
DETECTING FAKED PENILE
RESPONSES TO EROTIC STIMULI:
A COMPARISON OF STIMULUS CONDITIONS
AND RESPONSE MEASURES

Robert D. Card, PhD
Clinic for Counseling & Psychotherapy, Inc.
Salt Lake City, Utah

&

William Farrall, B.S.
Farrall Instruments, Inc.
Grand Island, Nebraska

Annals of Sex Research

ABSTRACT

Eighteen men were instructed to fake penile responses to discrete sexual stimuli presented auditorily or visually. The same stimuli were presented a second time following instruction on faking. Sexual responses were monitored using a penile plethysmograph (PPG). Further physiological response measures to detect faking attempts included the Galvanic Skin Response (GSR) and respiration (RESP). Faking attempts were found to be more successful to audio than visual stimuli, especially if the stimuli were relatively weaker. Arousal was much more difficult to fake than suppression. The more intense the efforts to fake, and especially following faking instructions, the more transparent the faking efforts. GSR and RESP added measurably to the detection of faked PPG responses which would have appeared otherwise to be genuine responses.

Card, R. D., & Farrall, W. (1990). Detecting faked responses to erotic stimuli: A comparison of stimulus conditions and response measures. *Annals of Sex Research*, 3, 381-396.

With the clinical application of penile plethysmography in the early 1960's, it was naively believed that somehow the plethysmograph would invariably measure an individual's sexual response patterns. With the passage of time and with the introduction of the plethysmograph into the legal system where the outcome of an evaluation might result in favorable or unfavorable consequences for the individual being tested, examiners began to notice deliberate attempts on the part of some individuals to manipulate the tests to their advantage (Freund, 1971). To make the tests less susceptible to distortion, examiners developed a number of techniques to discourage or detect faking when it was being attempted. Plethysmographic testing was performed while observing the client through a one-way mirror to see if he was attending to the visual stimuli, or to see if he was using his hands to manipulate the apparatus in order to produce a "normal" response by self-stimulation. To make it more difficult to fake reactions to audiotapes, the individual sometimes was required to repeat back to the examiner the content of the audiotape immediately after it was administered to him (Quinsey & Bergerson, 1976).

With increasing use of the plethysmograph, detecting efforts to fake have become an even more important issue. Clients, intent on faking, in addition to using any or all of the faking methods described above, may masturbate just prior to being evaluated, and may perform mental exercises to prevent or, to produce arousal (Quinsey & Bergerson, 1976). A number of studies have clearly demonstrated that a sizeable proportion of normal adults are able to suppress penile responses to sexual stimuli, while a relatively smaller number are able to voluntarily produce an enhanced erectile response (Freund, 1971; Hall, Proctor & Nelson, 1988; Henson & Rubin, 1971; Rosen, 1973; Quinsey & Bergerson, 1976).

The most common way to deal with faking is to make it as difficult as possible for the client to fake during the assessment period (Freund, 1971). For example, Quinsey & Chaplin (1987) noted that a tracking task given during the presentation of audio stimuli made it very difficult to fake. Certainly, the preferred method to deal with faking is to eliminate as many ways to do it as possible from the outset of testing. Nevertheless, attempts at faking will persist and methods to detect such attempts to fake should be addressed.

For the purpose of this paper, faking is defined as:

an attempt to manipulate sexual response measures to stimulus materials in order to present a response profile which is at variance with actual sexual response patterns.

To examine a test record for possible faking, ideally one wants to know:

1. The legal position of the person being tested.
 - a. Does he deny that the alleged offense occurred at all? We suspect that, with complete denial, the attempt to fake or manipulate the test data will be greatest.
 - b. If the client admits some contact with the victim or victims, how honest and open is he about the offense? In our experience, most sex offenders are motivated to minimize their sexual involvement with the victim(s), and try to present themselves as being as sexually "normal" as possible. We would, therefore, expect such men to produce minimal responses to deviant stimuli while attempting to enhance or claiming to have stronger sexual response patterns to adult female stimuli.
 - c. Is the client motivated to appear more "normal" following a period of treatment to obtain more favorable disposition or to "prove" he is ready for probation or outpatient treatment? If this were the case, we might see a pattern of faking showing increasingly less sexual response to deviant sexual stimuli, and increasingly greater response to "normal" sexual stimuli where such is not actually the case.
2. One needs to know the age and gender of the alleged or actual victim(s). This permits us to look at responses to certain age/sex categories more closely to examine for faking. The differential response can be quite revealing.
3. One needs to know the characteristics of faked test profiles to be able to compare these characteristics with genuine profiles.

