
THE SENSITIVITY OF PHALLOMETRIC
MEASURES WITH RAPISTS

Martin L. Lalumière, MPs,

&

Vernon L. Quinsey, PhD

*Queen's University,
Kingston, Ontario, Canada*

Annals of Sex Research

ABSTRACT

The sensitivity of phallometric measures with rapists was investigated using individual rape indices of rapists and non-sex offenders from 14 published and unpublished data sets. *Sensitivity* (the probability of detecting deviant sexual arousal among rapists) was determined by alternately setting the *specificity* (the probability of not detecting deviant sexual arousal among non-sex offenders) at 70%, 80% or 90%. Eight data sets showed sensitivity values equal to, or greater than, 60% when specificity was set at 90%. Stimulus sets that contained more brutal and graphic descriptions of rape produced higher sensitivity values. A conservative rape index cut-off score of 1.0 seems optimal in that it yields moderate sensitivity values while minimizing the number of individuals who are falsely classified as presenting deviant sexual arousal. Conditions that may increase the sensitivity of phallometric assessments are presented and discussed.

Lalumière, M.L., & Quinsey, V.L. (1993). The sensitivity of phallometric measures with rapists. *Annals of Sex Research*, 6, 123-138.

The phallometric assessment of rapists must meet certain criteria in order to be useful in clinical settings. One of these criteria is that rapists should show, on average, sexual response profiles that are different from profiles obtained from non-rapist groups. Another is that the sensitivity of phallometric measures should be high enough to permit meaningful decisions regarding individual rapists.

The first criterion pertains to discriminant validity. As we discussed elsewhere (Lalumière & Quinsey, in press), establishing discriminant validity allows the inference that rapists' sexual preferences are *deviant*. The assignment of a particular response profile (preference) to the category 'deviant' can only be done by comparing response profiles of rapists and non-rapists. We recently conducted a meta-analysis of studies that compared rapists' and non-sex offenders' penile responses to stimuli describing adult heterosexual mutually consenting sex and adult heterosexual sexual coercion (Lalumière & Quinsey, in press). We obtained a large and statistically significant average group difference (effect size, $d = 0.82$). Moreover, rapists could be discriminated from non-sex offenders using different stimulus sets and in different settings (mental health or correctional). Interestingly, stimulus sets that contained more graphic and brutal descriptions of rape discriminated the two groups more effectively. This particular finding mirrors results from individual studies that explicitly varied stimulus content. For example, Rice, Chaplin, Harris, and Coutts (1993) found that rape descriptions which included victim suffering told from the female point of view discriminated rapists from non-rapists better than descriptions that included victim enjoyment told from the male point of view. The findings of the meta-analysis suggest that the discriminant validity of phallometric assessments of rapists is quite good and that there are conditions that augment it.

The second criterion mentioned above pertains to the question of whether rapists' and non-rapists' sexual response profiles overlap substantially. A statistically significant group difference is a necessary condition to infer the utility of phallometric assessments, but it may not be sufficient if a large number of rapists exhibit sexual response patterns that are too close to the average pattern obtained from non-rapists. If this were the case, very few assessments (i.e., only the most extreme results) would provide useful information about individual offenders. This question refers to the sensitivity of phallometric measures, an extension of discriminant validity.

The sensitivity of phallometric measures with child molesters who do not admit a sexual interest in children has been estimated to be 55% (Freund & Blanchard, 1989). This means that 55% of child molesters who do not admit a sexual interest in children nevertheless show sexual response profiles that are deviant. The corresponding value for non-child molesters is 5% (or specificity = 95%). These values were determined using decision rules developed with admit- ters and non-child molesters and subsequently applied to non-admitters and a

new group of non-child molesters (see Freund & Watson, 1991, for more information on the sensitivity of phallometric measures with child molesters). These findings indicate that when deviant sexual arousal is observed in a given child molester (using decision rules developed from derivation samples), there is a very high likelihood that it is real.

In this study we attempted to determine a) the sensitivity of phallometric measures with rapists using individual data from 14 published and unpublished data sets; b) whether different stimulus sets produce different sensitivity values; and c) the cut-off scores that maximize both the sensitivity and the specificity of phallometric assessments.

