal to depictions of rape and self-reported history of sexual aggression. Thus the phallometric rape index accounted for about 19% of the variance in the criterion. In a study of long-term treatment effectiveness reported by Marshall and Barbaree (1988), unofficial records were better predictors of child molesters’ recidivism than were phallometric data. Phallometric data permit stronger conclusions concerning sexual preferences when considered along with other predictor variables (Malamuth, 1986).
Sample Selection

Selection is a threat to internal validity and also places constraints on generalizability (Cook and Campbell, 1979). As in any research paradigm, a data set reflects a given sample. Evidence of the diagnostic and prognostic value of phallometric data has in large part been based on convicted offenders, homosexuals, and transvestites (Earls and Marshall, 1983). The use of such groups prevents generalization to college samples and also to other forensic samples (e.g., alleged or undetected offenders). For example, Saunders and Awad (1988) noted a dearth of phallometric research with adolescent sex offenders.

Selection bias may also be related to the use of groups of subjects who may be differentially influenced by demand characteristics. Offenders may be motivated to exhibit normal sexual interests. Such motivations may not be operative among college samples. Little is known about potentially important differences among diverse samples in this and other respects or about the impact of selection on clinical and research data. Indeed, studies in this area have often provided insufficient detail concerning sample selection and subject characteristics. For example, Malcolm et al. (1985) cited unpublished selection criteria for research with sexually deviant subtypes. Comparative statements about groups and individuals require meaningful criteria and precise definition.

A related problem stems from the practice of combining various subtypes of sexual deviates into a single group. For example, Barbaree and Marshall (1988) conducted a recidivism study with nonfamilial child molesters who had offended against male and female. This study did not control for possible differences between homo- and heterosexual groups (see Groth and Birnbaum, 1978). The practice of combining deviant subtypes in a sample threatens external validity because it obscures differential treatment responses and reoffense patterns. Follow-up research cannot accurately represent the prognostic value of deviant arousal with respect to future functioning without recognizing the heterogeneity of sexual deviates both within and between broad diagnostic categories (Marshall and Barbaree, 1988; Quinsey et al., 1979).

One other selection-related issue deserves mention. Laboratory research subjects who exhibit no arousal are often dropped from the sample. For example, in an early study on faking, Laws and Rubin (1969) dropped three of seven subjects. More recently, Barbaree and Marshall (1989) excluded 11 extrafamilial child molesters and 11 incest offenders from their original sample because these subjects did not show at least 20% of full erection to age-graded stimuli. The practice of excluding nonresponders
compounds the bias associated with the use of heterogeneous opportunity samples and compromises both internal and external validity.

**Procedural and Contextual Confounds**

Whether deviant response patterns are observed is determined not only by the complex stimulus materials used to elicit arousal, but by a variety of environmental contingencies. Test results may be influenced by the clinician conducting the evaluation, the setting, instructional set, and the circumstances leading to assessment or treatment. In addition, the impact of situational variables is likely to be potentiated by a variety of individual differences variables. For example, some subjects may respond with arousal to certain stimuli only under certain circumstances; others may be much less influenced by situational factors. Little is known about the nature of such confounds or their effect on phallometric findings.

**CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH**

This review has focused on theoretical and methodological issues relevant to the validity of plethysmography in the assessment and treatment of sexual deviance. These issues are difficult to resolve, but can be discussed constructively in terms of the directions they offer for future research in this area.

True conclusions may be derived from false premises, but valid deductions require a valid basis. The use of phallometric findings for important clinical and legal decisions and scientific inquiry should reflect a full appreciation of the measurement technique and the assumptions underlying its use. This becomes possible given adequate empirical support and clear explication of general principles.

The validity and clinical utility of plethysmography in the assessment and treatment of sexual deviance remain to be established. At the same time, basic questions relating to the significance of physiological aspects of the human sexual response must be addressed. Future work in this area should focus not only on the appropriateness of plethysmography in typical clinical and research applications with forensic populations but also should respond to the need for empirical and theoretical knowledge about human sexuality in general.

The relationship of physiological arousal to other dimensions of sexuality must be further clarified. This calls for a better understanding of
the variables that affect phallometric findings. Failures of replication and discrimination in this area may reflect important differences between studies in terms of the apparatus used (volume vs. circumference transducers), test stimuli, administration and scoring techniques, and sample selection bias. The effects of relevant and extraneous sources of variation must be determined to provide a basis for the development of practical guidelines and minimal standards of technical adequacy for phallometric testing. But attention to methodological considerations alone cannot ensure the integrity and usefulness of a scientific paradigm. Theory and research must be mutually complementary. Thus the theoretical ideas upon which a paradigm is based must also be evaluated.

Crawford (1979) observed, “Although it is too simple to view sexual deviance as consisting only of sexual arousal to deviant stimuli, it would be premature to suggest that deviant arousal is irrelevant or unrelated to the problem” (p. 155). Likewise, if sexual deviates do not uniformly become aroused in the presence of deviant stimuli, this does not mean that they are not motivated by deviant sexual interests. Thus more precise inferences about individual differences in sexual preference require other levels of analysis. An integrated approach to the assessment of sexual preferences will take into account aspects of an individual’s functioning other than those of a purely sexual nature, as well as dimensions of an individual’s sexuality other than genital arousal.

REFERENCES


THE POLLS—A REPORT
AIDS

ELEANOR SINGER
THERESA F. ROGERS AND
MARY CORCORAN

In 1981, a new, mysterious, and deadly disease known as AIDS was identified. At present there is no cure for it and no vaccine to prevent its transmission. Efforts to control the spread of AIDS are hampered by the long latency between infection and symptoms (an average of at least four years) and by the fact that once an individual is exposed to the virus presumed to cause AIDS, he or she can infect others without knowing it. The disease has been termed a pandemic; by 1986, 1.5 to 2 million people in the United States were believed to have been exposed to it.

To the best of our knowledge, the first survey of public opinion about AIDS was carried out in June 1983. Between then and November 1986, the cutoff date for this article, we identified 22 surveys on this topic. Four were carried out in 1983, nine in 1985, and another nine in 1986. Almost all have been nationwide, and they represent more than 200 questions that the public has answered about AIDS—what they know about the disease, how vulnerable they think they are, what they think the government should do, and what, if anything, they are doing to protect themselves. These data provide the best gauge we have of the public’s knowledge about and reaction to AIDS and how their concerns and views may be changing.

In selecting questions for this article, our main criterion was the availability of trend data. In some cases, however, we have reported responses to items for which there are as yet no trends, because we believe they are likely to be asked again in the future.

We have organized this summary of the poll results in terms of six dimensions: awareness of the disease, concern about it both as a seri-