Journal of Consulting and Clinical Psychology
1989, Vol. 57, No. 1, 100-105

0022-006X/89/$00.75

Phallometric Diagnosis of Pedophilia

Kurt Freund and Ray Blanchard
Clarke Institute of Psychiatry
Toronto, Ontario, Canada

We investigated the sensitivity and specificity of our phallometric test for pedophilia (and hebephilia). An initial sample of subjects included 47 men accused of sexual offenses against minors and 26 control subjects—men accused of offenses against adult women (exhibitionism, rape, or sexually sadistic activity). A second sample included 107 offenders against minors and 30 control subjects. In both samples, the offenders against minors were further classified according to the targets of their sexual offenses (girls, boys, or both) and according to the extent to which they admitted an erotic preference for the immature physique. Computerized diagnostic rules were developed with the first sample and cross-validated with the second. The sensitivity of the test in detecting pedophilia or hebephilia in complete nonadmitters is probably greater than or equal to 55% but is certainly less than 100%. Its specificity appears to be over 95%.

Pedophilia is the erotic preference for prepubescent children; *hebephilia,* the preference for pubescent children. Both pedophilia and hebephilia may be further qualified as heterosexual or homosexual according to the sex of child preferred. In pedophilia, sex-of-partner preference may also be undifferentiated (Mohr, Turner, & Jerry, 1964). *Gynephilia* is the erotic preference for physically mature women; *androphilia,* the preference for physically mature men. Pedophilic and possibly also hebephilic preferences are almost exclusively found in men.

Pedophilia (or hebephilia) cannot always be diagnosed from a patient's history. A man's sexual approach to minors may only be a substitute object choice in an erotically normal but socially inadequate individual or it may reflect a positive preference for the immature physique. In the former case, only counseling, psychotherapy, or behavior therapy are indicated; in the latter, concurrent treatment with sex-drive-reducing medications may be desirable (Freund, 1980). Therefore, the differential diagnosis of pedophilia or hebephilia versus surrogate object choice is of great importance.

Self-report is an unreliable basis for the differential diagnosis of pedophilia versus surrogate choice because the large majority of patients who prefer the immature physique deny it.

The serious limitations of diagnosing from history or self-reports necessitate an alternative approach to the assessment of erotic age-preference. The most promising candidate so far is the phallometric method (Freund, Diamant, & Pinkava, 1958). In this procedure, penile blood volume, the dependent measure of erotic arousal, is monitored during the presentation of potentially erotic test stimuli. Phallometric tests for erotic age-preference have previously been shown to discriminate reliably between groups (Abel, Becker, Murphy, & Flanagan, 1981; Marshall, Barbaree, & Christophe, 1986; Murphy, Haynes, Stalgaitis, & Flanagan, 1986); there is far less research, however, on their usefulness for individual diagnosis.

In recent years, the phallometric method has been used by some diagnosticians as if it were a totally dependable instrument for the diagnosis of pedophilia. This is not the case. Although the phallometric method is much less readily affected by the subject's attitude than is verbal report, it is far from impervious to the influence of the subject's willingness to disclose or recognize his erotic interests. Several previous studies have shown that many subjects can influence the outcome of a phallometric test by suppressing penile responses to stimuli that normally arouse them, inducing tumescence by fantasizing in the presence of stimuli that normally do not, or both (see Freund, Watson, & Rienzo, 1988). When assessing the diagnostic power of the phallometric test of pedophilia or hebephilia, therefore, one must consider the extent to which the examinee denies an erotic preference for minors.

The present study investigated the ability of the phallometric test to distinguish men with sexual offenses against minors from men with sexual offenses against physically mature women. The first phase of the study involved the development of diagnostic rules for diagnosing pedophilia and hebephilia from phallometric test data. The second phase involved the cross-validation of these rules with a new sample of subjects.

Because we were primarily interested in the test's ability to diagnose pedophilia in uncooperative subjects, we looked separately at those offenders against children who acknowledged and those who denied an erotic preference for the immature physique. To compare the latter with control subjects who had an equivalent motivation to appear "normal" on the phallometric test, we used men who were also undergoing clinical assessment for erotic disorders but whose known sexual offenses (and self-reported preferences) involved only physically mature women.

