NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a confréré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé où si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, tests publiés, etc.) ne sont pas microfilmés.

PLETHYSMOGRAPHIC ASSESSMENT OF SEXUAL AROUSAL IN PEDOPHILES: THE RELATIONSHIP BETWEEN INTELLIGENCE, AS MEASURED BY THE WAIS-R AND AROUSAL

by

Conrad Bowden
B.A., Antioch University, 1978
M.A., Simon Fraser University, 1983

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY in the Department of Psychology

© Conrad Bowden 1987
SIMON FRASER UNIVERSITY November, 1987

All rights reserved. This work may not be reproduced in whole or in part, by photocopy or other means, without permission of the author.
Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

Name: Conrad Bowden
Degree: Doctor of Philosophy
Title of thesis: Plethysmographic Assessment of Sexual Arousal in Pedophiles: The Relationship Between Intelligence, as Measured by the WAIS-R and Arousal

Examining Committee:
Chairman: Dr. B. Krane

Dr. R. G. Ley
Senior Supervisor

Dr. R. Koopman

Dr. R. Roesch

Dr. A. Wydra

Dr. D. Cox

Dr. T. Dalby
External Examiner
Department of Psychology
Calgary General Hospital

Date Approved: November 23, 1987
PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay

Plethysmographic Assessment of Sexual Arousal in Pedophiles:

The Relationship Between Intelligence, as Measured by the WAIS-R and Arousal

Author:

(signature)

Conrad Bowden

(name)

December 16, 1987

(date)
ABSTRACT

The assessment and treatment of pedophilia are important concerns in the field of forensic psychology. Plethysmographic assessment of sexual arousal in pedophiles is routinely undertaken for these purposes. Two problems which commonly emerge in the plethysmographic assessment procedure are insufficient arousal and the occurrence of false negatives. Developments in cognitive-behavioral theory and some empirical findings suggest that the intelligence of subjects may influence plethysmographic assessment outcomes. Therefore, in the present study, it was hypothesized that intelligence is related to the outcome of plethysmographic assessment. Thirty-nine currently or previously federally incarcerated, self-confessed pedophiles were administered a plethysmographic assessment of sexual arousal, the WAIS-R, and other cognitive and personality measures. Subjects were assigned to either an "aroused" or "non-aroused" group on the basis of their sexual arousal profiles. There were significant differences between the groups on the Comprehension sub-test of the WAIS-R and on Verbal IQ - Performance IQ differences. No significant differences were found on file data variables, state anxiety, social desirability, or neurological impairment as measured by the Trail Making Test. Compared to the aroused group, the non-aroused group had significantly higher Comprehension subtest scores and greater Verbal IQ - Performance IQ differences. For subjects within the aroused group, no relationship was found
between a measure of deviance (PDQ) and intelligence. Using the Comprehension subtest as the independent variable, logistic regression analyses correctly classified 33 of 39 subjects. The results support the hypothesis that the outcome of plethysmographic assessment is related to an aspect of intelligence. Specifically, it is suggested that verbal social comprehension (as measured by the WAIS-R) moderates a subject's arousal in plethysmographic assessment; high Comprehension subjects exhibited less arousal than low Comprehension subjects. It may be that within the assessment situation pedophiles with high comprehension scores are influenced not only by the immediate qualities of the sexual stimuli, but also by verbally stored understanding about social functioning. It is possible that the absence of arousal is mediated by such understanding and may not reflect deliberate attempts to suppress arousal. It is suggested that plethysmographic assessments be conceived within a framework that accounts for stable cognitive factors. Implications for clinical practice and future research are considered.
I would like to acknowledge and thank the following for the critical roles they played in this research project. My fellow graduate students, especially Libby, Walter, Cheryl and Simon, gave invaluable support and encouragement. The administrative and technical staff, especially Malcolm, Franz, Mary and Bev, were always available and helpful. My friends and family watched and cared. My supervisory committee supplied the necessary guidance. The Correctional Service of Canada granted me permission to conduct my study with their inmates. The Corrections staff aided my project. Finally, the subjects made the study possible by their participation.
# TABLE OF CONTENTS

Approval ................................................................. ii
Abstract ........................................................................ iii
Acknowledgements ....................................................... v
List of Tables .................................................................. ix

I. Introduction ................................................................... 1
   Plethysmographic Assessment of Pedophiles .................. 3
   Pedophilia ................................................................. 4
   Diagnostic Problems .................................................. 4
   Sexual Offense Patterns ............................................. 8
   Advantages of Plethysmographic Assessment ............... 11
   Problems with Plethysmographic Assessment .............. 12

II. Theoretical Models ...................................................... 15
   Model 1: S – R ......................................................... 16
   Model 2: S – States – R ............................................ 16
   Model 3: S – Cognition – R ....................................... 22

Overview of the Current Study ..................................... 28

II. METHOD ................................................................. 30
   Subjects ....................................................................... 30
   Apparatus .................................................................... 33
   Penile Tumescence Measurement ................................ 33
   Stimuli ......................................................................... 37
   Procedures .................................................................... 40
   Sexual Arousal Assessment ....................................... 42
   Psychometric Measures ............................................. 45
   State Anxiety Inventory ............................................. 46
III. RESULTS ........................................................... 51

Section 1: Descriptive Analyses ........................................ 51
File Data .................................................................. 51
Intelligence Data .......................................................... 54
Other Psychometric Data ................................................. 55
Sex of Victim Data ......................................................... 59

Section 2: Arousal Data ..................................................... 59
Calculation of Maximal Arousal ........................................ 60
Classification as Aroused or Non-Aroused to Children .......... 65

Section 3: Relationships Between Arousal/Non-arousal and Other Variables ....................................................... 68
Intelligence ................................................................ 69
Verbal IQ - Performance IQ Differences .......................... 72
Age ........................................................................ 72
Other Variables ................................................................ 75
Classification of Subjects into Aroused/Non Aroused Groups by Means of Logistic Regression ......................... 78
Comprehension ............................................................ 79
Pedophilic Deviance Quotients ......................................... 82
Pedophilic Deviance Quotients and Intelligence: Correlations .......................................................... 83
Pedophilic Deviance Quotients and Intelligence: Cutoff Scores .................................................................. 83
Generalization Data ......................................................... 86
Intelligence ................................................................. 86
Other Variables .................................................. 88
IV. DISCUSSION .................................................. 90
Review of Hypotheses ............................................. 90
Hypothesis 1 - Average IQ .................................... 91
Hypothesis 2 - Percent Insufficiently Aroused ........... 91
Hypothesis 3 - Percent False Negatives .................. 92
Hypothesis 4 - Cognitive Ability and Plethysmographic Assessment ........................................... 94
Arousal/Non-Arousal and the Comprehension Subtest ... 96
Arousal/Non-Arousal and Verbal IQ - Performance IQ Differences .................................................. 101
Other Findings .................................................. 103
Arousal/Non-Arousal and Age .................................. 103
Arousal/Non-Arousal and Other Psychometric Variables 104
Pedophilic Deviance Quotients ................................ 107
Implications .................................................. 109
Theoretical Implications: Support for S - Cognition - R .......................................................... 109
Clinical Implications ........................................ 110
Generalizability ................................................. 112
Future Research ................................................ 114
Summary ....................................................... 116
Appendix A ....................................................... 117
Appendix B ....................................................... 118
Appendix C ....................................................... 121
References ...................................................... 122
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Intelligence of Pedophiles: Results of Previous Studies</td>
</tr>
<tr>
<td>2</td>
<td>Descriptive Subject Information from File Data</td>
</tr>
<tr>
<td>3</td>
<td>Classification of the Distribution of Intelligence Scores</td>
</tr>
<tr>
<td>4</td>
<td>WAIS-R Subtest Means and Standard Deviations</td>
</tr>
<tr>
<td>5</td>
<td>Recorded Minimum and Recorded Maximum Raw Data Scores</td>
</tr>
<tr>
<td>6</td>
<td>Percentage Arousal: Actual or Predicted, Under and Over Estimates</td>
</tr>
<tr>
<td>7</td>
<td>Arousal/Non-arousal: Intelligence</td>
</tr>
<tr>
<td>8</td>
<td>Arousal/Non-arousal: Sub-test Scores</td>
</tr>
<tr>
<td>9</td>
<td>Arousal/Non-arousal: Verbal IQ - Performance IQ Differences</td>
</tr>
<tr>
<td>10</td>
<td>Arousal/Non-arousal: File Data</td>
</tr>
<tr>
<td>11</td>
<td>Arousal/Non-arousal: Other Variables</td>
</tr>
<tr>
<td>12</td>
<td>Arousal/Non-arousal: Treatment</td>
</tr>
<tr>
<td>13</td>
<td>Results of Multiple Logistic Regressions as Aroused/Non-aroused</td>
</tr>
<tr>
<td>14</td>
<td>Pedophilic Deviance Quotients and Intelligence: Correlations</td>
</tr>
<tr>
<td>15</td>
<td>Pedophilic Deviance Quotients and Intelligence: Cutoff Scores</td>
</tr>
<tr>
<td>16</td>
<td>Participants/Non-participants: Shipley Hartfield Scores</td>
</tr>
<tr>
<td>17</td>
<td>Participants/Non-participants: Other Variables</td>
</tr>
<tr>
<td>18</td>
<td>Arousal/Non-arousal, Age, and Comprehension</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

The focus of the present study is the plethysmographic assessment of pedophilia. It is hypothesized that intelligence mediates the sexual response to sexual stimuli within the plethysmographic assessment.