METHOD

The Studies Selected

Fourteen data sets were used in this study. They form a subset of a larger data base ($N = 18$) used by Lalumière and Quinsey (in press). These 14 data sets are from 13 separate studies and were selected according to the following five criteria: a) At least one group of identified rapists and at least one group of non-sex offenders were tested; b) at least one stimulus category of mutually consenting sex between an adult male and an adult female and at least one stimulus category depicting a sexually aggressive interaction between a male protagonist and a non-consenting female were used; c) audiotaped or videotaped stimuli were used; d) data from studies of voluntary control of penile responses or studies using non-standard instructions such as "maintain arousal" and "suppress arousal" were excluded, as were data from groups within studies given special instructions or special experimental conditions; e) sufficient details about penile responses had to be provided in order to allow the computation of a rape index for each individual subject (missing information was requested from the authors when necessary). Criterion 'e' is responsible for the exclusion of four data sets used by Lalumière and Quinsey (in press).

The 14 data sets selected are presented in Table 1. Data from all sets were obtained using auditory stimuli. Italicized values are from studies that did not meet the first criterion, that is, only rapists were tested. However, rapists in these studies (hereafter, *secondary data sets*) could be compared with control subjects from studies that had the same setting, stimulus set, and instructions. Rapists from Quinsey and Chaplin (1982) and Looman (1989b) were compared with non-sex offenders from Quinsey, Chaplin, and Varney (1981). Rapists from Eccles (1990), Malcolm (1992), and Looman (1989a) were compared with a combined group of non-sex offenders from Baxter, Barbaree, and Marshall (1986), and Barbaree, Marshall, and Lanthier (1979).

Data from Malcolm (1992) are from 83 rapists that were assessed pre-treatment between 1983 and 1985 using Barbaree's stimulus set. Data from Looman (1989 a & b) are from rapists who were assessed pre-treatment in the 1980s; 68 rapists (Looman a) were assessed using the stimuli of Barbaree et al. (1979), and 79 rapists (Looman b) were assessed using the stimuli of Quinsey et al. (1981).

With the exception of Looman (1989 a & b), each study used independent samples (subjects used in two or more studies appear only once in Table 1 and in the analyses presented below). Looman (1989 a & b) tested some rapists twice with two different stimulus sets. In the present article these rapists are grouped according to the stimulus set used in their assessment and are compared with non-sex offenders tested with the same stimulus set; therefore Looman's (1989) study is divided into two partially independent data sets (a and b).

Although Rice, Chaplin, Harris, and Coutts (1990) was considered an outlier in Lalumière and Quinsey's meta-analysis, the data from that study were used here because in these analyses the focus is on individual cases, rather than on a study statistic. The reader is referred to Lalumière and Quinsey (in press) for more information about study selection, reasons for excision, case exclusion, and sample replication.

Rape Index

The dependent measure for the analyses was the rape index, originally developed by Abel, Barlow, Blanchard, and Guild (1977). This index provides a measure of arousal to rape stimuli relative to consenting stimuli. The rape index was calculated by dividing the average response to all heterosexual rape stimuli by the average response to all mutually consenting heterosexual stimuli. A rape index greater than 1.0 indicates greater responding to rape stimuli, hence a preference for rape.

Although other indices have greater discriminative power, and therefore probably superior sensitivity (e.g., difference scores calculated from standard scores, see Harris, Rice, Quinsey, Chaplin, & Earls, 1992), this particular index was used because it could be obtained or calculated in all 14 data sets.

Analytical strategy

The strategy to determine sensitivity was based on the following rationale. Assuming that phallometric measures produced perfect discrimination (i.e., that all rapists show deviant sexual arousal and all non-sex offenders show non-deviant sexual arousal), it would be expected that, within a given study, the smallest individual rape index in the rapist group would be larger than the largest individual rape index in the non-rapist group. Of course, this assumption is

unlikely to be true for at least three reasons: first, because of measurement error; second, because of faking; and third, because it is unlikely that all rapists are sexually deviant, and that all non-sex offenders are sexually non-deviant.

In the absence of a history of sexual behaviors and self-reported fantasies for all subjects to be used as an external criterion, it is difficult to determine, concurrently, the sensitivity and the specificity of phallometric assessments of rapists. Sensitivity in this context refers to the likelihood of identifying deviant arousal among rapists. Specificity refers to the likelihood of not detecting deviant arousal among non-sex offenders.

