Changes in physiological and verbal responses of child molesters during aversion therapy

VERNON L. QUINSEY, SIDNEY G. BERGERSEN,
AND CARY M. STEINMAN
Oak Ridge Division, Mental Health Centre, Penetanguishene, Ontario

ABSTRACT

Ten child molesters were each given 20 classical conditioning type aversion therapy sessions and three assessment (generalization) test sessions (before, after 10, and after 20 treatment sessions). Shock was the unconditional stimulus. Larger skin conductance responses were made to child slides than to adult slides in all sessions. Penile circumference responses indicated a small positive shift in sexual preference in the assessment tests when averaged over subjects. Pre- and post-treatment semantic differential and slide ranking sexual preference measures also showed increases in sexual preference for adults relative to sexual preference for children.

The assessment of therapeutic change among institutionalized child molesters is complex and difficult for theoretical as well as practical reasons (Quinsey, 1973). There have been very few behaviour therapy studies that have focused exclusively on child molesters and most investigators have included them with other sex offenders on the dubious assumption that the response to treatment of different types of sexual offenders is similar. Bancroft (1974) has observed:

There is very little evidence of the efficacy of behaviour modification techniques in this group [child molesters] ... The treatability of this group is of considerable importance as paedophiliac offences incur such heavy penalties. In the author's experience paedophiliac offenders frequently have personalities in which self-deception and deception of others is marked, making treatment extremely difficult to evaluate ... At the moment we lack any obvious prognostic criteria in this group.

Penile response measures currently appear to be the most meaningful index of changes in sexual preference. This measure has been used to assess deviant sexual preferences (Freund, 1967; Quinsey, Steinman, Bergerson, & Holmes, 1975). It has also been employed as a measure of progress in the treatment of a variety of sexual disorders (Bancroft, 1970; Callahan & Leitenberg, 1973; McConaghy, Proctor, & Barr, 1972). Unfortunately, there are certain ambiguities in interpreting penile circumference changes

because of the possibility that a patient could learn to control his penile responses during therapy to bias an assessment (Henson & Rubin, 1971; Laws & Rubin, 1969). Concurrent measures of other physiological responses may clarify the meaning of penile responses during generalization assessment tests or aversion therapy sessions.

Skin conductance has been used by Bancroft and Mathews (1971) as an adjunct to penile circumference measures in the assessment of sexual preferences among normal subjects. Steffy and Gauthier (Note 1) have employed skin conductance responses to monitor the effects of aversion therapy with child molesters. However, measures of skin conductance alone present great difficulties as indicators of changes in sexual preference during aversion therapy (Quinsey, 1973). Verbal measures of sexual preferences have also been found to be sensitive to treatment effects (e.g. Marks & Sartorius, 1968). Unfortunately, verbal measures appear more easily biased than physiological measures (Quinsey et al., 1975).

Each of the physiological and verbal measures described above is potentially valuable in evaluating the effects of aversion therapy. The usefulness of combination of these measures in interpreting aversion therapy effects remains to be explored. The purpose of the present experiment was to compare the changes in a number of measures of sexual preference that were exhibited in generalization tests administered at various times during aversion therapy. An additional goal was to compare skin conductance and penile circumference during a constant number of classical aversive conditioning sessions in which no instrumental behaviour was required of the patient.

METHOD

Ten in-patients of the Mental Health Centre, Penetanguishene, volunteered for aversion therapy. All of these patients admitted to sexual contacts with children and nine of them had been institutionalized for at least one such offence. Each patient had had a sexual encounter involving physical contact with a child 15 years of age or younger when the patient was over 16 years of age and at least 5 years older than the child. Two patients were included who had had such a contact when younger than 16 but who admitted such inappropriate sexual feelings when over 16. The characteristics of these patients are presented in Table 1.

Apparatus

The apparatus and recording procedure have been described elsewhere (Quinsey et al., 1975). Briefly, the subject was seated in a separate room from the apparatus and experimenter. Skin conductance was recorded from the distal volar surface of the first and second finger of the left hand and penile circumference was measured with a mercury in rubber strain gauge which the subject placed on the shaft of his penis.