This investigation was supported by Medical Research Council of Canada Grant MA-9134 to Kurt Freund and Ray Blanchard. We wish to thank Cathy Spegg, George Whiting, Robin Watson, and Douglas Rienzo for computational, programming, and technical assistance. Barry Smith and Ted Harris-Brandts designed and built the interface between the pressure transducer and the mainframe computer and automated the stimulus presentation.

Correspondence concerning this article should be addressed to Kurt Freund, Research Section of Behavioral Sexology, Clarke Institute of Psychiatry, 250 College Street, Toronto, Ontario, Canada M5T 1R8.

Method

Subjects

Subjects were nonpsychotic male outpatients of the sexology department of a psychiatric teaching hospital. All were between the ages of 18 and 55 years (inclusive). The sexual offenders against children were men accused of such activities, with or without the formal laying of charges. The control subjects (offenders against women) were men charged for exhibitionism, rape, or sexually sadistic activity. They were routinely tested as part of their clinical assessments to determine whether their erotic disorders might also include pedophilic tendencies, and they were so informed. We excluded 32 men from the control group (and from the study) because their charges also included offenses against minors.

There were 210 subjects in total. The initial sample included 47 offenders against children and 26 control subjects; the cross-validation sample included 107 offenders against children and 30 control subjects. The mean age of the offenders against children was 34.0 years (SD = 10.1), and their median and modal educational level—including or excluding the mentally handicapped—was 1 or more years of high school without graduation. The mean age of the control subjects was 27.6 years (SD = 7.1), and their median and modal educational level was also 1 or more years of high school without graduation. The racial composition of the offenders against children was 95.5% White, and that of the control subjects was 98.2% White.

The extent to which the subject admitted an erotic preference for minors was assessed by means of two items in Freund's unpublished Erotic Preferences Examination Scheme (EPES), a questionnaire routinely administered to all male patients examined for erotic disorders in this sexology department:

1. When lying in bed and sexually aroused, do you imagine or fantasize about intimately touching *a* boys or girls up to 10 years old, *b* boys or girls 12–15 years old, *c* boys or girls in both age ranges, *d* none of the above.

2. When lying in bed and sexually aroused, do you imagine or fantasize about intimately touching females 17 and over more often than touching children or teenagers 15 years or younger? *a* yes, *b* no.

Those patients who endorsed option *d* of Item 1 and option *a* of Item 2 were classified as nonadmitters. Those who endorsed only one of these two options were classified as partial admitters. Those few individuals who endorsed option *a*, *b*, or *c* of Item 1 and option *b* of Item 2 were classified as full admitters.[1]

Nonadmitting offenders against male minors occasionally claim to be androphilic rather than gynephilic. In this study, therefore, offenders against boys were administered an additional section of the EPES that included an item asking the preferred age of male sexual partners. An individual who indicated on this additional item that he preferred male partners of age 17 or older was considered to have claimed androphilia rather than gynephilia for his erotic preference, a distinction with later implications for assessing the validity of his phallometric test outcome.

Subjects who were of borderline intelligence or who were mentally retarded were grouped separately on the supposition that they would not be able to fake the test outcome as well as others. In all cases, the diagnosis of mental handicap was made prior to the patient's presentation at our department. None of these individuals was able to read or write.

Apparatus and Procedure

The phallometric hardware included a penile volume sensor, a pressure-to-voltage transducer, and mainframe and microcomputer interfaces. Output for individual subjects included penile response tracings (recreated by software from the digitized data) and diagnostic printouts summarizing subjects' responses to the various categories of test stimuli.

The test consisted of two sessions, each using a different modality of phallometric stimuli. In Session 1, the stimuli were 14-s film strips of nude children and adults walking slowly toward the camera. Each trial included a sequence of two such strips, both showing individuals of the same age and sex; this was done simply to increase the effective stimulus duration to 28 s. Nine categories of stimulus objects were used: physically mature women, early adolescent girls, 8–11-year-old girls, and 5–8-year-old girls; four corresponding male age categories; and erotically neutral stimuli (landscapes). The test consisted of three blocks of nine trials. Each block included one trial with each type of test stimulus in fixed pseudorandom order.