Because of the difficulty in determining sensitivity and specificity concurrently, the specificity was fixed at different levels to measure its effect on sensitivity. We selected three specificity levels, 70%, 80%, and 90%; these levels determined the percentage of non-rapists classified as showing non-deviant (normal) arousal. The rape index cutoff values that classified 70%, 80%, or 90% of non-rapists as sexually non-deviant were then used to determine the percentage of rapists that would be classified as sexually deviant.

The cutoff scores were simply the rape index values corresponding to the minimum percentile ranks of .70, .80, and .90, calculated on the non-sex offender subjects in a given study. For example, in a study with a sample of 20 non-sex offenders, the 18th largest rape index was used as the 90% specificity cutoff score; a rape index greater than that cutoff value would classify rapists in that study as sexually deviant.

RESULTS

Descriptive statistics for each study are presented in Table 1. Studies are presented in decreasing order of the obtained correlation between individual rape index and group membership. Also included is the newly developed Common Language Effect Size statistic (*CL*; McGraw & Wong, 1992). *CL* is the probability that a randomly chosen member of the rapist group has a higher rape index than a randomly chosen member of the non-sex offender group. In the absence of between-group differences, *CL* approaches .50.

Sensitivity Values

Sensitivity values are presented for each study in Table 2 for the three pre-selected levels of specificity. Also included in Table 2 are the rape index cutoff values, the classification accuracy values, the Relative Improvement Over Chance values (RIOC, Loeber, & Dishion, 1983), and the phi coefficients representing the relationship between predicted group membership (deviant or not) and actual group membership (rapists or non-sex offenders). The RIOC statistic is particularly

Table 1
Studies selected (italicized values are from secondary data sets)

Study	Sample size		<i>rb</i>	<i>CL</i>	Stimulus Set ^a
	Rapist	Control			
Rice et al. (1990)	10	13	.85***	.98	OR
Quinsey & Chaplin (1984)	15	15	.71***	.92	OR
Earls & Proulx (1986)	10	10	.62**	.86	A
Quinsey & Chaplin (1982)	24	20	.56***	.83	OR
Quinsey et al. (1981)	20	20	.52***	.80	OR
Quinsey et al. (1984)	20	20	.52***	.80	OR
Proulx et al. (1992)	10	10	.51***	.79	A
Abel et al. (1978)	19	15	.37*	.72	A
Barbaree et al. (1979)	10	10	.37	.70	B
Eccles (1990)	19	41	.28*	.66	B
Malcolm (1992)	83	41	.26**	.66	B
Looman (a) 1989)	68	41	.17	.61	B
Baxter et al. (1986)	50	31	.16	.59	B
Looman (b) (19489)	79	20	.12	.60	OR

Note: *rb* Point bi-serial correlation coefficient. *CL* common Language Effect Size. **a** OR = Oak Ridge, A = Abel, B = Barbaree. * $p < .05$; ** $p < .01$; *** $p < .001$.

useful for comparing classification accuracy across contingency tables that have different base rates (proportions of actual rapists) and different selection ratios (proportions predicted to be deviant).

First it can be observed that rape index cutoff values were quite variable across studies. Second, because of a lowering of specificity (higher selection ratios, or

Table 2
Sensitivity and classification accuracy for the three levels of predetermined specificity (italicized values are from secondary data sets)