Shocks from a constant current device were delivered through brass electrodes attached by means of a velcro band to the inside of the subject's left upper arm. Before

TABLE 1

Characteristics of the sample[1]

Patient	Age	Diagnosis	Phenothiazines during treatment	Victim age	Victim sex	Number of times institutionalized for child molesting
1	20	Behaviour disorder mental retardation	Y	4–6	F	1
2	30	Homosexual pedophilia	N	2–9	M	2
3	20	Mental retardation	Y			0
4	33	Pedophilia	N	8–11	F	2
5	28	Mental retardation	Y	5–7	M	1
6	30	Heterosexual pedophilia	N	6–11	F	4
7	24	Heterosexual pedophilia personality disorder	N	5–8	F	3
8	25	Personality disorder Mental retardation	N	4–5	M	2
9	19	Personality disorder	N	9–11	M & F	2
10	25	Pedophilia	N	8	F	1

[1] Information was obtained from the patients' clinical files. Their verbal reports often conflicted with the files in that they often claimed that the victims were older. Patient three was an exception in that he was institutionalized for arson. He reported an extensive history of molesting young children of both sexes which psychophysiological assessment supported.

the first aversion session, subjects set the intensity of the sine wave shock as high as they could tolerate and the shock was left at that intensity unless the subject requested it to be raised or lowered. Patients were encouraged to set the shock level high because of the well-documented occurrence of adaptation to shock (Quinsey, 1970) and the finding that aversion therapy is more effective with higher shock intensities (Tanner, 1973).

Shocks, slide changes, penile circumference, and skin conductance were recorded on a Beckman Type RS Dynograph. Experimental events were controlled by an aversive therapy unit built by Applied Automation Associates of Toronto.

Treatment and Psychophysiological Assessment Sessions

There was a total of 24 sessions in which physiological measures were taken. There was 1 preliminary assessment session, 3 assessment or test sessions, and 20 aversion therapy sessions. Subjects who did not respond favorably in the first 20 sessions were given additional treatment; the results of the additional treatment are not presented here. The first session was a standard diagnostic slide series which has been described elsewhere (Quinsey et al., 1975). It was used in conjunction with clinical file and verbal report data to choose a population of deviant slides for use in therapy. An age and sex category of deviant slides was selected if the patient showed large penile circumference responses (PCRs) to that category in the standard assessment session or selected victims from that category. Adult slides were chosen according to which sex the patient said he would like to respond. The three test sessions were identical in all respects: the first was given before aversion therapy began, the second after 10 aversion sessions, and the third after 20 aversion sessions. The test sessions consisted of a presentation of 20 slides – 10 from the deviant population, 5 of either adult males or females, depending on the patient's request, and 5 neutrals (which were defined as slides depicting scenes with no animal life and no fetish objects). These slides were shown in a fixed random sequence. Each slide presentation was 30 sec long and each interslide interval was 60 sec long unless the subject maintained an erection in which case the next slide presentation was postponed until the erection declined to baseline and ceased to fluctuate. The slides in the assessment test sessions were not used in aversion therapy and tests were conducted with shock electrodes removed. During test sessions, the peak PCR in the 60-sec period following slide onset, disregarding the first 2-sec period, was obtained. The reading at slide onset was subtracted from this figure to give the PCR used in the analyses. This method of scoring has been found to be superior to several others in differentiating child molesters from others (Quinsey, Note 2).

The skin conductance responses (SCRs) were of much shorter duration than the penile circumference changes. The largest magnitude response during the last 28 sec of the slide presentation was subtracted from the level at slide onset to obtain the SCR used in the analyses. Raw difference scores as well as Z score transformations were analysed for both physiological measures. The Z scores were computed for each subject and each session separately.

Each aversion therapy session consisted of 10 adult (male and/or female) slides and 10 slides of children. Slide presentations were 10 sec in duration and the slides were separated by a variable intertrial interval of 20 sec (the intervals were of 10, 20, and 30 sec duration and their order was determined from a table of random numbers, with the restriction that the intervals occurred equally often). A shock occurred throughout the last 2 sec of 80% of the deviant slides; none of the adult slides contained a shock. In each session, the slides were presented in one of five fixed sequences with different slides occurring in each sequence. A different fixed random order of slides was chosen

for each with the restriction that the first deviant slide in each sequence contained a shock and that not more than three slides of a kind occurred in a row. The five sequences were presented in a fixed order over 10 treatment sessions so that the subject saw each slide twice. A different random sequence of intertrial intervals occurred in each of the 5 sequences. The 100 slides were randomly rearranged after the first 10 sessions. The slides were mostly in colour and varied widely in degree of nudity and sexual attractiveness.

Skin conductance and penile circumference responses in the treatment sessions were both scored by forming two difference scores. In the first (SCR1 or PCR1), the reading at slide onset was subtracted from the reading at the eighth second. In the second (SCR2 or PCR2), the reading at slide onset was subtracted from the largest response occurring in the 2-sec to 8-sec interval following slide onset.