This procedure was varied slightly for the cross-validation sample. For these cases, the film strips were accompanied by audiotaped narratives, presented through headphones, that described (nonsexual) activities of persons of the same age and sex as those shown on the screen.[2] Each narrative consisted of approximately 60 words and lasted 22–25 s.

The stimuli in Session 2 consisted of audiotaped narratives presented through headphones and accompanied by slides. There were five categories of narratives that described explicitly sexual interactions with men, women, prepubescent boys, and prepubescent girls as well as solitary and nonsexual activities (neutral stimuli).[3] All narratives were written in the second person and present tense ("You are babysitting your neighbors' little girl for the evening"). The narratives describing heterosexual interactions were recorded with a woman's voice, and those describing homosexual interactions were recorded with a man's voice. Neutral stimuli were recorded with both. Each narrative was approximately 100 words long and took the actor or actress about 46 s to read; penile volume recording, however, always continued for 54 s, the fixed length of a trial. Each narrative was accompanied by three successive color slides showing the front view, rear view, and genital region of an individual corresponding in age and sex to the topic of the narrative; neutral narratives were accompanied by slides of landscapes. The test consisted of five blocks of five trials, and each block included one trial with each type of narrative in fixed pseudorandom order.

The procedures used in Session 2 were modified slightly for the cross-validation sample. For these subjects, we used three projection screens in a semicircular arrangement to fill the subject's visual field completely with pictorial stimuli. Slides showing different views of the same individual were simultaneously projected on all three screens. Three different stimulus persons were still shown during each trial so that nine slides were presented altogether.

Session 1 was preceded by three warm-up trials, and Session 2 was preceded by two. The warm-up trials followed exactly the same format as trials in the test proper; their purpose was to habituate the subject to

---

[1] Offenders who could not complete the EPES, either because of insufficient English language skills or because of mental deficiency, were classified as nonadmitters or full admitters by interview. The partial admitter classification could not be used in these cases.

[2] The following sample narrative accompanied film strips of pubescent boys: "Bill is 12 years old. His best friend, Peter, is sleeping over tonight. After Bill's parents have gone to bed, the boys decide to go skinny-dipping in the pool in the back yard. They sneak outside, quickly taking off their clothes, and jump in the water. The water feels hot compared to the cool night air."

[3] Written copies of the stimulus narratives are available upon request. Because we have assured our photographic models (or their parents) that their images would be used only within the Clarke Institute, we do not offer copies of slides or movies. Professionals interested in developing their own stimulus sets are welcome to visit the Clarke Institute to inspect any visual materials.

Table 1
Number of Subjects, Uninformative Outcomes, and Partially Informative Outcomes in Each Admitter Category

Admitter status	*n*	Uninformative diagnoses: Low OI	Uninformative diagnoses: Uncertain age preference	Uninformative diagnoses: Effort[a]	Partially informative diagnoses
Nonadmitter					
Sample 1	32	3	1	5	1
Sample 2	49	2	4	3	0
Partial admitter					
Sample 1	6	0	0	0	0
Sample 2	32	0	1	2	0
Full admitter					
Sample 2	16	0	1	0	0
Retarded					
Sample 1	9	0	0	0	2
Sample 2	10	1	0	0	1
Control					
Sample 1	26	2	1	7	0
Sample 2	30	1	2	5	0

Note. OI = output index.
[a] Diagnosis invalidated by signs of subject's deliberate effort to influence test outcome.

the laboratory situation. The onset of each trial in the test was delayed until penile blood volume had returned to its baseline value (or was oscillating closely around that value). To hasten detumescence during the intertrial interval, subjects were asked to read aloud from slide projections of nonerotic texts; illiterate subjects were engaged in casual conversation. Low-light cameras and closed-circuit television were used to monitor subjects' movements (e.g., turning their heads away from the projection area and closing their eyes). Both test sessions typically lasted a little over 1 hr. They were conducted on the same day, with a rest period for the subject provided in between.