Study	Rape Index Cutoff (>)	Sensitivity (%)	CA (%) ^a	RIOC (%)	Phi
≥ 90% Specificity					
Rice et al.	0.66	100	96	84	.92***
Quinsey & Chaplin	0.57	100	97	88	.94***
Earls & Proulx	0.90	100	95	82	.90***
Quinsey & Chaplin	<i>1.16</i>	67	77	76	<i>.57***</i>
Quinsey et al.	1.16	45	67	64	.39*
Quinsey et al.	0.75	85	88	79	.75***
Proulx et al.	0.90	70	80	75	.61**
Abel et al.	0.89	63	83	76	.58**
Barbaree et al.	0.63	60	75	71	.52*
Eccles	<i>0.94</i>	16	67	16	<i>.09</i>
Malcolm	<i>0.94</i>	19	43	40	<i>.12</i>
Looman (a)	<i>0.94</i>	22	43	38	<i>.16</i>
Baxter et al.	0.94	14	43	22	.06
Looman (b)	<i>1.16</i>	28	40	59	<i>.17</i>
≥ 80% Specificity					
Rice et al.	0.62	100	91	70	.84***
Quinsey & Chaplin	0.41	100	90	67	.82***
Earls & Proulx	0.81	10	90	67	.82***
Quinsey & Chaplin	<i>1.00</i>	92	96	66	<i>.73***</i>
Quinsey et al.	1.00	60	80	56	.41*
Quinsey et al.	0.54	95	88	65	.76***
Proulx et al.	0.80	70	75	56	.50*
Abel et al.	0.59	79	79	62	.59**
Barbaree et al.	0.53	60	70	50	.41
Eccles	<i>0.76</i>	26	67	10	<i>.08</i>
Malcolm	<i>0.76</i>	32	43	31	<i>.14</i>
Looman (a)	<i>0.76</i>	31	43	27	<i>.12</i>
Baxter et al.	0.83	20	43	02	.01
Looman (b)	<i>1.00</i>	34	43	36	<i>.12</i>

Table 2, Continued

Study	Rape Index Cutoff (>)	Sensitivity (%)	CA (%) ^a	RIOC (%)	Phi
≥ 70% Specificity					
Rice et al.	0.35	100	83	50	.70***
Quinsey & Chaplin	0.39	100	83	50	.71***
Earls & Proulx	0.72	100	85	54	.73***
Quinsey & Chaplin	0.71	96	84	54	.69***
Quinsey et al.	0.71	85	78	48	.56***
Quinsey et al.	0.44	95	82	52	.67***
Proulx et al.	0.70	90	80	50	.61*
Abel et al.	0.53	79	76	52	.52**
Barbaree et al.	0.48	60	65	33	.30
Eccles	0.56	47	63	16	.18
Malcolm	0.56	49	56	32	.19*
Looman (a)	0.56	38	50	16	.09
Baxter et al.	0.63	40	52	19	.11
Looman (b)	0.71	49	54	34	.16

Note: a Overall Classification Accuracy. $p < .05$; ** $p < .01$, * $p < .001$.**

lower cutoff values), sensitivity increased; this produces a higher number of true and false positives (more subjects, both rapists and non-sex offenders, would be classified as deviant). Third, all 14 RIOC values decreased when specificity went from 90% to 80%, or from 90% to 70%, and 10 out of 14 RIOC values decreased when specificity went from 80 to 70%. Fourth, all RIOC values were positive. Finally, the classification accuracy values and associated phi coefficients show that most studies achieved statistically reliable classification.

Effect of stimulus set

Because stimulus set was found to be a moderator variable in Lalumière and Quinsey's (in press) meta-analysis, the next analysis combined subject data according to the three types of stimulus sets that were used in more than one study (see Table 1): Oak Ridge (six data sets), Abel (three data sets), and Barbaree (five data sets). The first two stimulus sets describe more violent and graphic descriptions of rape, compared to the third set. Again, cutoff values were determined so that 70%,

Table 3
Sensitivity and classification accuracy for the validation samples for the
three levels of predetermined specificity as a function of stimulus sets
(all studies)

Stimulus set	Cutoff (>)	Sensitivity (%)	Specificity (%)	CA (%) ^a	RIOC (%)	Phi
≥ 90% Specificity						
Oak Ridge	0.82	69	82	71	71	.39*
Abel	0.80	77	72	75	55	.47*
Barbaree	0.93	20	86	26	28	.04
≥80% Specificity						
Oak Ridge	0.60	76	76	76	65	.42*
Abel	0.70	85	67	79	51	.51*
Barbaree	0.63	40	67	42	15	.04
≥70% Specificity						
Oak Ridge	0.44	83	70	81	60	.46*
Abel	0.59	90	56	79	41	.49*
Barbaree	0.52	50	57	51	14	.04

Note: a Overall Classification Accuracy. * p < .001.