Plethysmographic assessment of penile tumescence affords a quantitative measure of sexual arousal to sexual stimuli. In general, such assessment allows one to determine an individual's sexual preference. Plethysmographic assessment of penile tumescence has been used to assess sexual preference for more than twenty-five years (c.f., Freund, 1963). The plethysmographic assessment of sexual arousal is undertaken because sexual preference as measured by penile arousal predicts sexual behavior. Measured sexual preference has been related to sexual behavior (either self-confessed or evidenced by criminal convictions) in a number of diverse groups: homosexuals and heterosexuals (Mavissakalian, Blanchard, Abel, & Barlow; 1975) exhibitionists (Abel & Blanchard, 1980; Fedora, Reddon, & Yeudall; 1986); rapists (Barbaree, Marshall, & Lantheir; 1979; Abel, Becker, Blanchard, & Djenderedjian; 1978) and child molesters (Baxter, Marshall, Barbaree, Davidson, & Malcolm; 1984; Quinsey, Chaplin, & Carrigan; 1979).

It is important to note that although plethysmographic assessment is used with a variety of sexually deviant
populations including homosexuals, exhibitionists, and rapists, the technique is used most frequently with child molesters. Homosexuality is no longer classified as a deviant sexual behavior and is consequently less of an assessment priority. Exhibitionists are assessed for research purposes, however, exhibitionists' lack of physical contact with their victims lessens the seriousness of the offense and clinical forensic assessments of their sexual arousal are seldom made.

Recent research has cast doubt on the usefulness of the plethysmographic assessment for assessing rapists. Recent studies with larger samples than those published earlier have not found most rapists to be substantially different from normal men in terms of arousal to violence (Baxter, D.J., Barbaree, H.E., & Marshall, W.L.; 1986). Currently, much of the focus in the use of the plethysmographic technique is on child molesters.

Although child molesters are usually divided into two groups, incestuous child molesters and pedophiles, definitional problems exist and confound such divisions. Some clinicians reserve the diagnosis of pedophilia for those offenders who demonstrate substantial arousal to children upon plethysmographic assessment, regardless of familial relations. Other clinicians reserve the diagnosis of pedophilia for those offenders who have offended with non-related children. The latter definition is the one commonly used by researchers in this field and is the one adopted for the current study. The distinction between pedophiles and incestuous child molesters is
important since pedophiles (by either definition) have substantially higher recidivism rates than do incestuous, non-pedophilic child molesters (Davidson, 1979; Quinsey, 1977). For this reason alone, pedophiles warrant intensive study, and are the population of interest in this study.

**Plethysmographic Assessment of Pedophiles**

The present study focuses on the plethysmographic assessment of pedophilia; that is, upon the detection of sexual responses to child stimuli. In the section that follows, background information about pedophilia is presented. Following this review of pedophilia, the advantages and disadvantages of plethysmographic assessment are discussed. Two major problems, namely, those of insufficient arousal and false negatives are identified. To assist in our understanding of these problems three theoretical models that underlie the interpretation of plethysmographic assessments are proposed. These models are discussed and the evidence supporting each model is presented. The third model proposes that cognitive ability is an important and under researched variable that may be related to the problems of plethysmographic assessment. A rationale for this conjecture is advanced leading to the proposition that intelligence mediates sexual response in a plethysmographic assessment. The experimental hypotheses are then presented.
In the opening overview of pedophiles and pedophilia, then, diagnostic problems associated with pedophilia and information that highlights the seriousness of the crime are provided. The need for accurate assessment of this population is emphasized, and forms the basis for the current study.

Pedophilia

Diagnostic Problems

By definition, pedophilia is a type of philia. Philias are characterized by non-normative arousal-activity patterns. The non-normative arousal activity in pedophilia is arousal to children, and a sexual preference for children is the essence of pedophilia.

The DSM-III (APA, 1980) definition of pedophilia requires: A. that acts or fantasies of sexual activity with prepubertal children are the preferred or exclusive method of achieving sexual excitement; and B. that the child be ten years younger if the individual is an adult, or the age difference is unspecified if the individual is an adolescent.

Although DSM-III is the recognized standard for psychiatric diagnosis, few research studies select pedophilic subjects according to its criteria. One problem with the DSM-III criteria is that the age limits are loosely defined; more precise age limits are preferred by researchers in the field. Quinsey (1977) proposed that pedophiles be defined as men of at least age 16
who engage in sexual acts with children of thirteen years of age or under. The man would also need to be at least five years older than the child. The definition used for the present study was that of Avery-Clark and Laws (1985) and Marshall and Christie (1981) which appears to be gaining consensus within the research community. This definition stipulates simply that sexual acts have been committed with individuals age fourteen or under. This definition is gaining support at least partly because the law differentiates between intercourse with those over and under fourteen years of age.

Another diagnostic problem which is highlighted in the research literature, but not in the DSM-III, is the distinction which may be made between sex with related versus non-related children. It is often assumed that sex with related children occurs primarily because of opportunity, or as a result of marital conflict and stress, but not because of a sexual preference for children. DSM-III puts it this way, "It is preferable not to include in this category (pedophilia) individuals who perform deviant sexual acts when normal sexual outlets are not available to them." In spite of a trend in the literature to view incestuous child molesters as constituting a category separate from that of pedophiles, in many cases the distinction may be more semantic than real. For example, it is recognized that many pedophiles marry in order to increase their access to children and to appear as "normal" as possible. These individuals would probably be classified as incestuous child
molesters according to a familial definition and as pedophiles by a sexual arousal definition.

Behavioral researchers have proposed that labelling an individual as a pedophile should be based on the presence of arousal to children as measured by the plethysmograph and not by reason of the victim's relationship to the offender. Evidence suggests that some but not all incestuous child molesters are aroused by children. Research findings have been equivocal, but generally show that incestuous child molesters do not show as great an arousal to children as do pedophiles. Quinsey, Chaplin, & Carrigan (1979) concluded that inappropriate sexual arousal was less a factor in incest than in pedophilia, and that natural fathers demonstrated less arousal than did step fathers. Similar results were obtained by Marshall, Barbaree, & Christophe (1986). However, Abel, Becker, Murphy & Flanagan (1979) observed no difference between the levels of arousal exhibited by incestuous child molesters and pedophiles. The different findings may have resulted from differences in the populations studied by the researchers. A population that has engaged in incestuous sex for a greater period of time, with more incestuous partners, and with penetration is more likely to demonstrate arousal than a population with more limited time contact, a single partner and no penetration. Learning theory predicts that a greater number of orgasmic experiences leads to a greater conditioned sexual response to the orgasm-inducing object.
There are, then, two distinct ways of identifying an individual as a pedophile. The first, based on behavioral or psychological observation, defines pedophilia as occurring with non-relations and when there is a demonstrated preference for children rather than for adults. The second definition is based upon whether sexual aroused to children exists, regardless of whether sexual acts have been committed with related children or whether sexual acts have also occurred with adults.

For the purposes of this study, pedophiles were defined as those who had committed sexual offences with non-relations. Incestuous child molesters were excluded from the present study. The group of child molesters selected for study are those who are most likely to demonstrate sexual aroused upon plethysmographic assessment.

Pedophiles, even by this definition, do not form a homogeneous group (Quinsey, 1977). Three categorizations can be made simply depending upon the sex of the victim. There are heterosexual pedophiles, homosexual pedophiles and bisexual pedophiles. Bisexual pedophiles are generally considered to be most similar to homosexual pedophiles. Of the three categories, homosexual pedophiles are believed to have the greatest number of victims, to have the highest recidivism rate and to prefer younger victims. Homo-, hetero- and bisexual pedophiles were all accepted into the present study.
Sexual Offense Patterns

Pedophiles have sometimes been dismissed as "gentle fondlers", a label that suggests that they are harmless and not to be taken seriously. It is apparent, however, that this population merits serious clinical and research consideration.

Pedophiles are now acknowledged to be predatory, seeking out great numbers of victims. Estimates of the average number of victims have ranged from seven (Groth, Longo, & McFadin, 1982) to seventy (Abel, cited in Marshall, Earls, Segal, & Drake, 1983). The actual number of victims is difficult to determine since the number of convictions represents only a fraction of actual acts. Rapists (perhaps the closest comparison sex offender group to pedophiles) are judged to have perhaps a fifth as many victims as pedophiles (Abel, cited in Marshall et. al., 1983).

It is not only that pedophiles are believed to have more victims than rapists, pedophiles differ in one other striking way. Unlike rapists, pedophiles are acknowledged to often have multiple sexual contacts with each of their victims. Of all child molesters, incestuous child molesters typically have the lengthiest sexual relationship with a single victim. Groff and Hubble (1984) found that the average length of an incestuous relationship was 2.5 years in a sample of 42 cases. Typically, non-incestuous pedophiles do not have ongoing relationships of similar length, although, ongoing relationships do occur.
this way, most pedophiles engage in an extremely high number of sexual contacts, both because there are numerous victims and because of the tendency to engage in multiple sexual acts with each victim.