Both the raw scores and their Z score transformations were analysed. The Z scores were calculated for each session separately for each subject. In addition, a $\Delta\phi$ ix transformation (Lykken & Venables, 1971) was performed on the raw skin conductance data. In this transformation, each skin conductance response was divided by the largest response to shock in that session. The largest response was scored by subtracting the reading at slide onset from the peak response following shock.

Semantic Differential Measure

Each subject was interviewed regarding his sexual preferences and given Marks and Sartorius's (1968) semantic differential measure of sexual attitude. This scale is a 6-item inventory which yields a general evaluative score and a sex evaluative score for each concept rated. Three types of concepts were rated: (a) deviant sex concepts such as "6–10 year old girls," (b) normal sex concepts, for example, "women about your own age," and (c) a sexually "neutral" concept, "Oak Ridge." Oak Ridge was the building in which the treatment was conducted. The average rating of each concept on the general and evaluative scale was obtained for each subject before and after therapy.

Slide Sort Test

Because a number of subjects required help in performing the ratings because of their limited ability to perform abstract tasks and poor reading ability, a more concrete verbal test of sexual preference was also given before and after treatment. Two arrays of six black and white slides each were shown to the patient, one array at a time on a lighted slide sorting tray. A seventh colour slide depicting a sadomasochistic act was included in each array. In one array the figures were nude and in the other the figures were partially clothed. There was one representative of each sex at each of three age levels: adult, pubescent, and child. The subject was shown an array in a standard irregular sequence and asked to choose the most sexually attractive slide. This slide was then removed and the process repeated until the subject either refused to choose any more or only one slide was left.

Each slide was assigned a number corresponding to its order of choice and the two arrays were combined by averaging over categories. When a patient refused to rank a relevant slide by dismissing it with others as of no interest, the relevant slide was given the average of the unused ranks.

Data Analysis

The semantic differential measures and assessment session measures were subjected to similar analyses of variance. Responses to adult slides or concepts were compared with responses to relevant child slides or concepts both before and after therapy. The ranking

data were treated similarly, but separate non-parametric analyses were performed on the pre- and post-treatment data. Physiological data from the treatment sessions were averaged within a slide category for each subject over the 20 sessions and subjected to analyses of variance.

RESULTS

Assessment Sessions

As shown in Figure 1 the PCRs elicited by child slides decreased across assessment sessions relative to the other slides and the PCRs to the adult slides increased. An analysis of variance on the Z scores for the child and adult slides for the 10 subjects revealed a significant Slide Type by Session (pre- or post-) interaction ($F(1,18) = 7.29$, $p < .025$). There was, however, no significant Slides by Session interaction in the raw PCR data.
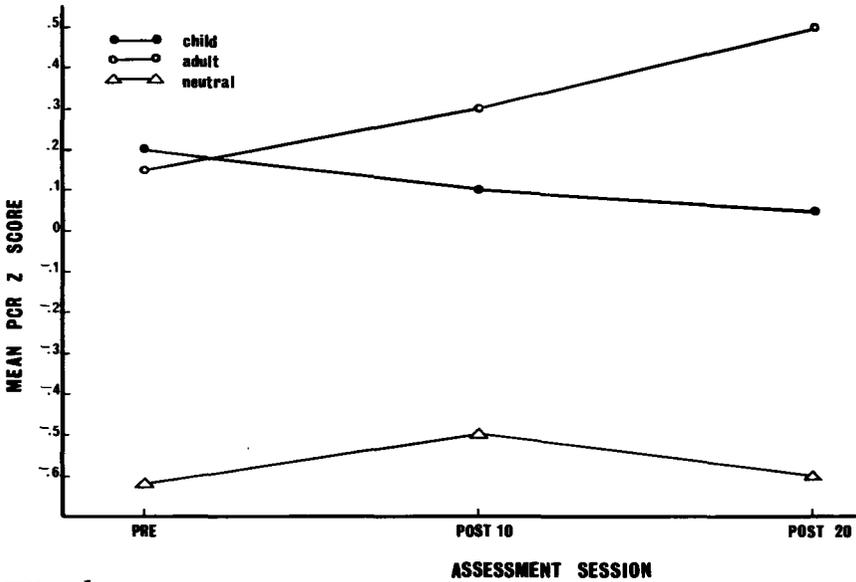

FIGURE 1
Average PCR *Z score for the assessment sessions.*

There were no significant Slide Type by Session interactions in the SCR analyses. There was, however, a significant difference between the adult and child slides with the SCR Z scores ($F(1,18) = 25.42$, $p < .001$), indicating that there were larger SCRs to the child than to the adult slides.