Results

Phallometric test results in which the degree of penile blood volume change was so small that it could be attributed to random physiological "noise" were excluded as invalid. The criterion variable used for this decision was the output index (OI; Freund, 1967). The OI for a test session equals the mean of the three largest blood volume increases that occur in response to any stimulus category or categories except neutral. For the results of Session 1 to be considered valid, the OI for the session had to be greater than or equal to 0.5 ml. The criterion for Session 2 (in which responses were generally larger) was 1.0 ml. If a subject's OI was too low for both test sessions, he was excluded from further diagnosis.[4] The number of subjects for whom this occurred is presented in Table 1.

The sequence of computations and rule-based decisions leading from a subject's raw phallometric data to his final diagnosis is summarized in the following.[5] The diagnostic procedure consisted of three phases: (a) conversion and condensation of raw data, (b) computation of erotic preference indexes, and (c) rule-based diagnosis.

Basic Diagnostic Procedure

Conversion and condensation of raw data. The purpose of this phase was to convert the penile volume changes observed during the individual trials of a session into a smaller set of scores reflecting the subject's relative response to the various stimulus categories presented during the session. The measure used (a mean composite *z* score) took latency as well as amplitude of response into account and facilitated between-subjects comparisons of relative response. The end product of this phase was, for each subject, a set of nine scores for Session 1 (one for each stimulus category) and of five scores for Session 2.

Erotic preference indexes. The mean scores were next used to compute two types of erotic preference index. *Sex-preference indexes* were measures of the subject's relative tendency to respond to male versus female targets. *Age-preference indexes* were measures of the subject's relative tendency to respond to adults versus children or to adults versus pubescents. Erotic preference indexes were computed separately for each test session. These indexes were then used as input for the diagnostic procedure described next.

Rule-based diagnosis. The development of objective rules for the phallometric diagnosis of pedophilia and hebephilia had two aspects. The first was the identification of cutting scores for the erotic preference indexes, which were used to assign subjects to discrete diagnostic categories (e.g., heterosexual versus homosexual). The second was the development of a hierarchical decision tree to direct the flow of successive categorizations.[6]

The diagnostic rules (cutting scores plus decision tree) were embodied in the Phallometric Expert System (PES), a computer program designed to emulate Freund's clinical decision-making process in diagnosing pedophilia or hebephilia from phallometric test results. An informal trial-and-error procedure was used to develop cutting scores for the age- and sex-preference indexes that, when "plugged into" the PES, would maximize the number of offenders against children classified as pedophilic or hebephilic while minimizing the number of control subjects so classified.

The cutting scores were located, and the PES diagnostic rules were finalized, using the initial sample of subjects. These rules were then applied, with no further changes, to the cross-validation sample.

As previously noted, a full listing of the PES diagnostic rules is available upon request. The following discussion, however, requires that two specific points be explained. First, if the subject's age-preference indexes showed neither a clear preference for children nor a clear preference for adults, then his age-preference diagnosis was labeled *uncertain*. Second, the phallometric method appears presently unable to diagnose heterosexual

[4] If the subject's OI was valid for Session 1 only, the diagnostic rules were essentially unaltered. If his OI was valid for Session 2 only, however, he could not be diagnosed as hebephilic because Session 2 included no stimulus category of pubescents.

[5] Readers interested in the specifics should write to the senior author for a copy of the technical supplement to the present report.

[6] The operation of the decision tree can be illustrated with a simple example. If the subject's age-preference index for either session fell beyond the critical value for pedophilia, then his final diagnosis was pedophilia and his sex preference was not assessed. If he was not diagnosed as pedophilic, then his sex-preference index was evaluated, and he progressed to further decisions according to his classification as heterosexual or homosexual.

hebephilia. The PES, therefore, yields no such diagnosis, and offenders who approached only 11–16-year-old girls were not included in the study.

Internal Validity Indicators

Phallometric test outcomes were evaluated as uninformative, partially informative, or fully informative. Three types of outcome were evaluated as uninformative: (a) The OI was too low on both test sessions to make a diagnosis, (b) the subject's erotic age-preference was uncertain and his phallometrically diagnosed sex-preference was the same as his verbal claim, and (c) a "socially acceptable" diagnosis of androphilia or gynephilia was invalidated by signs that the subject had made a deliberate effort to influence the outcome of the test (Freund et al., 1988). The partially informative category pertains to patients whose known offenses involved male minors and who verbally claimed an erotic preference for adult women rather than adult men. In these cases, the diagnosis was considered partially informative if it was positive for homosexuality without being positive for pedophilia or hebephilia.