The kinds of sexual acts that pedophiles engage in has been a subject of increasing research concern. The myth that pedophiles are 'gentle fondlers' implies both an absence of aggression and an absence of sexual contact other than touching. Research has discounted both of these implications, although surveys of the type of sexual acts engaged in by child molesters are few. The research of Marshall et al. (1983) indicated that 38% of all victims under the age of 12 were vaginally or anally penetrated. Further, this figure increased to 50% when boys 12-14 years of age were considered and to 60% when girls of the same age range were considered. Moreover, regardless of the large percentage of pedophiles already engaging in penetration, there is evidence that those who were not engaging in penetration were thinking about it and desiring it (Marshall, 1985). The evidence is clear: pedophiles do much more than fondle.

An other aspect of the 'gentle fondler' myth is that pedophiles are non-aggressive or non-violent. A number of studies have examined records of injury to victims. Marshall and Christie (1981) studied medical and other records pertaining to 41 convicted pedophiles. They found that only eleven cases (25%) were free from evidence of threats or violence. In twenty-four
cases, physical force was used and, "The degree of force used appeared, in all cases to be in excess of that necessary...." (Marshall & Christie, 1981, p.149). These results are similar to those reported by Christie, Marshall, & Danthier (1978) (cited in Abel et al., 1979) who reported that fully fifty-eight percent of pedophiles had used excessive force during sexual contacts with victims. Forty-two percent of the child victims had sustained noticeable injury.

This section has detailed evidence which indicates that pedophiles often have large numbers of victims and have ongoing relationships with their victims; that they engage in penetration and that they do use excessive force. However, since most of the above cited evidence describes incarcerated populations, the evidence may be biased. That is, it is likely that only the more serious offenders are incarcerated and that the frequency of more serious offences will be over-represented in an incarcerated population. Nevertheless, the evidence does lay to rest the myth of the gentle fondler.

Consideration of the effects of pedophiles' actions upon their victims is beyond the focus of this section. However, it suffices to say that the actions of pedophiles described above are morally and legally wrong because they create pain and suffering in their victims. In recent years much publicity has been given to the prevalence and long lasting effects of child molestation. Prospective studies are difficult, yet the little empirical evidence available indicates that the type of sexual
act and degree of force used, as well as other factors (i.e., relationship to child, parental reaction) are related to the degree of subsequent psychological trauma to the victim.

Pedophiles are difficult to assess (Dreiblatt, 1982). Paper and pencil measures do not reveal a uniquely pedophilic profile. Pedophiles' self-reports usually are unreliable since truthful statements about their sexual preferences, fantasies and behaviors result in longer sentences and delayed parole. Recognition of the disadvantages of relying on a pedophile's description of his sexual tendencies has focused assessors on the objective measurement of sexual arousal, namely plethysmographic assessment. Physiological measurement of sexual arousal is more reliable than self-report in differentiating pedophiles from non-pedophiles (Quinsey, Steinman, Bergersen, & Homes, 1975).

Advantages of Plethysmographic Assessment

Plethysmographic assessment of suspected pedophiles (e.g., individuals who are at the pre-adjudication or pre-sentence phase of legal proceedings) is routinely undertaken in order to determine whether sexual arousal to children exists. The assessment method is also used to measure response to treatment and to inform parole decisions. Sexual arousal to children during assessment has been related consistently, both clinically and empirically, to previous sexual acts with children and is
viewed by assessors as highly indicative of a sexual preference for children. Indeed, the absence of sexual arousal upon a successful plethysmographic assessment towards children in a suspected pedophile has been considered clinically to rule out a diagnosis of pedophilia. Further, reductions in deviant arousal as recorded by plethysmographic assessment usually indicate a changed sexual preference and a significantly reduced likelihood of reoffending against children. In general, sexual arousal patterns are considered to be stable, whether they indicate a preference for adult women, adult men, or for children. Unless interventions aimed specifically at changing sexual arousal patterns (e.g., aversive conditioning), are implemented, no evidence exists to support the notion that sexual arousal patterns will change with the passage of time.

Problems with Plethysmographic Assessment

Notwithstanding its merits, the plethysmographic assessment of pedophiles has two serious methodological and interpretative problems. These are that insufficient arousal may occur in the assessment procedure and that some subjects may be wrongly identified as non-pedophiles. These latter subjects are henceforth referred to as false negatives.

With respect to insufficient arousal, between twenty and forty percent of pedophiles fail to display sufficient arousal to allow their results to be interpreted. Regarding this
problem, Marshall et al. (1986, p. 435) state, "These are serious numbers and represent a serious disadvantage in the use of erectile measures as part of an overall assessment of these men." The absence of arousal makes physiological assessment of sexual preference impossible. In this situation the clinician is restricted to more traditional and less reliable methods of assessment (e.g. interview, paper and pencil measures).

Several explanations for why known pedophiles fail to display sufficient arousal for assessment purposes have been offered. These include successful deliberate faking of arousal patterns, accidental reduction of arousal by cognitive distractions, the use of unprovocative stimuli (Laws & Osborne, 1983) and anxiety (Marshall et al., 1986). However, these explanations have remained speculative to date.

The second problem for clinicians undertaking the plethysmographic assessment of pedophilia is that of false negatives. Some pedophiles exhibit arousal to adults and no arousal to children. Examination of the arousal patterns in these men would lead a clinician to diagnose them as normal and not as pedophilic. These individuals, actually are deviant, but when tested with a plethysmograph they appear normal.

For research purposes, an individual is labelled pedophilic when his profile score exceeds a cut-off score derived from his Pedophilic Deviance Quotient (PDQ). Pedophilic Deviance Quotients are a ratio of arousal to children versus arousal to
adults. A ratio greater than one indicates that more arousal was manifested to child stimuli than to adult stimuli. Different cut-off scores are employed in different studies and result in different hit rates. A cut-off score of .6, for example, results in fewer false negatives. Cut-off scores of .6 and .8 are commonly used in research studies. Lowering the cut-off score too much increases the problem of false positives (identifying someone as deviant who is in fact normal). Clinically, someone with a ratio of .6 is demonstrating considerable arousal to children, but not as much as to adults. In a recent study, however only sixty percent of known pedophiles were correctly identified when a cut-off score of .6 was used (Marshall et al., 1986).

It is not clear why some men who are known pedophiles (e.g., through self-confession or overwhelming proof) are not physiologically identified as pedophiles (that is they do not show arousal to children). Marshall (1985) suggests that higher intelligence may be associated with the display of less deviant arousal. Specifically, Marshall et al. (1986) reported data indicating a negative correlation of .33 between intelligence (as measured by the Ravens Progressive Matrices) and the degree of arousal to children as compared to adults. Thus more intelligent pedophiles exhibited significantly less deviant arousal than did less intelligent pedophiles. Marshall's provisional findings suggest that the outcome of plethysmographic assessment may be a function of intelligence as
well as deviant sexual arousal. The relationship between intelligence and plethysmographic assessment may help account for the large number of pedophiles who produce arousal yet fail to appear deviant. More specifically, it is possible that pedophiles who fail to produce sufficient arousal for assessment purposes differ in their intelligence from the pedophiles that are successfully assessed. However, as yet, no data exist to support this conjecture.

Marshall's findings are particularly important since they suggest that more intelligent pedophiles may be particularly difficult to successfully diagnose using plethysmography. These more intelligent pedophiles may be the most dangerous, because they are most likely to escape detection and therefore have the most potential to continue to victimize children. Clarification of the relationship between intelligence and the outcome of plethysmographic assessment clearly seems a worthwhile goal.

In order to better understand the role intelligence might play in the problems of insufficient arousal and the incorrect classification of pedophiles, a number of theoretical models underlying plethysmographic assessment are presented.

Theoretical Models
Model 1: S - R

Plethysmographic assessment has been traditionally understood within a stimulus-response paradigm. Within this paradigm, a sexual stimulus is presented and an unconditioned sexual response is elicited. The individual's response to the sexual stimulus is dictated by a combination of his biological makeup and his learning history. The classes of stimuli are selected to represent various real-life sexual objects. The response measured is sexual arousal, typically penile tumescence. The response is believed to be invariant, reflecting a combination of a highly learned (conditioned) and biologically-based sexual response. No sexual response is expected to occur to the assessment stimulus other than the response typically elicited by the corresponding sexual object. For example, if an individual experiences sexual attractions to adult males in daily life, then it is expected that sexual arousal will be exhibited to sexual stimuli involving adult males within the assessment situation. Responses to classes of sexual stimuli within the assessment situation are believed to reflect the responses to analogous classes of sexual objects in daily life.

Model 2: S - States - R

A stimulus-response paradigm (Model 1) does not consider mediational factors which may influence the response to a particular stimulus. Subsequent modifications to the
stimulus-response paradigm have been necessary to account for the observation that various changes in subjects' states alter their sexual responses in a systematic way. These state-altering factors include: relaxation, anger, alcohol, distraction and instruction.

Relaxation, or at least the absence of anxiety, is almost universally regarded as facilitating sexual arousal (Beck & Barlow, 1986). Non-sex offenders with sexual dysfunctions are often treated directly by means of relaxation within a behavioral treatment program. In a study using plethysmographic assessment with non-offender subjects, Langevin, Stanford, and Block (1975) found that relaxation instructions resulted in significantly greater arousal to preferred, but not to non-preferred stimuli in both homosexual and heterosexual males. In clinical practice, efforts are made to induce the subject into a relaxed state during the assessment. Thus, although no study has directly examined the impact of relaxation instructions on measured arousal with sex offenders, it is likely that relaxation is a state which facilitates arousal to preferred stimuli with sex offenders as well as with normals.