Despite the significant interaction found in the group analysis of the PCR Z score data, significant treatment effects were not often found within the data from individual subjects. The pre- and post-treatment PCR Z

scores for each subject were analysed separately with Wilcoxon Signed Ranks tests and the data were scored so that an increase in response magnitude to adult slides was the same mathematically as a decrease in magnitude to child slides. There were 10 child slides and 5 adult slides in each analysis. Of the 10 analyses only one subject showed significant improvement ($p < .02$) even though the mean PCR Z score data plotted for individual subjects indicated increased preference for adult slides relative to child slides in 7 out of the 10 cases.

Treatment Sessions

Neither of the PCR measures showed appreciable differentiation between the adult and child slides during the course of therapy. When the sessions were averaged for each subject and the adult and child PCRs compared using analysis of variance, neither the raw PCRs nor their Z score transformations approached significance.

In contrast, all the SCR measures differentiated between the adult and child slides. The largest SCRs were to child slides. Averaging over the 20 treatment sessions for each subject and comparing the adult and child slides with analysis of variance, the largest effect was found with the SCR1 Z scores ($F(1/9) = 85.10, p < .001$) and the second largest with the SCR2 Z scores ($F = 40.64$). The raw scores and the $\Delta\phi$ ix transformations for SCR1 and SCR2 yielded Fs in the twenties.

A one-tailed binomial test on the SCR1 Z score data for each subject was performed by noting the number of sessions of the 20 in which the mean SCR was higher for the child slides than for the adult slides. This test was significant for eight of the subjects at beyond the .001 level. The probability values lay between .05 and .01 for one of the remaining subjects, and between .002 and .001 for the other.

The SCR data presented above indicate that each subject discriminated between the slide categories. They do not, however, indicate acquisition of this differentiation. Data from the pre-treatment assessment session indicated that larger SCRs were made to child slides before therapy began and visual inspection of the SCR data from the first treatment session for each subject failed to give evidence of increasing differentiation.

Semantic Differential

The Semantic Differential data are graphed in Figure 2. Surprisingly, the changes of the adult and child concepts along the evaluative and sex dimensions appeared to be similar. As in the PCR Z score data, there was a significant Slide Type (adult or child) by Session (pre- or post-) interaction for both the evaluative ($F(1/16) = 12.04, p < .01$), and sexual
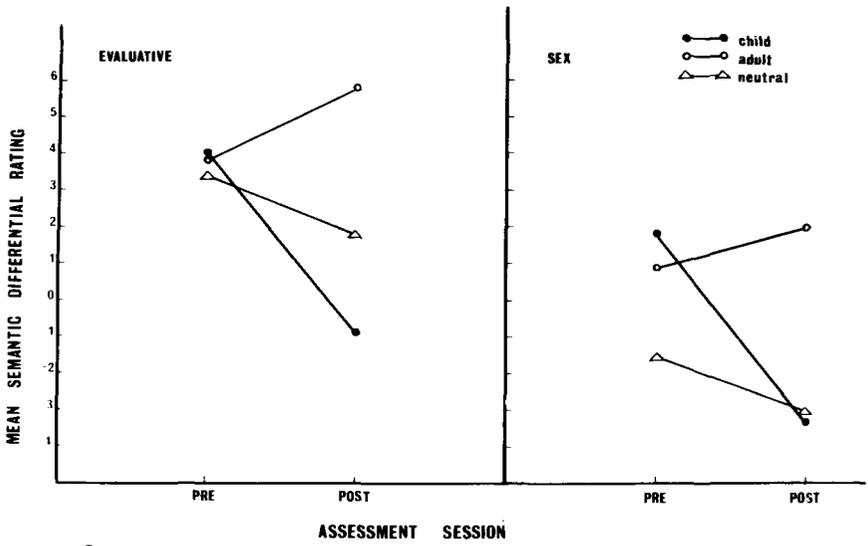
FIGURE 2
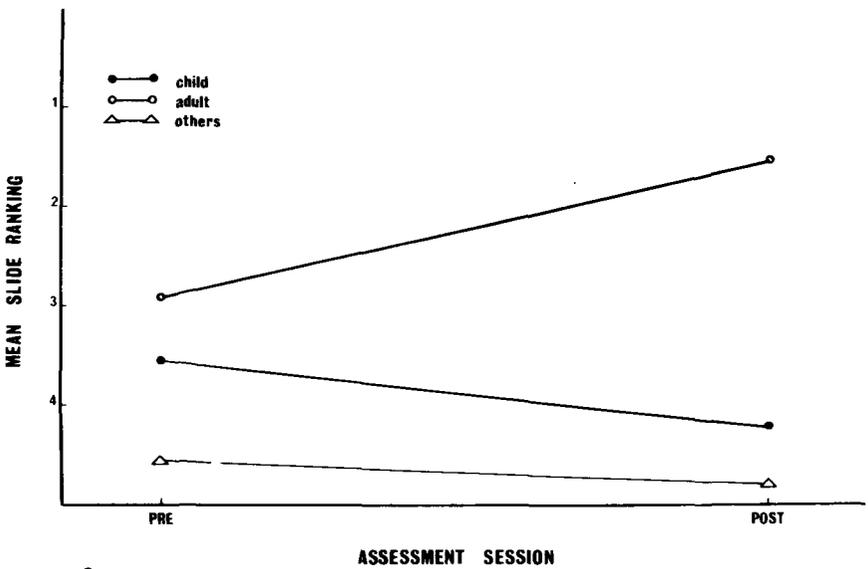*Average pre- and post-treatment semantic differential ratings.*