80%, and 90% of control subjects were classified as non-deviant. Cutoff values were determined, however, using one-half of each of the three samples of non-sex offenders, and then applied to the other half of non-sex offenders and to all rapists for a given stimulus set. As an illustration, for the Oak Ridge stimulus set, the derivation sample consisted of 34 non-sex offenders and the validation sample consisted of 34 non-sex offenders and 168 rapists. This procedure permitted an assessment of the cross-validation consistency of the cutoff values regarding specificity. The results of this analysis are presented in Table 3. The sample sizes in the validation samples are presented in Figure 1, along with the classification distribution at 90% specificity, for the three stimulus sets. Thus, the values presented in Figure 1 correspond to the percentages presented in Table 3 at 90% pre-set specificity.

Table 3 shows that there was a large difference between sensitivity levels across stimulus sets, for all three specificity levels: Only the Oak Ridge and Abel stimulus sets reached a statistically significant classification accuracy. This shows that a

Oak RidgePredicted
Deviant Non-deviant

Actual Group	Rapists	116	52	168
	Normals	6	28	34
		122	80	N = 202

AbelPredicted
Deviant Non-deviant

Actual Group	Rapists	30	9	39
	Normals	5	13	18
		35	22	N = 57

BarbareePredicted
Deviant Non-deviant

Actual Group	Rapists	47	183	230
	Normals	3	18	21
		50	201	N = 251

Figure 1. Classification of rapists and non-sex offenders as a function of stimulus set for 90% preset specificity.

Table 4
Specificity and sensitivity using a rape index cutoff value of 1.0
(all studies)

Stimulus Set	Sensitivity %	Specificity %	CA % ^a	RIOC %	Phi
Oak Ridge	67	90	67	77	.44*
Abel	69	94	81	85	.65*
Barbaree	14	95	26	61	.10

Note: a Overall Classification Accuracy. * $p < .001$.

reliable assignment to group membership (deviant vs. non-deviant) on the basis of the rape index alone can be made using the first two stimulus sets but not the third.

Although classification accuracy increased with decreases in preset specificity, the RIOC decreased as well, suggesting that one would do better to assign subjects using cutoff scores that allow 10% or less of false positive errors. The decrease in specificity in the validation sample could be a result of the fact that the cutoffs were derived from small samples (34, 18, and 20, respectively); however, a conservative strategy would be to use slightly higher cutoff values in order to avoid false positive errors.

In the last analysis, the specificity and sensitivity was examined using a convenient and more conservative rape index cutoff value of 1.0. This analysis used all subjects. Results are presented in Table 4. Again, the first two stimulus sets, but not the third, achieved statistically significant classification. Specificity was high (90% and higher) for all stimulus sets. Sensitivity was moderate for the Oak Ridge and Abel stimulus sets, and low for the Barbaree stimulus set.

DISCUSSION

Results of this study satisfy the second criterion identified in this article for the utility of phallometric assessment of rapists: The rape index is moderately sensitive under certain conditions. A rape index cutoff score of 1.0 is conservative in that it leads to very few false positive errors while allowing the identification of a substantial number of rapists as sexually deviant. The larger specificity than sensitivity of phallometric measures using a cutoff score of 1.0 signifies that the presence of deviant sexual arousal among rapists is more informative than its absence. This latter point has also been stressed by Freund and Watson (1991, p.

259) with regard to the phallometric assessment of child molesters.

The next task for researchers is to identify conditions that may increase sensitivity values. An increase in sensitivity of course implies an increase in discriminant validity. Five possible conditions are discussed here.

First, many subjects, but especially rapists, are motivated to simulate non-deviant arousal. To the extent that rapists are successful in simulating non-deviant arousal (Wydra, Marshall, Earls, & Barbaree, 1983), the probability of detecting deviant sexual arousal among rapists in phallometric assessments is reduced. Modifications of the assessment procedure to abolish or diminish efforts to dissimulate would inevitably increase sensitivity. Among recent attempts at such modification (e.g., Lalumière & Earls, 1992; Quinsey & Chaplin, 1988), Quinsey and Chaplin's (1988) method of semantic tracking developed using non-sex offenders is promising and is presently being evaluated with sex offender subjects.

The second possibility refers to the scoring method used to summarize phallometric scores. Harris et al. (1992) found that discriminant validity is enhanced by scoring methods that use scores that compensate for individual differences in responsivity (preferably z scores) and scores that reflect relative arousal (preferably difference scores) rather than arousal to a single category of stimuli.