The effects of anger upon measured arousal have likewise been studied with normals, but not with sex offenders. Yates, Barbarée, and Marshall (1984) demonstrated that when angered, normals produced rapist-like sexual arousal profiles. In their experiment, one experimental group was angered by a female confederate prior to testing. The anger resulted in greater
arousal to rape cues. Although the effects of anger upon arousal within a pedophilic population are untested it is likely that anger would serve to increase pedophiles' arousal to aggressive cues.

Alcohol is another factor whose effects on sexual arousal have been studied with normals, but not with sex offenders. Alcohol ingestion has not been shown to consistently effect arousal, although all studies finding an affect report that normals' sexual arousal profiles become more similar to those of rapists following the ingestion of alcohol (e.g. Barbaree, Marshall, Lightfoot-Barbaree, & Yates, 1979). The induced belief that subjects had consumed alcohol also has been reported to have had a greater effect on measured arousal than actual consumption of alcohol (Wilson & Lawson, 1976; Bridell, Rimm, Caddy, Klawitz, Sholis, & Wunderlin, 1978). It may be that alcohol, or the belief that one has consumed alcohol, have disinhibiting effects on individuals. Thus, although the exact mechanism of the effects of alcohol upon arousal are not known (psychological or pharmacological) and have not been specifically studied with pedophiles, alcohol is commonly believed to have a disinhibiting effect upon existing sexual arousal.

The effects of distraction upon sexual arousal are commonly known. Distraction is believed to prevent erections in impotent men and is used as an aid to reduce premature ejaculation in sexually dysfunctional men. The effects of distraction on sex
offenders' arousal have been studied in two contexts: first, in attempts to measure subjective arousal concurrently with measuring physiological arousal, and second, in attempts to measure the ability of sex offenders to fake arousal patterns.

Attempts to measure subjective arousal concurrently with physiological arousal have emerged through synchrony-dysynchrony theory (Hall, Binik, & Di Tommaso, 1985). Briefly, this theory holds that behavior, physiological response and self-report do not necessarily covary. In one study, subjects were asked to continue twisting a dial calibrated from 0 to 100 throughout the plethysmographic assessment procedure, as part of an attempt to obtain moment by moment estimates of arousal. This ongoing process of self-monitoring produced decrements in the measured physiological arousal (Wincze, Venditti, Barlow, & Mavissalakian, 1980). In short a concurrent cognitive task reduced arousal. Similar reductions in physiological arousal attributable to a competing cognitive task have been reported by other researchers (Farkas, Sine, & Evans, 1979; Abrahamson, Barlow, Sakheim, Beck, & Athanasiou, 1985). Probably the best demonstration of the effects of cognitive distraction upon arousal is that by Geer & Fuhr (1976). Using four levels of complexity of a concurrent cognitive task they demonstrated a strong negative relationship between the complexity of the cognitive task and the level of arousal observed. Increased complexity of task resulted in decreased arousal.
The effects of distraction upon arousal have also been studied in the context of the deliberate faking of plethysmographic assessments. Faking of sexual arousal has been studied sufficiently such that several broad conclusions can be drawn. Early studies with small numbers of normal subjects (Henson & Rubin, 1971; Laws & Rubin, 1969) suggested, and later studies with larger numbers of sex offenders confirmed, that some normals and sex offenders can suppress arousal when they are instructed to (Wydra, Marshall, Earls, & Barbaree, 1983; Malcolm, Davidson, Marshall, 1985). However, pedophiles may not be able to suppress as effectively as rapists (Avery-Clark & Laws; 1985).

Individuals have reported that they controlled their arousal either by cognitive (commonly used), muscular or penile-manipulative (uncommonly used) strategies. The muscular strategy involves tensing the muscles in the lower abdomen and can be detected during testing by changes in breathing pattern. The manipulative strategy involves touching the penis either to create arousal or suppress it. Penile manipulation to create arousal produces a recognizable pattern of arousal and can also be detected. Manipulation to suppress arousal is prevented in some laboratories by requiring that a switch be depressed by each hand. Failure to depress the switches results in a signal to the tester.

Cognitive strategies have been used both to suppress deviant arousal and to create appropriate arousal. To suppress deviant
arousal, asexual stimuli are focused upon, such as lyrics to popular songs, and multiplication tables (Laws & Holman, 1978). In contrast, to create arousal, erotic fantasies are entertained (Henson & Rubin, 1971).

Attempts have been made to decrease the use of cognitive strategies as a means of faking. The subject is asked to describe the stimuli subsequent to their presentation. However, faking is still possible even though this requirement is imposed (Henson & Rubin, 1971).

The research demonstrating that competing cognitive tasks reduce arousal is consistent with the research demonstrating that sex-offenders use competing cognitive tasks to alter their arousal patterns. Both lines of research demonstrate that non-stimuli related thoughts decrease the normally occurring stimulus response chain. At first glance, the ability of sex offenders to fake arousal patterns suggests that the assessment technique is not useful. However, the general opinion is that most sex offenders do not try to alter their arousal pattern during testing (Wydra et al., 1983, Avery-Clark & Laws, 1985).

In summary, a number of state factors have been demonstrated to affect sexual arousal, including relaxation, anger, alcohol, cognitive distraction and intention. These are all state factors in that they may not be predictably present in the same individual from one testing session to the next. The potential effects of each factor are clearly such that a simple
stimulus-response model cannot account for all of the data. A model of plethysmographic assessment must account for and include the state of the individual. Such a model still holds that real world behavior is related to the sexual arousal profile on plethysmographic assessment, but recognizes that the results of a sexual response assessment are variant and reflect the individual's changing states.

Model 3: \( S - Cognition - R \)

Recently cognitive factors have become increasingly prominent in the theory and application of behavioral psychology (see, for instance, Beck, 1979; Bandura, 1964). Cognitive factors are hypothesized to influence physiological arousal and to be stable without consistent and concerted intervention. In light of this trend it seems appropriate to consider cognitive factors in attempting to understand the results of plethysmographic assessment.

The cognitive-behavioral model proposes that cognitive appraisals mediate behavior in all situations. Thus, an individual's physiological response can be altered through modification of the cognitive appraisal being undertaken. Cognitive behaviorists employ this model in accounting for the physiological and psychological states and traits of an individual. For example a depressive state may result from applying negative cognitive appraisals in a given situation, while a depressive trait may result from the habitual
application of similar negative appraisals. The cognitive-behavioral model predicts that cognitive factors of a stable (habitual) nature can be expected to modify a response to a given stimulus in a predictable way. Since individuals appraise situations differently, they can be expected to respond with different physiological responses.

The cognitive-behavioral model, applied to plethysmographic assessment, predicts that the cognitive appraisal of the assessment situation may significantly alter the assessment outcome. Sexual arousal within the assessment situation may, or may not resemble sexual arousal and behavior in the real world, where different cognitive appraisals may occur. Given this model, the cognitive ability and style used by the individual in appraising the assessment situation, through which he may subsequently modify his sexual response are matters of considerable interest. It is likely that not all cognitive styles are equally efficient at modifying the fundamental sexual drive. Therefore, according to Model 3, the mental abilities with which an appraisal of the plethysmographic assessment situation is made are crucial, as the assessed sexual arousal is a function of the cognitive appraisal and related cognitive functions as well as of sexual preference.

Intelligence is acknowledged to be the most stable of individual cognitive characteristics and it can be measured easily with an impressive degree of reliability.
Intelligence has not, however, been systematically studied as a factor which may mediate sexual arousal in plethysmographic assessments. As mentioned above, Marshall et al. (1986) reported a correlation of .33 between the Pedophilic Deviance Quotient and intelligence as measured by the Ravens Progressive Matrices, indicating a possible relationship between intelligence and assessed sexual arousal.

In considering this finding, however, Marshall stated that a simple linear function in which PDQ was plotted against IQ would not reveal all that the data had to show. Rather he found that grouping his subjects into those with IQs greater than and less than 80 to be most useful. In his sample, pedophiles with IQs below 80 were very likely to appear deviant on plethysmographic assessment, while those with IQs greater than 80 were much less likely to appear deviant. Although not stated in his paper, it can be inferred from Marshall's data that only fifty percent of those with IQs greater than eighty were correctly identified as pedophiles by using a standard cut-off score. In contrast it would appear that all of the pedophiles with IQs less than eighty were correctly identified.

A number of studies have reported on the intelligence of pedophilic subjects. The results of a number of these studies are summarized in Table 1. As can be seen, pedophiles' IQs have fallen within the average range, slightly below a mean IQ of 100. In this respect, pedophiles are similar to other incarcerated offender populations who have been reported to have
average IQs in the 90s (Ruff & Templar, 1976; Segal & Marshall, 1985).

Some researchers have examined the neuropsychological functioning of pedophiles. Scott, Cole, McKay, Golden, & Liggett (1984) examined the neuropsychological performance of pedophiles and other sexual assailters using the Luria-Nebraska Neuropsychological Battery. They found that 36% of their pedophilic sample (5 of 14) showed evidence of cognitive impairment. Hucker, Langevin, Wortzman, Bain, Handy, Chambers, & Wright (1986) tested a larger sample of pedophiles (N=39) using the Luria-Nebraska Neuropsychological Battery, the Reitan Neuropsychological Battery, the WAIS-R and CT scans. Results from the different test batteries were not highly correlated, however 67% of the pedophiles were impaired on at least one of the test batteries. The results were interpreted as reflective of left temporo-parietal impairment. The authors conclude that brain pathology may be relevant to an understanding of pedophilia.