FIGURE 3
Average pre- and post-treatment slide ranking scores.

($F(1/16) = 10.56, p < .01$) dimensions. The semantic differential data for one subject were not collected by error.

Ranking Test

The ranking data shown in Figure 3 show an interaction between Slide Type and Session similar to those found with the PCR and semantic differential measures. Data for two subjects were missing. The difference between the targeted child slides and the targeted adult slides was not significant with a two-tailed Wilcoxon Signed Ranks Test before therapy ($T = 15, p > .10$) but was significant after therapy ($T = 0, p < .01$).

Relationships among Measures

The PCR and SCR measures of change appeared to be unrelated. In the treatment sessions, SCRS differentiated between child and adult slides, but the PCRs did not. In the generalization sessions, PCRS showed a treatment effect whereas the SCRS did not. The most striking aspect of the data was the correspondence between the ranking, semantic differential, and penile circumference measures in which all three indicated that adult stimuli became more sexually attractive whereas child slides became less sexually attractive.

In order to determine how patients' changes on one measure were related to their changes on the other measures, difference scores were computed for each patient based on his mean pre-test and mean post-test score for each of the four measures (PCR Z scores, semantic differential sex and evaluative scores, and the slide ranking data). Difference scores were computed for the adult and child categories separately. These difference or change scores were then ranked in order of magnitude within each measure and a rho statistic was calculated between each pair of measures for both the adult and child categories. None of the correlations approached significance except the correlation between the semantic differential sex and evaluative change scores (for the child category $rho = .63, N = 9$, $p < .05$, for the adult category, $rho = .70, p < .05$). In view of the fact that the measures as averaged over subject all showed similar treatment effects, it is somewhat surprising to find these measures of change uncorrelated.

DISCUSSION

The results of this study support the contention that the combination of skin conductance and penile circumference is useful in the assessment of changes in sexual preferences. These two measures provided different kinds of useful information: skin conductance responses indicating discrimination of the slide categories within treatment and penile circumference

responses showing treatment effects in assessment sessions. The verbal measures also yielded treatment effects but the absence of a close relation among the various pre- and post-treatment measures of change remains puzzling.

The group PCR data from the assessments indicate that the aversion therapy had a desirable effect. Semantic differential and slide ranking measures support this conclusion. Much less favorable, however, was the finding that the PCR assessment measure usually did not show significant improvement within individual subjects and the finding that even though the different pre- and post-treatment measures showed similar mean changes, subjects did not change similarly on them except for the two semantic differential measures.

The question of major importance is "how could the treatment be altered so as to produce larger magnitude effects within individual subjects?" The measure of most concern is the PCR because it is most unambiguously related to sexual preference in this population (Quinsey et al., 1975). No acquisition of a differential PCR to child and adult slides was observed in the treatment session data, perhaps because the interstimulus interval was too short. It seems reasonable that the demonstration of acquisition of an adult-child discrimination in the treatment session PCR data would be a step towards an increase in the magnitude of the treatment effect in the assessment tests. Longer slide duration would be required during treatment sessions.