In the past, attempts were made to detect faking by carefully examining the penile plethysmographic (PPG) trace, especially when motivation to fake was clearly present. That was insufficient, especially in the case of minimal response. Then self-reported arousal obtained concomitantly with the PPG was examined in attempts to reveal faking. However, where an intention to fake exists, the client will probably fail to self-report accurately. More importantly, self-report may be a measure of denial.

Having used the Galvanic Skin Response (GSR) in biofeedback for stress management, it was noted that when a client holds his breath we can expect a GSR spike. It was hypothesized that when a client tries to fake he will probably hold his breath during this attempt, producing a telltale GSR spike concurrent with a noticeable drop in the PPG. It was hypothesized further that a respiration monitor would provide additional evidence showing when abdominal muscles

are being used to enhance or suppress a response to stimulus materials.

This study was undertaken to explore the relationship between the penile response (PPG), Galvanic Skin Response (GSR), and Respiration (RESP) to a variety of stimulus conditions and instructions to produce an arousal response at variance with the stimuli being presented.

METHOD

Research Participants

Eighteen men were paid ten dollars each to participate in the experiment. All were ongoing or previous clients of the clinic. No "normals" were solicited. Sixteen of the eighteen men had been previously tested on the plethysmograph. None had been previously exposed during treatment at this clinic to stimulus materials of the type used in the present study. Two participants who had not previously been tested with the PPG were basically homosexual in orientation. The majority of the men were a mixture of homosexual and heterosexual pedophiles, at various stages in treatment. None of the volunteers were currently involved in criminal proceedings.

Apparatus

Equipment used to examine physiological responses included the Farrall Instruments Co. CAT 200 system, plus an outboard CAT 100. The standard Farrall PPG module and GSR module were used. Farrall constructed a respiration module for this experiment. It was decided to sacrifice the self-report module for the respiration module and to use a Likert-type scale for the individual to record his own response patterns at the end of each segment. An event marker was needed so that the individual could indicate by pressing a button when he began faking efforts. It was decided to use a momentary spike on the respiration trace as an event marker. The PPG transducer was a Barlow strain gauge produced by Farrall Instruments.

All standardized testing methods for the penile plethysmograph were used. The participant's lap was covered with a board on which was placed the event button in easy reach of the right hand. The first and third fingers of the left hand were used to collect GSR data. The respiration transducer was placed around the abdomen just beneath the rib cage.

The clients sat in a room in a reclining chair approximately 8 feet away from a 19" color TV screen. They wore a headset with a boom microphone for communication with the examiner who conducted the testing in an adjacent room in

which all of the CAT equipment was installed. Total testing time involved about one hour per client.

Study Design & Procedure

The study was a 2 x 2 x 2 x 2 design illustrated in Table 1. Unfortunately, the lower stimulus values and instructional set were confounded in the testing conditions.

I. Stimulus Type (Audio Vs. Visual)

Visual: Two highly stimulating 2-1/2 minute segments were selected from VHS pornographic movies, one depicting heterosexual activity and one depicting a males-only homosexual encounter. The sound tracks were stripped from the segments leaving only the visual representations.

Auditory: Two 2-1/2 minute audio segments (one homosexual and one heterosexual) were written by the first author as a monologue to parallel approx-

Table 1
Study Design

Independent Variables	Dependent Variables
I. Stimulus Type -Audio -Visual	Penile Response (PPG) Respiration (RESP)
II. Stimulus Novelty Value -Higher (novel stimulus) -Lower (repeated presentation)	Galvanic Skin Response (GSR)
III. Instructional Set -No instructions (naive participant) -Faking Instructions (informed participant)	Self-Report (SR)
IV. Faking Response -Suppress -Enhance	

imately the visual segments. The scripts were then professionally narrated and recorded on the sound track of a VHS cassette, accompanied by a neutral visual background.

II. Stimulus Novelty Value (Higher [Novel] Vs. Lower [repeated] Presentation)

The first stimulus presentation was considered to represent higher erotic stimulus value because of the nature (pornography) and novelty of the stimuli. The second tape, presented immediately following the first tape and in the same testing session, was considered to be a lower intensity stimulus because the stimulus was no longer novel.