Neither of the studies examining neuropsychological deficits in pedophiles undertook sexual arousal assessments so that no information on the relationship between neuropsychological deficits and the outcome of a plethysmographic assessment is available. Because sexual disinhibition can occur subsequent to a head injury, it is conceivable that the demonstrated cortical impairments in some pedophiles do contribute to disinhibited sexual behavior. Impaired mediational processes would also
### Table 1

**Pedophiles' intelligence: Results of previous studies**

<table>
<thead>
<tr>
<th>Study</th>
<th>Subjects</th>
<th>IQ Measure</th>
<th>VIQ</th>
<th>PIQ</th>
<th>FSIQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mohr (1964)</td>
<td></td>
<td>WAIS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24 hetero-</td>
<td>20 homo-</td>
<td></td>
<td>100.5</td>
<td>99.5</td>
<td>100.0</td>
</tr>
<tr>
<td>3 bi-</td>
<td>47 total</td>
<td></td>
<td>99.0</td>
<td>97.4</td>
<td>98.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Beta Test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seagull &amp; Marshall (1985)</td>
<td>20 total*</td>
<td>Quick Test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marshall, Barbaree, &amp; Christie (1986)</td>
<td>40 hetero-</td>
<td>Ravens Progressive Matrices</td>
<td>91.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hucker, Langevin, &amp; Wright (1986)</td>
<td>15 hetero-</td>
<td>WAIS-R</td>
<td>94.5</td>
<td>96.9</td>
<td>94.9</td>
</tr>
<tr>
<td></td>
<td>14 homo-</td>
<td></td>
<td>93.1</td>
<td>95.6</td>
<td>93.6</td>
</tr>
<tr>
<td></td>
<td>10 bi-</td>
<td></td>
<td>107.7</td>
<td>102.5</td>
<td>106.1</td>
</tr>
<tr>
<td>Bain, Hardy, Chambers, &amp; Wright (1986)</td>
<td>39 total</td>
<td></td>
<td>97.4</td>
<td>97.9</td>
<td>97.3</td>
</tr>
</tbody>
</table>

*Breakdown by type of subject not provided*
decrease the likelihood that arousal would be inhibited within the assessment situation.

It is interesting to speculate on the ways in which cognitive ability might mediate physiological arousal. It may be for instance, that more intelligent pedophiles understand that exhibiting deviant arousal will lead to negative consequences, and therefore strive to inhibit their arousal. Conversely, it may be that all pedophiles regardless of their IQ are equally motivated to inhibit their deviant arousal, but that higher IQ pedophiles are better at doing so, because of their greater intelligence. Also, it may be that higher IQ pedophiles may be thinking more and thereby reducing their arousal by distracting themselves from the stimuli.

Finally, it is possible that it is not overall IQ that is relevant, but rather the ratio of Performance to Verbal IQ. For instance, it is known that individuals with impulsive behavior problems often have a higher Performance IQ than Verbal IQ score. Thus, physiologically identifiable pedophiles may have less verbal ability, and be generally poorer at verbally directing their behavior. In general, they may be more externally focused with their behavior stimulated more by the immediate environment rather than by their cognitive processes.
Overview of the Current Study

The problems of insufficient arousal and false negatives discussed earlier can be considered in light of three models: Model 1: S - R; Model 2: S - States - R; and, Model 3: S - Cognition - R. Adherents to Model 1 might maintain that the sexual stimuli used during assessment were insufficiently strong and not representative of real world stimuli. This approach to the problem of insufficient arousal is currently being pursued by Farrell and his colleagues, who are attempting to produce powerful stimuli (combination slide and audio) in order to achieve arousal in all subjects. An adherent of the second Model might consider that too much anxiety within the individual is responsible for his failure to become aroused, and that therefore increased attention should be paid to relaxing individuals within the assessment situation. An adherent of Model 3 might want to know more about the individuals' cognitive abilities and their systematic effects upon the results of plethysmographic assessments.

According to both Models 1 and 2, the problems of insufficient arousal and false negatives are problems of error variance. Attempts are made to reduce the error variance as much as possible. However, no statements can be made about those individuals whose results are uninterpretable because of the aforementioned problems. From the viewpoint of Model 3, the two problems may be the result of a predictable and measurable
characteristic of the individual. This characteristic may have treatment and prognostic value as well as being predictive of response within the assessment situation.

The following study is a test of the applicability of Model 3 to plethysmographic assessment. Both empirical and theoretical rationales reviewed previously suggest cognitive abilities may influence the outcome of plethysmographic assessment. The current research will examine whether cognitive ability mediates sexual responsiveness within an assessment situation.

Therefore, it is predicted that:

1) The average IQ of the sample will be between 90 and 100.

2) Twenty to forty percent of the sample will fail to produce sufficient arousal such that their results can be interpreted.

3) Using a Pedophilic Deviance Quotient score of .6, approximately forty percent of subjects will be classified as false negatives.

4) Pedophiles' intelligence will be related to the outcome of plethysmographic assessment.
CHAPTER II

METHOD

Subjects

The subjects were thirty-nine convicted pedophiles. Thirty-seven subjects were incarcerated at the time of testing, and two subjects were on parole. All subjects had engaged in sexual acts with one or more children to whom they were not related and with whom they did not reside. All but two subjects had been convicted of sexual offences against children or adolescents aged fourteen or under. The other two subjects had been convicted of murder, although in both cases the subjects had committed sexual acts during their offence(s). Every subject in the study acknowledged he had committed sexual acts with children as charged. The sexual acts committed ranged from fondling to intercourse. Hetero-, homo- and bisexual pedophiles were all accepted into the study.

Subjects were drawn from three federal penitentiaries located in British Columbia: the Regional Psychiatric Center, Mountain Institution, and Kent Institution; and from two sex offender treatment groups run by the National Parole Board of Canada. Because the circumstances of the subjects varied greatly from one subject to another a brief description of the settings from which subjects were drawn is given below.
The Regional Psychiatric Center is a maximum security penitentiary and an accredited hospital. Most sex offenders in this institution participate in an intensive sex offender treatment program. The two year program consists of a minimum of twenty hours of group therapy per week. Rapists, pedophiles and incestuous child molesters are treated. The sex offenders are taught that they will always experience deviant arousal but must learn not to act upon it. Aversive conditioning procedures aimed at reducing deviant arousal are not part of the treatment program. Plethysmographic testing is not used routinely for either assessment or treatment. Names of eighteen potential subjects were provided by correctional staff and fifteen of these agreed to participate in the study. Twelve of the fifteen were members of the sex offender treatment program.

Mountain Institution is a medium security penitentiary housing a population which is ninety three percent sex offenders. Individual treatment is provided to inmates who seek it by staff psychologists. Plethysmographic testing is not routinely used for either assessment or treatment, although it is occasionally conducted for parole review reports. Eighty potential subjects were identified by reviewing all psychology department files. Twenty-two agreed to participate in the study.

Kent Institution is a maximum security penitentiary. All sex offenders are housed in the protective custody unit. They are locked in their cells twenty three hours per day and separated from the general prison population. Psychological treatment at
this institution is minimal. Plethysmographic testing is not routinely used for either treatment or assessment. A list of three potential subjects was provided by the staff psychologist. None of these agreed to participate in the study.

Acceptance rates varied according to institution. The acceptance rate at the Regional Psychiatric Center (RPC) was the highest, because only inmates who viewed their pedophilia as a problem were admitted into the treatment program. Therefore the inmates at RPC perhaps viewed research into pedophilia favorably and so volunteered to participate at a high rate.

The National Parole Board of Canada runs treatment groups for sex offenders released on parole. Attendance by parolees is compulsory. The groups run for one and a half hours per week and contain rapists, incestuous child molesters and pedophiles. Presentations about the study and requests for participation were made to both groups in which potential subjects were known to exist. Seven parolees in each group heard the presentation and in each group one parolee volunteered. It is not known how many in each group actually would have met the criteria for the study, that is, were pedophiles rather than rapists or incestuous child molesters, so the acceptance rate can not be computed.

A total of ninety-eight potential subjects were identified at the three institutions; thirty-seven of these agreed to participate in the study. This represents an acceptance rate of
40% and limits the representativeness of the sample. Two additional subjects were recruited from the Parole Board groups. Subjects were paid four dollars for participating.

**Apparatus**

The plethysmographic assessment apparatus and response measurement procedure were as follows. Subjects were seated in a reclining chair located in a partitioned room which provided a substantial visual barrier between the experimenter and the subject. Voice communication was unrestrained. Testing occurred in a total of four rooms, two at Mountain Institute and one each at the other two institutions. The experimental settings were made as similar as possible; however, the rooms did differ in terms of size and soundproofing. With respect to conducting plethysmographic assessments, one room designed for biofeedback purposes was excellent; the other three rooms were adequate.

**Penile Tumescence Measurement.** Plethysmographic assessment of penile tumescence has been used to assess sexual preference for more than twenty-five years. Unlike other measures of sympathetic arousal (e.g., heart rate, pupillary dilation), penile measures are specific to sexual arousal and do not occur concurrently with other strong emotional states or startle reflexes (Zuckerman, 1971; Bancroft, 1971). Penile arousal does not correlate with autonomic arousal.
Measurement of the response to sexual stimuli is accomplished by monitoring changes in one of three dimensions of penile tumesence: volume, circumference or length. Volumetric measurement devices consist of either an airtight metal (McConaghy, 1967) or glass (Freund, 1961, 1963) cylinder which is fitted over the penis. Changes in penile volume changes the air pressure within the device. Volumetric devices are the most sensitive measures of mild sexual arousal (Abel & Blanchard, 1975).