The third possibility pertains to stimulus content. Individual studies that systematically varied stimulus content have reliably found that some descriptions of rape (usually brutal, violent, and graphic) discriminate rapists from non-rapists better (Proulx, 1992; Quinsey & Chaplin, 1984; Quinsey, Chaplin, & Upfold, 1984; Rice et al., 1993). Although in Lalumière and Quinsey's (in press) meta-analysis, different stimulus sets tended to produce variation in only rapists' rape indices, it remains to be determined whether these larger variations among rapists are due to an inhibitory or an excitatory process.

The fourth possibility is the number of stimulus exemplars used in the assessment. From measurement theory it can be expected that the larger the number of items, the greater the reliability, and therefore the greater the discriminative potential. Lalumière and Quinsey (in press) found a high but non-significant correlation, $r(9) = .40$, between number of exemplars and study effect sizes.

Finally, it is unlikely (although theoretically possible) that sexual preference is a relevant issue for all rapists. Rapists are heterogeneous in their characteristics (Prentky & Knight, 1991), including their motivation for rape. Langevin et al. (1985) separated their rapists into sadist ($n = 7$) and non-sadist ($n = 9$) groups and compared their average rape index. No difference was found. However, neither group could be differentiated from non-rapists ($n = 15$), thereby preventing any conclusion about the presence or absence of rapist types. Barbaree, Seto, Serin, Amos, and Preston (in press) used Prentky and Knight's (1990) rapist typology to classify their rapists into two groups and four subgroups: non-sexual (opportunistic and vindictive) and sexual (non-sadistic and sadistic). They found that the sexual group ($n = 12$) had a higher rape index than the non-sexual group

(n = 22). There was no difference between the sadistic and the non-sadistic subgroups. However, none of these rapists were compared with control subjects, preventing the conclusion that only the sexual group showed *deviant* preferences. Rice, Chaplin, Harris, and Coutts (1993) compared seven sadistic rapists with seven non-sadistic rapists using a large number of different rape scenarios. Both groups could be differentiated from a control group (n = 14) and, unexpectedly, the non-sadist group showed more deviant arousal. Although rapist types with regard to phallometrically measured sexual arousal have yet to be found, their discovery would undoubtedly increase phallometric sensitivity if an independent marker of the types could be identified.

Results from this study, in combination with results from studies on discriminant and predictive validity, support the use of phallometric assessments of rapists in clinical settings. The most important contribution of phallometric assessments in clinical settings is to improve the construction of individual theories of offending that can subsequently be used in treatment planning, including the context in which the treatment is delivered. (i.e., degree of supervision and incapacitation).

Acknowledgements

Preparation of this article was supported by a Fellowship from the Social Sciences and Humanities Research Council of Canada accorded to the first author and by a contract between the second author and the Kingston Psychiatric Hospital. The authors would like to thank C.M. Earls, G.T. Harris, M.E. Rice, and M.C. Seto for their comments on the manuscript of this article, and H.E. Barbaree, T.C. Chaplin, C.M. Earls, A. Eccles, A. Gordon, J. Looman, B. Malcolm, J. Proulx, and the Correctional Services of Canada, for providing us with raw data.

REFERENCES

- Abel, G.G., Barlow, D.H., Blanchard, E.B., & Guild, D. (1977). The components of rapists' sexual arousal. *Archives of General Psychiatry*, 34, 895-903.
- Abel, G.G., Blanchard, E.B., Becker, J.V., & Djenderedjian, A. (1978). Differentiating sexual aggressives with penile measures. *Criminal Justice and Behavior*, 5, 315-332.
- Barbaree, H.E., Marshall, W.L., & Lanthier, R.D. (1979). Deviant sexual arousal in rapists. *Behaviour Research and Therapy*, 17, 215-222.
- Barbaree, H.E., Seto, M.C., Serin, R.C., Amos, N.L., & Preston, D.L. (in press). Sexual arousal to rape in sexual and nonsexual rapist subtypes. *Criminal Justice and Behavior*.
- Baxter, D.J., Barbaree, H.E., & Marshall, W.L. (1986). Sexual responses to consenting and forced sex in a large sample of rapists and nonrapists. *Behaviour Research and Therapy*, 24, 513-520.
- Earls, C.M., & Proulx, J. (1986). The differentiation of francophone rapists and nonrapists using penile circumferential measures. *Criminal Justice and Behaviour*, 13, 419-429.