Stimulus Presentation Format:

The 2-1/2 minute segments were assembled in a random order on two separate tapes with a space of ten seconds between each segment to allow for stopping and starting the tape and to facilitate interaction time between the examiner and the participant. This interaction time permitted assessment of the stimulus value of the segment and faking direction.

III. Instructional Set (Naive vs. Informed Participant)

The 18 men were instructed to fake their responses to segments of auditory or visual erotic stimulus materials while PPG, RESP, and GSR were monitored. In one condition, participants were to use their own ingenuity to fake. This was compared to faking efforts made after they had received instruction in "faking strategies".

Following the first presentation of the four stimulus segments (higher intensity stimuli), and prior to presentation of the second set of four stimuli (lower intensity or repeated stimuli), the participants were given further instructions about faking. They were told that the second part of the test was also a measure of their ability to "beat" the plethysmograph but with a more informed approach. They were told that some people learn to suppress their arousal by sucking in or tightening the muscles around the penis, or they try to distract themselves by thinking about something else. To produce arousal, some individuals try to energize or lift up the muscles associated with the erectile response while thinking sexual thoughts. They were also told, "You may have your own method to increase or decrease the arousal." They were then informed that all of the other testing conditions remained the same.

Research participants were also asked to press an event marker any time they started or stopped using a particular faking strategy (which might happen more than once in a segment).

IV. Faking Response (Suppress Vs. Enhance)

Between each segment the participants were asked to score a Likert-type scale for the arousal value of that segment and to write out briefly the faking method used.

The instructional set read to the participants asked them to "beat the plethysmograph" by trying to become aroused if the materials were not arousing and to suppress arousal if the stimulus materials were found to be arousing. In order to be able to interpret from the computerized printout whether they tried to suppress or enhance their arousal, the men were asked to rate their self-perceived sexual response to each segment on a 9-point Likert-type scale (-4 extremely unarousing, -3 very unarousing, -2 moderately unarousing, -1 mildly unarousing, 0 neutral, +1 mildly arousing, +2 moderately arousing, +3 very arousing, and +4 extremely arousing). If the self-report score was positive (arousing), they were expected to suppress; if negative (nonarousing), they were expected to attempt to produce an enhanced response. Eight identical Likert scales were arranged on an 8-1/2" x 11" sheet of paper, one for each of the eight stimulus segments. The sheet was taped on the board within easy reach of the right hand, with space on the paper for marking the Likert-type scale indicating degree of perceived arousal to the stimulus materials. Under each Likert-type scale was space to write out faking methods tried during each segment.

The faking response was examined for the interrelationship of the PPG, RESP, and GSR.

At the conclusion of the testing, participants were "debriefed" about their experience with the test. They were asked (1) whether they found it easier to fake to the audio or visual stimuli; (2) what method seemed to work best, and (3) if they thought the instructional set helped them to fake.

RESULTS

The participants' ratings of the erotic value of each stimulus segment was used to group arousal and suppress attempts. A negative value indicated that the participant should try to enhance his arousal, while a positive value indicated he should try to suppress. More than one segment during the testing could be enhanced or suppressed by a single client; however, many of the participants were not consistent with their enhancement or suppression attempts compared to their rating of the segment. Where there was a discrepancy and where arousal or suppression could not be clearly determined, the segments were dropped from the study, resulting in varying numbers of segments evaluated in each of the categories shown in Table 2.

Table 2
Successful Faking of Penile Responses

	Number Trials	Number Successes	Percent Successes	Fisher Exact p
Audio	47	25	53.19	.0614
Visual	53	19	35.85	
Higher Stimulus Novelty	45	17	37.78	.1759
Lower Stimulus Novelty	55	27	49.09	
Suppress	53	30	56.60	.0061
Enhance	47	14	29.79	

Independent Variable I. Audio vs. Visual Stimuli

When debriefed following the testing, every participant said, without hesitation, that it was easier to fake the audio than the visual segments. No one said it was easier to fake because the audio was inherently less stimulating than the visual; generally, they said the audio was easier to fake than the visual because they could more easily focus their eyes and attention on something else in the room, or divert their attention using some simple mental scheme.

To determine if participants were actually able to more successfully fake to audio than visual stimuli, their PPG response to the various segments was examined. Successful enhancement was determined to be an increase at any time during the segment of at least 15%, while successful suppression was rated where the PPG was not allowed to rise more than 15% above baseline. Because the numbers of cases in the eight cells was small, the Fisher Exact probabilities were computed. Only one comparison reached significance (see Table 2).