Although most earlier research was conducted using volumetric devices (c.f., Freund 1963, 1965, 1967), most contemporary researchers employ circumferential measures, which are less bulky and less expensive. Two circumferential devices are in common use: the Barlow or mechanical gauge, and the mercury-in-rubber strain gauge (Laws, 1977). The Barlow gauge (Barlow, Becker, Lettenberg & Agras, 1970) is made of spring steel, while the mercury-in-rubber strain gauge consists of a slim column of mercury encased in rubber. Both devices measure changes in penile circumference. A comparison of the operating characteristics of the two gauges revealed no statistical difference in the manner in which the gauges measured penile tumescence (Laws, 1977). The mercury-in-rubber strain gauge is preferred because it is less obstructive and less expensive (Laws, 1977).

The measurement of changes in penile length holds promise since penis length will alter up to fifty percent before
noticeable changes in circumference occur (Earls & Marshall, 1983). However, at present the technology for penile length measurement remains experimental. Therefore, the preferred means of measuring degree of sexual arousal remains the mercury-in-rubber strain gauge (Earls, 1986). Although not as sensitive as the volumetric gauge, the mercury-in-rubber strain gauge has satisfactory sensitivity for research and clinical purposes.

The operating characteristics of the mercury-in-rubber strain gauge have been examined in two studies. Farkas, Evans, Sine, Eifert, Wittlieb & Vogelmann-Sine (1979) studied the reliability and validity of the gauge. They found test-retest reliability over one week to be .94 for the basal circumference measure and .75 for the maximum circumference measure. Both of these reliability figures were judged to be satisfactory. The lower reliability of the maximum circumference measure (full erection) reflects individual variations in fullness rather than measurement error. The validity of the mercury-in-rubber strain gauge was established by comparing changes in circumference with changes in the angle of penile erection. As expected, circumference change and angle change covaried, providing support for the use of circumference measurement as a measure of erectile response.

Earls and Jackson (1981) studied the effects of temperature on the mercury-in-rubber strain gauge. The surface temperature of the penis varies by as much as six degrees Celsius (Fisher,
Gross & Zuch, cited in Earls and Jackson, 1981) and it was thought that a change in temperature might effect the electrical properties of the gauge. Upon study, the gauge was found to be unaffected by changes in its temperature from 0 to 50 degrees Celsius.

In summary, the mercury-in-rubber strain gauge is the assessment measure of choice for monitoring changes in sexual arousal. It is known to be reliable and valid and to be unaffected by temperature. Additionally, its results are highly correlated with other more cumbersome and expensive measures of penile tumescence.

In the present study, penile responses were measured by means of mercury-in-rubber strain gauges which the subject fitted onto the shaft of his penis. The mercury-in-rubber strain gauge works as follows: once the gauge is fitted on an individual, a change in the circumference of the penis changes the circumference of the gauge, causing a change in the diameter of the mercury and a consequent change in the mercury's electrical resistance. The resistance changes, monitored as changes in voltage output, are amplified by means of a Pacific Design and Instrument (PDI) plethysmograph. The voltage output is sampled every two seconds by an analog-digital card and a digital equivalent is stored in an Apple IIe clone (Mega IIe).

Calibration was made of each mercury-in-rubber strain gauge prior to use and periodically throughout use. This process
assured that a linear relationship between change in circumference and voltage output was maintained throughout the entire range within which penile measurements were made.

Stimuli

Plethysmographic assessment measures sexual response to different classes of stimuli. The content of the stimuli influences subsequent arousal. For example, a male heterosexual achieves greater arousal to pictures of nude women than to nude men. The medium through which stimuli are presented also influences the level of sexual arousal. Three mediums are commonly employed: slides, audio-tapes, and video tapes (Abel et al., 1975).

Slides are the least arousing of the three modalities. Slides are static and do not evoke as much response as evolving sequences of sexual behavior. Although the slide content can be made very specific, slides do not typically evoke large amounts of arousal. As a result, slides are used with sex offenders primarily for assessing arousal to very specific stimuli. For example, with pedophiles, slides may be used to determine the preferred age of the child victim.

Audio tapes are the next most arousing stimulus modality. These are commonly used to define broad categories of preference (e.g. adult, child; male, female; coercive, non-coercive). While the taped descriptions of sexual behavior are often very
explicit, it is difficult to assess specific age preferences using the audio tapes, since the use of descriptive terms such as "small" is open to individual and widely varying interpretations. The advantages of audio tapes are that behavior and feelings can be described.

Videotapes are the most arousing of all modalities of stimuli. Individuals' sexual preferences cannot usually be assessed with video tapes since regardless of content, maximum arousal is often achieved to each stimulus, creating a ceiling effect. The extremely arousing qualities of video tapes make them excellent for inducing maximum arousal in an individual and they are sometimes used for this purpose.

Arousal is affected by the duration of the stimulus presentation as well as by the sensory modality of the presentation. Duration of the stimulus presentation is most important when audio tapes are used; significantly greater arousal is achieved with four minute tapes than with two minute tapes (Avery-Clark & Laws, 1984). Thus, in undertaking a sexual arousal assessment, modality of stimulus presentation and length of exposure must both be taken into consideration. In general, longer exposures and more intense modalities produce the greatest arousal.

In the present study the stimuli were six audiotapes, three describing sex with adults and three describing sex with children. Five of these tapes were selected from a group of
twelve standardized tapes used by a number of research teams. The remaining tape was prepared specifically for this study. Order of presentation of the tapes was counterbalanced, except that no two tapes of the same type followed one another and an adult tape was always played first. Counterbalancing was necessary since previous research has demonstrated serial order effects on arousal (Kolarsky & Madlafousek, 1977). An adult tape was played first in order to make the test situation as socially acceptable as possible.

Child and adult tapes were designed to be identical in the presentation of sexual material. Child and adult tapes were considered to be generally equivalent in terms of their ability to generate arousal. The primary difference between the adult and child tapes was in the adjectives and nouns used to describe the age and size of the sexual partners. All tapes described consenting sex. Some tapes contained references to slight coercion, since in general mild coercion facilitates arousal. Two sets of tapes were used: one in a female voice describing sexual activities between a male adult and a female partner (child or adult) and one in a male voice describing sexual activities between a male adult and a male partner (child or adult). The sex of the partner in the tapes used was determined by the sex most frequently victimized by each subject. If a victim sex preference was not evident from file information, subjects stated which victim sex they preferred. Each tape was over six minutes in length, in keeping with the Avery-Clark and
Laws (1984) finding that four minute tapes elicited higher levels of arousal than did two minute tapes. All of these tapes have been shown to be adequate for the elicitation of sexual arousal (Marshall et al., 1986).

Audiotaped stimuli were presented by means of a Panasonic cassette player and headphones.

**Procedures**

This study met the ethical guidelines established in consultation with the Simon Fraser University Ethics Review Board (Dr. Calvert, Chairman); and its procedures were also reviewed by the Regional Research Committee of the Corrections Service of Canada, Pacific Region (James Murphy, Deputy Commissioner); the Medical Staff Ethics Committee, Regional Psychiatric Center, Pacific (Dr. Moffat, Chairman); and an external reviewer, (Dr. Beiser, Head, Division of Social and Cultural Psychiatry, University of British Columbia).

A list of potential subjects was derived from an exhaustive review of all available inmate files as well as through consultation with clinical staff. Potential subjects were individually interviewed in the room where testing was to occur. The equipment (chair, screen, plethysmograph, computer, tape recorder) was in place and the experimental procedures were explained in detail. Prospective subjects' questions were answered accurately and honestly. Each potential subject's file
was reviewed to confirm that pedophilic acts had occurred. Additionally, potential subjects were asked whether they had committed the offense for which they were convicted. Only subjects admitting to sexual acts with children were accepted into the study.

Subjects indicating that they would consider participating in the study were recalled a minimum of forty-eight hours after the initial interview. At this time, subjects were asked if they still wished to participate in the study. Any inmate declining at this point was not assessed. Those agreeing to participate were tested immediately or given an appointment for testing.

Testing began with the subjects reading a memo from the director of the institution. In accordance with an Ethics Review Board request, the memo (Appendix A) stated that the Corrections Services of Canada would not make any reward or penalty contingent upon inmates participating or not participating in the research study. The memo also disclaimed any therapeutic benefits from the procedure.

After reading the correctional services memo, subjects read and signed the consent form (Appendix B). Prior to signing the consent form, subjects were asked if they understood the consent form or if they had any questions. In cases where there was doubt as to the subject's ability to read, the consent form was either read aloud, or the points on the consent form were explained and discussed. All questions were answered, and the
Confidentiality of the results was reiterated (as it had been stated in the consent form and in the initial interview).

Considerable effort was made to establish rapport with the subjects. Rapport building was attempted in order to elicit agreement to participate in the study and to promote acceptance of the study within the general institution population. Most importantly however, rapport was sought with the goal of allowing the subjects to be as relaxed and open to the sexual arousal assessment as possible. The efforts undertaken to ensure confidentiality and promote rapport are considered important, since both factors may affect a subject's willingness to become aroused.