- Eccles, T. (1990). *The clinical utility of the Rape Index*. Unpublished doctoral Dissertation, Queen's University, Kingston.
- Freund, K., & Blanchard, R. (1989). Phallometric diagnosis of pedophilia. *Journal of Consulting and Clinical Psychology, 57*, 100-105.
- Freund, K., & Watson, R.J. (1991). Assessment of the sensitivity and specificity of a phallometric test: An update of phallometric diagnosis of pedophilia. *Psychological Assessment, 3*, 254-260.
- Harris, G.T., Rice, M.E., Quinsey, V.L., Chaplin, T.C., & Earls, C.M. (1992). Maximizing the discriminant validity of phallometric data. *Psychological Assessment, 4*, 502-511.
- Lalumière, M.L., & Earls, C.M. (1992). Voluntary control of penile responses as a function of stimulus duration and instructions. *Behavioral Assessment, 14*, 121-132.
- Lalumière, M.L., & Quinsey, V.L. (in press). The discriminability of rapists from non-sex offenders using phallometric measures: A meta-analysis. *Criminal Justice and Behavior*.
- Langevin, R., Bain, J., Ben-Aron, M.H., Coulthard, R., Day, D., Roper, V., Russon, A.E., Webster, C.D., & Wortzman, G. (1985). Sexual aggression: Constructing a predictive equation: A controlled pilot study. In R. Langevin (Ed.), *Erotic preference, gender identity, and aggression in men: New research studies* (pp. 39-76). Hillsdale, NJ: Erlbaum.
- Loeber, R., & Dishion, T. (1983). Early predictor of male delinquency: A review. *Psychological Bulletin, 94*, 68-99.
- Looman, J. (1989). *Rapists/pedophiles: A distinct group?* Unpublished honour's thesis, University of Saskatchewan, Saskatoon.
- Malcolm, B. (1992). *Rapists assessment data from the Sex Offender Treatment Program of the Regional Treatment Centre, Kingston*. Unpublished manuscript.
- McGraw, K.O., & Wong, S.P. (1992). A common language effect size statistic. *Psychological Bulletin, 111*, 361-365.
- Prentky, R.A., & Knight, R.A. (1991). Identifying critical dimensions for discriminating among rapists. *Journal of Consulting and Clinical Psychology, 59*, 643-661.
- Proulx, J. (1992, June). *Les réponses pénitentes de violeurs à deux types de stimuli de viol*. Paper presented at the annual convention of the Canadian Psychological Association, Québec City.
- Quinsey, V.L. & Chaplin, T.C. (1982). Penile responses to nonsexual violence among rapists. *Criminal Justice and Behavior, 9*, 372-384.
- Quinsey, V.L., & Chaplin, T.C. (1984). Stimulus control of rapists' and non-sex offenders' sexual arousal. *Behavioral Assessment, 6*, 169-176.
- Quinsey, V.L., & Chaplin, T.C. (1988). Preventing faking in phallometric assessments of sexual preference. In R.A. Prentky & V.L. Quinsey (Eds.) *Human sexual aggression: Current perspectives* (pp. 49-58). New York: Annals of the New York Academy of Sciences.
- Quinsey, V.L., Chaplin, T.C., & Upfold, D. (1984). Sexual arousal to nonsexual violence and sadomasochistic themes among rapists and non-sex-offenders. *Journal of Consulting and Clinical Psychology, 52*, 651-657.
- Quinsey, V.L., Chaplin, T.C., & Varney, G. (1981). A comparison of rapists' and non-sex offenders' sexual preferences for mutually consenting sex, rape, and physical abuse of women. *Behavioral Assessment, 3*, 127-135.
- Rice, M.E., Chaplin, T.C., Harris, G.T., & Coutts, J. (1990). *Empathy for the victim and sexual*

arousal among rapists and nonrapists (Report No. 8). Penetanguishene: Mental Health Centre Research.

Rice, M.E., Chaplin, T.C., Harris, G.T., & Coutts, J. (1993). *Empathy for the victim and sexual arousal among rapists and nonrapists*. Manuscript submitted for publication.

Wydra, A., Marshall, W.L., Earls, C.M., & Barbaree, H.E. (1983). Identification of cues and control of sexual arousal by rapists. *Behaviour Research and therapy*, 21, 469-476.