Diagnosis After Validity Assessment

Table 1 shows the number of subjects in each admitter category, the number of each type of uninformative test outcome in each admitter category, and the number of partially informative test outcomes in each admitter category. There were a total of 41 uninformative outcomes and 4 partly informative outcomes.

The remaining 165 test outcomes were considered fully informative. These are presented in Table 2, which shows the number of subjects phallometrically diagnosed as pedophilic or hebephilic in each admitter category.[7] In this table, each admitter category is further subdivided into three groups according to the sexual targets in the subjects's known offense history: (a) at least one girl younger than 11 years (and no boys), (b) at least one boy younger than 17 years (and no girls), and (c) at least one boy and one girl younger than 17 years.

Table 2 suggests that admitters are more likely to be diagnosed as sexually anomalous than nonadmitters and that offenders against boys are more likely to be diagnosed as anomalous than offenders against girls. These trends were tested using backward logistic regression. In this analysis, offenders against boys and girls were combined with offenders against boys only, and full admitters were combined with retardates. Control subjects were not included in the analysis. The results confirmed that the proportion of pedophilic or hebephilic diagnoses was related to admitter status, F-to-remove $(2, 121) = 4.93$, $p < .01$, and to offense history, F-to-remove $(1, 122) = 5.15$, $p < .05$. The interaction between admitter status and offense history was not significant, F-to-enter $(2, 121) = 0.44$. It should be noted that the finding regarding offense history is difficult to interpret because we did not include men who offended only against 11–16-year-old girls in the study and because the PES does not offer a diagnosis of heterosexual hebephilia.

The main point of this study was to determine the sensitivity and specificity of the phallometric test with subjects motivated to avoid a diagnosis of pedophilia or hebephilia. Of the 40 nonadmitters in the cross-validation sample, 55% were diagnosed as pedophilic or hebephilic. For reasons discussed later, this figure may be taken as a minimum estimate of the test's sensitivity. Of the 22 control subjects in the cross-validation sample, 95% were diagnosed as gynephilic. This figure may be taken as the best estimate of the test's specificity with motivated subjects.

Table 2
Number of Cases Phallometrically Diagnosed as Pedophilic or Hebephilic

Admitter status	Targets of sexual offenses: Girls	Boys	Girls and boys	Adult women
Nonadmitter				
Sample 1				
Ped/Heb	3	8	3	—
Total	7	11	4	0
Sample 2				
Ped/Heb	5	14	3	—
Total	14	23	3	0
Partial admitter				
Sample 1				
Ped/Heb	3	2	—	—
Total	4	2	0	0
Sample 2				
Ped/Heb	7	13	2	—
Total	11	15	3	0
Full admitter[a]				
Sample 2				
Ped/Heb	4	9	2	—
Total	4	9	2	0
Retarded				
Sample 1				
Ped/Heb	2	4	1	—
Total	2	4	1	0
Sample 2				
Ped/Heb	1	5	1	—
Total	1	6	1	0
Control				
Sample 1				
Ped/Heb	—	—	—	0
Total	0	0	0	16
Sample 2				
Ped/Heb	—	—	—	1
Total	0	0	0	22

Note. Ped/Heb = pedophilic/hebephilic.
[a] Full admitters were not used to develop the diagnostic rules. Therefore, all such cases were put into Sample 2.

Discussion

Only 1 control subject in the cross-validation sample was diagnosed as pedophilic or hebephilic. It thus appears safe to conclude that nearly all offenders against minors who were phallometrically diagnosed as pedophilic or hebephilic were diag-

[7] During the course of this study, an additional 5 patients presented at our department, complaining of sexual attraction to female children and requesting treatment for this problem. Because they had never, to our knowledge, been accused of or charged with any sexual offense, they were not included in the study proper. In all 5 cases, the PES diagnosed pedophilia.

nosed correctly. The high specificity of the test implies that a positive phallometric diagnosis of pedophilia may be the decisive factor in clinical assessment when the patient's age preference cannot be determined with confidence from his history.