More successfully suppressed responses overall were produced to audio over visual stimuli. The participants were accurate in saying they were more successful in faking responses to the audio segments but results were marginally nonsignificant ($p < .10$). The basis for the difference found may be in the weaker stimulus value of the audio stimuli themselves, or due to the construction of the stimuli.

It can be reasonably argued that comparisons between the fakability of the audio vs. the visual stimuli cannot be made because one cannot construct exactly parallel audio and visual stimuli. A comparison of the participants' ratings of erotic stimulus value of the audio versus the visual stimuli, and first

showing versus second showing, was made.

Participants found the audio stimuli to be significantly less arousing than the visual at both the novel and repeated stimulus levels ($F = 22.04$, $p < .000$). The stimuli became significantly less arousing after the second exposure, combining self reports to the audio and visual stimuli.

Independent Variable II. Stimulus Novelty Value

Repeating the stimulus aided in faking efforts, but results were not statistically significant. While 37.78% of faking attempts were successful on first presentation (novel stimulus), 49.09% were successful upon its repeated presentation.

Independent Variable III. Instructional Set

The instructional set given to the participant just prior to second exposure to the stimulus materials was confounded with the high/low stimulus value of the stimulus materials (repeated presentation). It is not known the degree to which the instructions assisted in producing a faked response to enhancement or suppression efforts. Nevertheless, in the debriefing period following the testing and in written statements made after each segment, the participants stated they used the instructions and found them helpful in faking. Examination of the individual segment printouts showed where the new insights in faking were being used. The new efforts, however, were easily recognizable as faked.

Independent Variable IV. Suppression vs. Enhancement

Suppression of a response was clearly easier to produce than enhancement. Whereas 56.60% of attempts at suppression were successful, only 29.79% of attempts at enhancing penile responses were successful; a statistically significant difference. Enhancement produced more exaggerated efforts, especially following instructions on how to fake, and consequently such faking efforts would be unlikely to deceive an examiner. Suppression was generally easier to produce, especially if the arousal value of the stimulus was low. Where a genuine response was being suppressed, it frequently showed up on the printout by a PPG "rebound" during the detumescence period.

Interactions of Independent Variables

Suppression (audio vs. visual): Participants exposed to the *visual stimuli* the first time were significantly less able to suppress their arousal than during the second exposure ($p < .032$) (Table 3). Fifty percent of the participants were able to

Table 3
Percent Successful Faking of
Penile Responses in
Stimulus Type, Novelty, and Fake Instruction Subcategories

		Arousal		
		High	Low	Exact p
Audio	Suppression	75.00	100.00	.055
	Enhancement	25.00	15.39	.566
Visual	Suppression	13.33	50.00	.032
	Enhancement	60.00	25.00	.258

Note: Fisher Exact Test was used to compute exact p values.

successfully suppress to the repeated visual stimuli, compared to 13.33% for the novel stimulus.

Referring again to Table 3, six (75%) of the participants were able to suppress their response to the novel audio stimuli. For the repeat audio stimulus and following faking instructions, all fourteen participants were able to suppress their response.

Even though all men were able to successfully suppress their response to the *audio stimuli* during second exposure and following faking suggestions, this difference was not statistically significant, possibly because of the small number of cases in that condition. Results were trending toward significance ($p > 0.055$), suggesting that faking to repeat audio stimulus materials may be more successful than to novel audio stimulus materials.

Self-reported arousal results indicated that participants found repeated presentation of stimuli less arousing than first presentations ($F = 11.67, p < .005$).

Thus, as in the audio versus visual stimulus comparison, participants may have faked more successfully to the repeated stimulus because it was less arousing.

Examining the audio and visual results separately, the major portion of this variance was due to audio stimuli. In other words, the audio stimuli lost significantly more arousal value as a result of a second exposure than the visual stimuli did.

Thus, statements involving comparisons of fakability between audio and visual stimuli, using these stimulus materials, should be regarded with caution. Differences in fakability within a stimulus modality (e.g. visual, high stimulus novelty value vs. visual, low stimulus novelty value) can be interpreted with greater confidence.

Enhancement (audio vs. visual): In the audio mode, only three out of a total of twelve (25%) attempts at enhancement were successful. For the repeat stimulus and following faking instructions, only two out of thirteen (15.39%) attempts to produce an enhanced response were successful.