The entire experimental procedure was completed in one session with the sexual arousal assessment completed prior to the psychometric testing.

Sexual Arousal Assessment

After the consent form was signed the sexual arousal assessment began. Subjects were instructed how to place the strain gauge upon the penis. They were then told to put on the strain gauge, sit in the reclining chair and cover themselves with a towel. The towel served to keep the subjects warm and decrease their visual awareness of their erection.

Once the subject was seated, the output of the strain gauge was checked to make sure that it was within normal limits and to
ensure that the gauge had been put on correctly. The subject was asked whether he was comfortable and upon acknowledgement was handed a number of adult magazines with which to stimulate himself. Heterosexual men viewed female erotica and homosexual men viewed male erotica. Subjects were told to inform the experimenter when three quarters of an erection was achieved.

Twelve minutes were allowed for the production of arousal. Some subjects became fully aroused within seconds, while other subjects did not become aroused within the twelve minutes. If no arousal was observed on the monitor after six minutes, subjects were asked if they were comfortable, encouraged to search the erotica for sequences they found arousing and told that it was important that they become aroused.

After either three quarters of an erection was reported, or twelve minutes had passed, subjects were asked to put aside the magazines and put the headphones on. At this point subjects were told:

"You will be hearing six tapes, three describing sexual activities with adults and three describing sexual activities with children. Each of the tapes will be about six minutes long. These tapes are designed to be arousing so don't be surprised when you find yourself becoming aroused. Treat the tapes as sexual fantasies. Do what you can to get yourself involved with the fantasies. Imagine that you're there being aroused by what is going on. Make the fantasies as vivid and as real as you can. When you
find yourself getting aroused let it happen and don't try to control it. So relax and let it happen. You will need to tell me if the volume is loud enough and I want you to tell me when each tape is over. All right? Then let's begin."

The instructions take into consideration the research on the effect of focusing on imagined arousal as well as on stimulus scenes (Dekker, Everaerd, & Verhelst, 1985). The instructions also take into account the research on the effects of instructions on arousal (e.g., Abel, Barlow, Blanchard, & Mavissakalian, 1975) and the effects of relaxation on arousal (Langevin, Stanford, & Block, 1975).

At this point the first tape was played and the subject was asked whether the volume level was comfortable. Appropriate adjustments were made to the volume in accordance with subjects' instructions.

Each subject's arousal was monitored and stored by the computer software developed for this study. Manually recorded notes on the subject's demeanor and arousal levels were also made.

The tapes were played sequentially, while allowing time for return to baseline arousal when necessary. After each tape, subjects were asked how aroused they had been to that tape.

If no arousal was observed to the first tape, the subject was asked if he was comfortable and if he had been able to get
involved with the fantasy. The conversation was to encourage the subject to get as involved as possible with the stimulus material. Subjects who had become aroused to the first tape were not further encouraged.

When all stimulus tapes had been played the subjects were advised that the sexual arousal assessment was over. After the gauge was removed and subjects had dressed they were seated at a desk for the remainder of the testing. As the sexual arousal assessment was potentially disturbing, subject's reactions were discussed. Subjects' reactions included interest, disinterest, disgust, anger, sadness, and enjoyment. Few subjects had intense emotional reactions to the stimuli, although one man cried in response to an adult tape that reminded him of his wife.

**Psychometric Measures**

During the next phase of testing a series of questionnaires was administered, beginning with the screening questionnaire (Appendix C). The screening questionnaire included questions regarding subjects' previous experience with plethysmographic assessment and with treatment for their sexually offensive behavior.

The following questionnaires were then administered: the State Anxiety Inventory, the Trail Making Test, the Crowne-Marlowe Social Desirability Scale, and the Wechsler Adult Intelligence Scale - Revised.
State Anxiety Inventory

The State Anxiety Inventory (SAI) is part of the State-Trait Anxiety Inventory (STAI) (Spielberger, Gorsuch, & Lushene, 1970). The SAI (form X-1) consists of twenty statements each followed by a four point response scale ("not at all", "somewhat", "moderately", "very much so") which collectively measure an individual's state anxiety level. The SAI has been shown to have adequate reliability and validity and to be responsive to mood changes.

This inventory was included because anxiety may interfere with sexual responding (Masters & Johnson, 1970). Inclusion of the SAI permitted comparisons to be made between the anxiety level of individuals who did and did not produce arousal during the plethysmographic assessment.

Trail Making Test

The Trail Making Test consists of two parts, A and B. Trail Making Test A consists of twenty-five numbers (1-25) set in circles scattered on a sheet of paper. The subject's task is to connect all the numbers in order as quickly as possible, without taking the pencil from the paper. Trail Making Test B consists of thirteen numbers (1-13) and twelve letters (A-L) set in circles scattered on a sheet of paper. The subject's task is to join numbers and letters alternately and in an ascending fashion (1,A,2,B,...) as quickly as possible. Subjects are timed on each part of the test and error scores are derived from mistakes made.
in completing the trails. Reliability on Trail Making Test A as measured by the coefficient of correlation is .78 at 6 and 12 month intervals. Test-retest reliability on Trail Making Test B as measured by the coefficient of concordance is .67 at similar intervals.

The Trail Making Test is commonly used in neuropsychological practice and is considered to be one of the most sensitive tests of brain damage (Lezak, 1983). Therefore, the test was included as a screening measure of neuropsychological dysfunction. The WAIS-R is also commonly used in making statements about the presence or absence of brain dysfunction. However the Trail Making Test measures a substantially different function than does the WAIS-R as the correlation between the two tests is only .50 (Marceau, 1986). Individuals with no deficits on the WAIS-R may be significantly impaired on the Trail Making Test. The deficits on the Trail Making Test are related to difficulties in dealing with more than one stimulus or thought at a time; that is, cognitive flexibility in the course of an ongoing activity (Lezak, 1983).

Crowne-Marlowe Social Desirability Scale

The Crowne-Marlowe Social Desirability Scale (CMSDS) is a thirty-three item, True-False, self-report scale designed to measure an individual's need for social approval. The scale is made up of items based on culturally sanctioned and approved behaviors which may be aspired to, but which are seldom
achieved. The internal consistency coefficient for the scale, using the KR-20 has been reported as .88 and the test-retest reliability has been reported as .89 (Crowne & Marlowe, 1960).

The Crowne-Marlowe Social Desirability Scale was included because subjects who failed to show arousal may have had a greater need for social approval than those who did show arousal. The need for social approval may have led to increased efforts to appear non-deviant.

The majority of subjects were administered the CMSDS after the Trail Making Test; however the subjects at the Regional Psychiatric Center were not administered the scale until one month after completion of the other testing. This delay occurred because the decision to include this questionnaire was made after testing had been completed at that institution.

Wechsler Adult Intelligence Scale - Revised

The Wechsler-Adult Intelligence Scale-Revised (WAIS-R) consists of eleven subtests and three summary IQ scores based on the subtest scores. The WAIS-R is the most commonly used test of intellectual ability (Lezak, 1983). The subtests of the WAIS-R measure different aspects of intellectual ability. The WAIS-R, and its predecessor the Wechsler Adult Intelligence Scale (Wechsler, 1955) have been thoroughly researched. It is a highly reliable and valid test of intelligence.
Some of the subjects in this study had been previously administered the WAIS-R as part of their correctional service psychological testing. These subjects were not retested because repeat testing may have artificially increased subject's scores through practice effects. No results were more than three years old and no appreciable change in WAIS-R scores is to be expected over this period (Matarazzo, Carmody, & Jacobs, 1980). Subjects whose last intellectual assessment employed the WAIS were retested using the WAIS-R. Over fifty percent of subjects were tested with the WAIS-R expressly for this study.

After the subjects completed the psychometric testing they signed a receipt for four dollars. This amount was subsequently deposited into their institutional savings account as payment for participating in the study. Subjects had been informed of the payment during the initial presentation of the experimental procedures.

All procedures took between two and a half and four hours to complete with an average time of approximately three and a half hours. Subjects' questions about the study were answered before the session was terminated. Interested subjects were told that they would be recalled to discuss their results when the scoring was completed.

Upon recall, subjects were shown their arousal graphs and were given a verbal summary of the other test results. Subjects received this feedback approximately one month after completing
the testing. Approximately one-half hour was spent in giving feedback to most subjects. Efforts were made to help the subjects integrate the information into their notions of themselves.
CHAPTER III
RESULTS

The Results chapter is divided into three sections. The first and third sections report results of the various analyses, with the first section detailing descriptive analyses and the third section detailing relationships between variables. The second section is primarily technical in nature and describes the solutions adopted for various methodological and statistical problems. The reader more interested in the results of the analyses than the underpinnings of the analyses may wish to focus on sections one and three. The more technically oriented reader may appreciate the rationale given in the second section.

Section 1: Descriptive Analyses

For ease of presentation and discussion the descriptive data are divided into four sections: File data, Intelligence data, Other Psychometric data and Sex of Victim data.

File Data

In this section, descriptive subject information derived from file sources is presented. The information is summarized in Table 2.