The sensitivity of this test, for reasons discussed next, cannot be completely determined from the present data. Its sensitivity with nonadmitters, however, is probably much lower than its specificity. Negative diagnoses should therefore be regarded as less informative than positive diagnoses.

To point estimate the sensitivity of the test for nonadmitters, we would need to divide the number of positive diagnoses by the number of true pedophiles. We do not, however, know the number of true pedophiles in the nonadmitter group. (If one could reliably make this determination, one would not need the phallometric test in the first place.) We only know the total number of nonadmitters, and it is most likely that this group included some "surrogate offenders" who were not truly pedophilic or hebephilic but rather who approached children as substitutes for adult sexual partners. It is, however, possible to estimate the minimum sensitivity of the test if one makes the assumption (reasonable, from the control subjects' data) that there were virtually no false positives in this group.

If there were no surrogate offenders in the nonadmitter group, then the number of pedophiles would equal the total number of cases, and the number of positive diagnoses (22) divided by the total number of cases (40) would equal the sensitivity of the test (22/40 = 55%). If there were any surrogate offenders among the nonadmitters (e.g., 7), then the correct denominator for computing sensitivity would be less than the total number of cases (40 − 7 = 33), and the same number of positive diagnoses would necessarily represent a greater proportion of correct identifications (22/33 = 67%). It can therefore be seen that the observed proportion of positive diagnoses (55%) estimates a lower bound for the true sensitivity of the test.

Nothing in our data rules out the possibility that the true sensitivity of the test is as low as 55% or even as high as 100%. (It is theoretically possible that the test identified every true pedophile in the nonadmitter group.) Other considerations, however, make it unlikely that the test's true sensitivity is as high or low as these extremes.

It is likely that the test's sensitivity is greater than 55% because it is likely that the nonadmitter group included at least some (nonpedophilic) surrogate offenders. To minimize such heterogeneity, we did exclude two types of offenders who often seem to interact sexually with female children without having an actual preference for the immature physique: incest offenders (Quinsey, Chaplin, & Carrigan, 1979) and exhibitionists who expose to children as well as adults (Freund & Blanchard, 1986). It remains, nonetheless, highly probable that at least the offenders against female children included a significant admixture of nonpedophiles; this can be inferred from the relatively low recidivism rate observed for this type of offense (Abel, Becker, Mittelman, & Cunningham-Rather, 1986; Gebhard, Gagnon, Pomeroy, & Christenson, 1965; Mohr et al., 1964; Quinsey, 1977).

Certain other factors might also have spuriously depressed the apparent sensitivity of the test. For example, some of the alleged offenders against children may have been falsely accused. It must also be recognized that all patients who undergo phallometric testing sign a consent to do so and are therefore a self-selected population. It is possible that patients who know themselves to be truly normal in age preference are overrepresented among those who consent to testing.

On the other hand, it is well-established that age-of-partner preference is relatively easy to fake on the phallometric test (Freund et al., 1988). Thus, it is virtually certain that the true sensitivity of the test, whatever it may be, falls well short of 100%.

A final point regarding sensitivity and specificity is in order. Our diagnostic rules were deliberately adjusted (using Sample 1) so that no control subject was diagnosed as pedophilic or hebephilic. To achieve this low rate of false positives, we had to accept an increased number of negative (and probably false) diagnoses for the offenders against children. We did this in the belief that falsely labeling a man as pedophilic is more unacceptable than failing to identify a true pedophile.

An additional limitation of phallometric testing is that it does not always yield a clear diagnosis. In this study, about 21% of patients were not diagnosed: Some responded insufficiently, some failed to show a clear age preference, and some showed signs of attempting to manipulate the test outcome.

The 12 control subjects in Table 1 whose phallometric diagnoses were invalidated by the last-mentioned criterion require some explanation. What these signs (called "faking" signs by Freund et al., 1988) actually indicate is that the subject made some effort to influence the outcome of the test as opposed to simply responding in a relaxed and passive manner.