It is of interest that increases in sexual preference to adult categories occur in the assessment test PCR scores during aversion therapy. Even though this increase is a consequence of the Z score conversion in certain of our data, it appeared in the raw scores and has been observed elsewhere (Bancroft, 1970; 1971). This increase in preference for stimuli with no programmed reinforcement may be explained by a "contrast" effect (Rescorla, 1969; Quinsey, 1970). Whatever the explanation for the phenomenon, however, its existence and obvious therapeutic relevance suggest that the variables which increase its magnitude ought to be studied. It may be that a discriminated operant paradigm with symmetrical contingencies (reward for appropriate increases in PCR and punishment for inappropriate increases) would produce the desired results.

RÉSUMÉ

Dix patients souffrant de pédophilie sont chacun soumis à 20 séances de conditionnement classique de type aversif et à trois séances de tests de mesure (avant tout traitement, après 10 séances, et après 20 séances de conditionnement). Le stimulus non conditionnel est un choc électrique. Les résultats montrent que, à toutes les séances, les réponses de conductance dermique sont plus fortes quand on projette des diaposi-

tives d'enfants plutôt que d'adultes. La mesure de la circonférence du pénis indique un léger changement des préférences sexuelles d'un test à l'autre. Les mesures de différenciation sémantique et de préférence sexuelle des diapositives montrent également une préférence sexuelle pour les adultes après le traitement, par rapport à la préférence sexuelle pour les enfants observée avant le traitement.

REFERENCE NOTES

1. STEFFY, R.A., & GAUTHIER, R. *Psychophysiological factors in aversive behavior therapies: A contrast between successfully and non-successfully treated pedophile offenders*. Paper read at the meeting of the Ontario Psychological Association, Kingston, Ontario, 1970.
2. QUINSEY, V.L. Progress report to the Ontario Mental Health Foundation, 1974.

REFERENCES

BANCROFT, J. *Deviant sexual behaviour: modication and assessment*. Oxford: Clarendon Press, 1974.
BANCROFT, J. A comparative study of aversion and desensitization in the treatment of homosexuality. In L.E. BURNS & J.L. WORSLEY (Eds.), *Behaviour therapy in the 1970's*. Bristol: Wright, 1970.
BANCROFT, J., & MATHEWS, A. Autonomic correlates of penile erection. *Journal of Psychosomatic Research*, 1971, *15*, 159–167.
CALLAHAN, E.J., & LEITENBERG, H. Aversion therapy for sexual deviation: Contingent shock and covert sensitization. *Journal of Abnormal Psychology*, 1973, *81*, 60–73.
FREUND, K. Erotic preference in pedophilia. *Behaviour Research and Therapy*, 1967, *5*, 339–348.
HENSON, D.C., & RUBIN, H.B. Voluntary control of eroticism. *Journal of Applied Behavior Analysis*, 1971, *4*, 37–44.
LAWS, D.R., & RUBIN, H.B. Instruction control of an automatic sexual response. *Journal of Applied Behavior Analysis*, 1969, *2*, 93–99.
LYKKEN, D.J., & VENABLES, P.H. Direct measurement of skin conductance: A proposal for standardization. *Psychophysiology*, 1971, *6*, 656–672.
MCCONAGHY, N., PROCTOR, D., & BARR, R. Subjective and penile plethysmography responses to aversion therapy for homosexuality: A partial replication. *Archives of Sexual Behavior*, 1972, *2*, 65–78.
MARKS, I.M., & SARTORIUS, N. A contribution to the measurement of sexual attitude. *Journal of Nervous and Mental Disease*, 1968, *145*, 441–451.
QUINSEY, V.L. Some applications of adaptation-level theory to aversive behavior. *Psychological Bulletin*, 1970, *73*, 441–450.
QUINSEY, V.L. Methodological issues in evaluating the effectiveness of aversion therapies for institutionalized child molesters. *Canadian Psychologist*, 1973, *14*, 350–361.
QUINSEY, V.L., STEINMAN, C.M., BERGERSEN, S.G., & HOLMES, T.F. Penile circumference, skin conductance, and ranking responses of child molesters and "normals" to sexual and nonsexual visual stimuli. *Behavior Therapy*, 1975, *6*, 213–219.
RESCORLA, R.A. Establishment of a positive reinforcer through contrast with shock. *Journal of Comparative and Physiological Psychology*, 1969, *67*, 260–263.
TANNER, B.A. Shock intensity and fear of shock in the modification of homosexual behavior in males by avoidance learning. *Behavior Research and Therapy*, 1973, *11*, 213–218.