In the visual mode, six of ten attempts (60%) to produce an enhanced response to novel stimulus materials were successful, while, for the repeat stimuli, three of twelve attempts (25%) were successful. Greater success in enhancement to visual stimuli may have been possible due to the participant being able to transfer stimulus cues from reportedly "unarousing" to "arousing" mental imagery.

From these data, we can expect far less success in faking an enhanced response than a suppressed response. An enhanced response was more successful to visual than audio cues, although not significantly so. Suppression was more successful to audio than visual stimuli; significantly more so if compared to the highly stimulating visual stimuli used in this study.

Novel visual stimuli made successful *suppression* difficult (13%). Repeating the stimulus and adding faking instructions improved faking suppression (50%). *Enhancement* of responses to novel visual stimuli was successful in over 60% of the segments examined, but as the stimulus intensity dropped, even with faking instructions provided, the success rate dropped (25%).

Dependent Variables: PPG, versus GSR and RESP

A total of 108 faking segments were examined to investigate the contribution of the GSR and Respiration measures to the detection of faking. Eight additional segments were found to be scorable in this comparison, and were added to the 100 in Table 2. Faking could be detected using the PPG alone in 49 cases (45.4%) (Table 4). Fifty-nine (54.6%) of the faking attempts would have been undetectable without the aid of the GSR and/or the respiration traces. In all of these cases, the GSR showed a flattened pattern or extreme variability. Both

Table 4
GSR and Respiration Enhancement of Faking Detection
(N = 108)

	Faking detected by PPG Alone	Respiration Assisted	GSR Assisted
Naive	20	21	32
Informed	29	22	27
Totals	49	43	59

the GSR and the respiration traces were helpful in detecting faking attempts in 42 cases. The respiration trace showed great variability when faking was being attempted. Although the GSR was helpful in nearly all of these cases, it was comforting when a corresponding variation was found in the respiration trace.

From the cases examined, there was no significant difference in using the additional measures (GSR and Respiration) to detect faking attempts at enhancement vs. suppression or in naive vs. an informed participant.

Comparing audio and visual faking attempts, 24 audio- and 25 visual-related attempts could be detected by use of the PPG alone, a non-significant difference. Faking was detected on 38 auditory stimuli with the GSR and/or respiration traces, while faking was detected on 37 visual stimuli in this manner. Again, the difference was non-significant. It appears that the use of the GSR and respiration measures are equally helpful in detecting faking whether the stimuli are auditory or visual.

SUMMARY

Based on the analysis of the responses obtained above, the following suggestions for detecting faked responses would appear warranted:

Assuming that a motivation to produce a faked response exists, determine from the client's arrest history which segments the client might want to enhance or suppress. For example, a homosexual pedophile would probably try to suppress responses to male child stimuli and enhance his responses to adult female stimuli. Using the PPG with a concurrent GSR and RESP printout, our data would suggest that the client will have little success in faking an enhanced response to

an adult female stimulus, even if he is successful in suppressing responses to the deviant stimuli.

A close examination of age/gender-related victim segments using the PPG alone will sometimes reveal faking efforts in extreme variability of the PPG trace, PPG below baseline or detumescent rebound. A concurrent GSR and/or RESP trace adds an informative dimension in detecting faking efforts. This is especially true where a very low level PPG response is being examined for faking.

Where previously-tested individuals are motivated to fake to obtain favorable consideration while in a correctional setting and where faking feedback from other clients is available, the problem in detecting faking becomes more difficult. Since production of an enhanced response is still very difficult even under these conditions, a client's claim to have "normal" heterosexual urges is suspect if a suppressed response to the deviant stimuli is detected without a corresponding significant response to the heterosexual stimuli. Again, addition of the GSR and RESP to the PPG trace adds a significant measure of confidence in detecting efforts at faking both enhancement and suppression. It is comforting to note that the greater the motivation and efforts to fake, the more transparent those efforts become.

The participants in the present study *claimed* and *were* more successful in faking to the audio than the visual segments, but the reasons they were able to do so is not clear.

The primary purpose of this study was to investigate the possibility of detecting faking of the PPG through concurrent GSR and RESP measures. Although no consistent interacting relationships between these three measures was observed, faking attempts became increasingly more transparent as the participant increased his efforts to fake, and employed more physical means to try to fake. Enhancement of the PPG was generally more difficult than suppression, necessitating greater efforts, and revealing faking most easily. Without the GSR and RESP measures, over half of the faking efforts would have gone undetected based on the PPG alone.