In the phallometric investigation of men accused of sexual approaches to children, these signs invalidate a socially acceptable diagnosis of gynephilia or androphilia because they suggest that the subject has made some effort to obtain this diagnosis. We applied the same diagnostic rule to our control subjects purely for the sake of consistency in this study. In the case of our control subjects, however, the presence of these signs was very unlikely to mean that the subject needed to fake to obtain a diagnosis of gynephilia. What these signs meant among our control subjects was that the subject (under investigation for sexual offenses, albeit offenses involving women) was strongly motivated to "help" the test produce a normal diagnosis.

We conclude this article with three caveats regarding the inappropriate use of the phallometric test for pedophilia. First, this test was developed for the differential diagnosis of men charged with sexual offenses and might be less valid with subjects who agree to testing for some other reason. A paid volunteer control, for example, might be more prone to allow his mind to wander back to previous stimuli involving women during (uninteresting) stimulus presentations involving children and, thus, to produce a false diagnosis of pedophilia. Thus, the test should not be used without appropriate validity studies to survey the general population. Second, a phallometric diagnosis of pedophilia or hebephilia should not be interpreted as evidence that a man has committed the specific offense (or offenses) for which he has been charged.

Our third point concerns the appropriate qualifications for users of the phallometric test. The limitations of phallometric testing in general necessitate caution when this method is used in making differential diagnoses with sex offenders. The results of clinical diagnostic phallometric tests should be exclusively

available to and interpreted by behavioral sexologists with medical or psychological degrees—professionals reasonably acquainted with scientific and ethical standards and accountable to the governing bodies of their professional organizations. We do not condone the use of diagnostic phallometric tests by narrowly informed technical experts or their application in the same manner in which lie detector procedures are sometimes administered.

References

Abel, G. G., Becker, J. V., Mittelman, M. S., & Cunningham-Rather, J. (1986, February). *The self-reported molestations of non-incarcerated child molesters.* Paper presented at a meeting on the Assessment and Treatment of Sex Offenders, sponsored by the National Institute of Mental Health and the Florida Mental Health Institute, Tampa, FL.

Abel, G. G., Becker, J. V., Murphy, W. D., & Flanagan, B. (1981). Identifying dangerous child molesters. In R. B. Stuart (Ed.), *Violent behavior: Social learning approaches to prediction, management, and treatment* (pp. 116–137). New York: Brunner/Mazel.

Freund, K. (1967). Diagnosing homo- or heterosexuality and erotic age-preference by means of a psychophysiological test. *Behaviour Research and Therapy, 5,* 209–228.

Freund, K. (1980). Therapeutic sex drive reduction. *Acta Psychiatrica Scandinavica, 63*(Suppl. No. 287), 1–38.

Freund, K., & Blanchard, R. (1986). The concept of courtship disorder. *Journal of Sex and Marital Therapy, 12,* 79–92.

Freund, K., Diamant, J., & Pinkava, V. (1958). On the validity and reliability of the phalloplethysmographic diagnosis of some sexual deviations. *Review of Czechoslovak Medicine, 4,* 145–151.

Freund, K., Watson, R., & Rienzo, D. (1988). Signs of feigning in the phallometric test. *Behaviour Research and Therapy, 26,* 105–112.

Gebhard, P. H., Gagnon, J. H., Pomeroy, W. B., & Christenson, C. V. (1965). *Sex offenders.* New York: Harper & Row.

Marshall, W. L., Barbaree, H. E., & Christophe, D. (1986). Sexual offenders against female children: Sexual preferences for age of victims and type of behaviour. *Canadian Journal of Behavioural Science, 18,* 424–439.

Mohr, J. W., Turner, R. E., & Jerry, M. B. (1964). *Pedophilia and exhibitionism.* Toronto, Ontario, Canada: University of Toronto Press.

Murphy, W. D., Haynes, M. R., Stalgaitis, S. J., & Flanagan, B. (1986). Differential sexual responding among four groups of sexual offenders against children. *Journal of Psychopathology and Behavioral Assessment, 8,* 339–353.

Quinsey, V. L. (1977). The assessment and treatment of child molesters: A review. *Canadian Psychological Review, 18,* 204–220.

Quinsey, V. L., Chaplin, T. C., & Carrigan, W. F. (1979). Sexual preferences among incestuous and nonincestuous child molesters. *Behavior Therapy, 10,* 562–565.

Received February 24, 1988
Revision received June 7, 1988
Accepted July 7, 1988 ■