It was comforting to note as an aside to the primary purpose of the study that high intensity stimuli, especially visual stimuli, are frequently very difficult to control. Using the PPG alone, we would usually be able to recognize faking in such segments, but without additional concurrent physiological measures such as GSR and RESP, we might miss additional crucial data, *especially if the participant has been successful in completely controlling his PPG response.*

DISCUSSION

The data from this study clearly reveals the ability of volunteers to successfully fake their penile response even under the most difficult conditions

(naive participants presented with highly arousing stimulus materials). However, fewer men were able successfully to produce an enhanced response than suppress penile responses. Viewed on a case-by-case basis, it is more alarming to report that under the most favorable conditions (uninformed participants presented with a novel visual stimulus), 18% of our cases were able to enhance their penile response, while in the worst case scenario (informed participants presented with a repeated auditory stimulus) 100% of the participants were able to successfully suppress responses. Confounding the lower stimulus value with faking instructions leaves us not knowing why all participants were successful in suppressing responses to the low novelty audio stimuli following faking instructions. The difficulty in detecting faking using the PPG alone was greatest with repeated stimulus materials, either auditory or visual (Farrall & Card, 1988) or with auditory materials in a test/retest situation where the participant is motivated to suppress his response to deviant stimulus materials. Dr Gene Abel's oft-repeated warning not to use the plethysmograph in the courtroom where the question of guilt or innocence is at stake appears to be well founded. When the results of the PPG are used in sentencing an offender, he may be less motivated to appear innocent, but he will probably still want to appear as normal as possible to obtain a more favorable sentence.

An examination of some of the experimental computer printouts shows that some efforts at faking might go undetected using the PPG alone. Using additional concurrent physiological response measures greatly increases our ability to detect efforts to fake. Respiration irregularity appears to be the most reliable additional measure to detect faking efforts, but the GSR, despite its unreliability, often adds an additional check on suppressive efforts, especially where there is a concurrent PPG drop, a GSR spike, and RESP irregularity.

Having used this experimental configuration of response measures for the past two years (PPG, GSR, RESP), very few plethysmograph evaluations were uninterpretable. Even minimal PPG responses or PPG responses which drop below an assigned resting baseline, when examined with the GSR and RESP measures, often revealed a pattern of suppression to stimuli depicting the age and gender of the victim and/or the coercive efforts of the offender. Faking efforts in such cases were detected by addition of GSR and RESP as effectively as the PPG alone where positive responses were available for examination.

Possibly there are better configurations of concurrent response measures which will aid in detecting faking efforts. Use or non-use of self report as a measure of denial might aid in detecting faking efforts to stimuli related to the sexual offense. Sometimes RESP and GSR measures indicate an emotional response to certain stimuli unrelated to the offense pattern. It remains a problem in interpreting this as distinct from an attempt at faking.

The usefulness of the PPG as a valid measure of sexual response patterns is well documented. However, the PPG used within the judicial system where criminal sanctions are at stake introduces the possibility of faking. We have suggested a means to assist in detecting faking efforts.

REFERENCES

- Farrall, W.R., & Card, R.D. (1988). Evaluation of assessment and treatment of the sexual aggressor. In R.A. Prentky & V.L. Quinsey (eds.). *Human sexual aggression: Current perspectives*, p.528, New York: The New York Academy of Sciences.
- Freund, K. (1971). A note on the use of the phalometric methods of measuring mild sexual arousal in the male. *Behavior Research & Therapy*, 2, 223-228.
- Hall, G.C.N., Proctor, W.C., & Nelson, G.M. (1988). The validity of physiological measures of pedophilic sexual arousal in a sexual offender population. *Journal of Consulting and Clinical Psychology*, 56, 118-122.
- Henson, D.E., & Rubin, H.B. (1971). Instructional control of eroticism. *Journal of Applied Behavior Analysis*, 4, 37-44.
- Quinsey, V.L. & Bergerson, S.G. (1976). Instructional control of penile circumference in assessments of sexual preference. *Behavior Therapy*, 7, 489-493.
- Quinsey, V.L. & Chaplin, T.C. (1987). Preventing faking in phalometric assessments of sexual arousal. Unpublished manuscript.
- Rosen, R.C. (1973). Suppression of penile tumescence by instrumental conditioning. *Psychosomatic Medicine*, 35, 509-514.