Thirty nine men participated in this study. They ranged in age from 25 years to 65 years with a mean age of 43.6 years.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Lowest Value</th>
<th>Highest Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>43.6</td>
<td>10.6</td>
<td>25</td>
<td>65</td>
</tr>
<tr>
<td>Education (years)</td>
<td>10.4</td>
<td>3.9</td>
<td>3</td>
<td>18</td>
</tr>
<tr>
<td>Sentence Length (years)</td>
<td>7.5</td>
<td>7.4</td>
<td>2</td>
<td>30</td>
</tr>
<tr>
<td>Time Served (years)</td>
<td>2.4</td>
<td>2.7</td>
<td>0.1</td>
<td>12</td>
</tr>
<tr>
<td>Youngest Victim (years)</td>
<td>-8.5</td>
<td>3.2</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>Oldest Victim (years)</td>
<td>12.4</td>
<td>2.4</td>
<td>4</td>
<td>14</td>
</tr>
</tbody>
</table>
(S.D. = 10.6 years). Compared to other pedophilic samples reported recently in the literature, this sample is relatively old. For example, Avery-Clark and Laws (1984) reported a mean age of 27.4 and Marshall et al. (1986) reported a mean age of 34.5. Level of education for the present sample ranged from grade 3 to grade 18 (high school plus 6 years of university). The mean educational level was 10.4 years. The distribution was primarily bi-modal with modes at grade 7 (n=7, 17.9%) and grade 12 (n=12, 25.6%).

Excluding two men who had indeterminate sentences, sentence length for the sample ranged from 2 to 30 years with a mean sentence length of 7.5 years (S.D. = 7.4 years). The length of time served prior to participating in the study ranged from 1 month to 12 years with a mean time of 2.4 years (S.D. = 2.7 years). Time served prior to the current incarceration was not considered in this calculation.

An approximate estimate of the number of victims of each subject was obtained from his file. According to this information, each subject was classified into one of four groups: 1-5 victims, n=4 (10.3%); 6-10 victims, n=15 (36.5%); 10-20 victims, n=12 (30.8%); and 21 or greater victims, n=8 (20.5%). The fact that very few subjects had less than 6 victims would suggest that this sample is indeed a representative sample of pedophiles. The number of victims reported in each file was regarded as a baseline minimum and is probably an underestimation of the actual number of victims.
The age of each subject's youngest victim ranged from 1 to 14 with a mean age of 8.5 years. The age of the subject's oldest victim ranged from 4 to 14 with a mean age of 12.4 years. Although, some of the men had committed sexual offences against individuals older than 14 years, these were not recorded since they were not the offences of interest for this study.

In summary, these descriptive data suggest that the sample studied is similar to other studies of pedophiles, except that the current sample was older.

*Intelligence Data*

Intellectual functioning was assessed by the WAIS-R. Full Scale IQ scores ranged from 72 to 141. The mean intelligence score was 100.6 (S.D. = 18.7). The distribution was positively skewed with a median intelligence score of 97.5. This sample of pedophiles is similar in intelligence, then, to both the general population (mean IQ M=100, S.D.=15) and to samples of pedophiles previously studied (see Table 1).

Verbal IQ scores ranged from 67 to 145 with a mean verbal intelligence score of 99.0, (S.D. = 19.0). The distribution was positively skewed with a median of 97.5. Performance IQ scores ranged from 75 to 136 with a mean performance IQ of 102.2 (S.D. = 16.9). This distribution was not skewed and had a median of 102.6.
The distribution of intelligence scores according to Wechsler's categories (Superior, High Average, etc.) can be found in Table 3.

The frequency of Verbal/Performance intelligence score disparities ('splits') of 10 or more points was examined. Sixteen instances were found. In eleven of these cases, the Performance IQ was greater than the Verbal IQ. The percentage of splits greater than or equal to 10 found in the present study (41%) is equivalent to the percentage of such splits found in the standardization sample of the WAIS-R (350 of 940, 37%) (Matarazzo, Bornstein, McDermott & Noonan; 1986).

WAIS-R subtest means and standard deviations are presented in Table 4. The means varied from a low of 7.9 for Digit Symbol to a high of 10.1 for Block Design. All of the remaining subtest means were between 9.1 and 9.7.

As a group, then, the pedophiles in this study are similar in intelligence to other samples of pedophiles previously studied.

Other Psychometric Data

A number of other psychological tests besides the WAIS-R were administered, including the State Anxiety Inventory (SAI), the Crowne-Marlowe Social Desirability Scale (CMSDS) and the Trail Making Test (TMT).
Table 3

**Classification of the distribution of intelligence scores**

<table>
<thead>
<tr>
<th>Classification</th>
<th>Verbal</th>
<th>Performance</th>
<th>Full Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very superior</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>(&gt;129)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Superior</td>
<td>2</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>(120-129)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High average</td>
<td>4</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>(110-119)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>15</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>(90-109)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low average</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>(80-89)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Borderline</td>
<td>5</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>(70-79)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Retarded</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(&lt;70)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4

WAIS-R subtest means and standard deviations

<table>
<thead>
<tr>
<th>Subtest</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Lowest Score</th>
<th>Highest Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information</td>
<td>9.5</td>
<td>3.5</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>Digit Span</td>
<td>9.6</td>
<td>3.4</td>
<td>2</td>
<td>18</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>9.7</td>
<td>3.5</td>
<td>4</td>
<td>18</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>9.5</td>
<td>3.6</td>
<td>3</td>
<td>17</td>
</tr>
<tr>
<td>Comprehension</td>
<td>9.5</td>
<td>3.5</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>Similarities</td>
<td>9.3</td>
<td>4.1</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>Picture Completion</td>
<td>9.5</td>
<td>2.7</td>
<td>4</td>
<td>17</td>
</tr>
<tr>
<td>Picture Arrangement</td>
<td>9.1</td>
<td>3.5</td>
<td>4</td>
<td>17</td>
</tr>
<tr>
<td>Block Design</td>
<td>10.0</td>
<td>3.5</td>
<td>3</td>
<td>19</td>
</tr>
<tr>
<td>Object Assembly</td>
<td>9.5</td>
<td>2.9</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>Digit Symbol</td>
<td>7.9</td>
<td>3.0</td>
<td>4</td>
<td>14</td>
</tr>
</tbody>
</table>
The mean score on the SAI was 41.6 (S.D. = 11.6) with a range of 24 to 70. This mean can be compared to those obtained in two normative groups, male prisoners and adult working males. Norms for a prison population (N=212) indicate a mean of 46.0 (S.D. = 11.0) (Spielberger et al., 1983). Thus, the current sample is less anxious than a similar prison population, which scored over the 36th percentile by comparison. Norms for a working adult male population (N=1387) indicate a mean of 35.7 (S.D. = 10.4) (Spielberger et al., 1983). Thus, prisoners in the current sample are more anxious than the average working adult, scoring on average at the 70th percentile compared to working adult males.

The mean for the CMSDS was 16.1 (S.D. = 7.0), with a range of 2 to 29. Data were available for only 33 subjects (6 missing) because the measure was introduced after data collection was underway. The general population norms (N=120) for the CMSDS indicate a mean of 13.7 (S.D. = 5.8) (Crowne & Marlowe, 1960). This difference in means is not significant \( t(152) < 2.01,\) \( p = .05 \). Thus the subjects in the present study present themselves on average in a more socially desirable light than did the normative group.

The mean for Trail Making Test A was 33.4 seconds (S.D. = 18.4). This is an average score for normals aged 40-49 (Davies, 1968; cited in Lezak, 1983). One subject did not know the alphabet and could not complete Trail Making Test B. The mean for Trail Making Test B was 90.6 seconds (S.D. = 47.9). This score
is at the 25th percentile for normals aged 40-49. The subjects performed less well than most normals on the Trail Making Test B. In order to determine how many individuals were impaired, classification of each subject's score was made according to the appropriate age norms. Individuals scoring at the 10th percentile (poor) or lower on either Trails A or B were scored as impaired. According to this criterion, 5 (12.5%) of the subjects were impaired on one or more parts of the Trails Making Test.

Sex of Victim Data

The majority of the subjects, 21 of 39 (53.8%) were heterosexual pedophiles, 12 of 39 (30.8%) were homosexual pedophiles and 6 of 39 (15.4%) were bisexual pedophiles.

All of the variables discussed in the previous section were reanalysed with sex of victim as a grouping variable. No significant relationships were found between sex of victim and any other variable.

Section 2: Arousal Data

The following section is technical. Readers primarily interested in the principal findings of the analyses may wish to proceed to the next section (which commences on page 68) entitled Relationships Between Arousal/Non-arousal and Other Variables.
Plethysmographic assessment provides information about an individual's sexual arousal in response to a variety of sexual stimuli. For the purpose of analysis the raw arousal data is converted into percentages of arousal for each subject. The percentages are examined to see how much arousal was obtained in response to adults and how much to children.

Calculation of Maximal Arousal

The raw arousal data is a digital conversion of an analog signal. A particular penile circumference level is represented by a particular electrical current which is then converted into a digital equivalent.

The raw arousal data (digital equivalents of penile circumference levels) is converted into percentages of arousal by using information about the minimum and maximum circumference levels for each subject. Once minimum and maximum circumference levels are determined for each subject, all of the raw scores for each subject can be easily converted to percentages of the range of arousal, that is the maximum circumference level minus the minimum circumference level. The percentages indicate what proportion of the range of arousal the subject obtained in response to a given stimulus. For example, "80% arousal" indicates that the subject's penis expanded from its minimum circumference level to 80 percent of its maximum circumference level (full erection) in response to a given stimulus. The greater the penis expansion, the higher the percentage and the
greater the sexual arousal.

The determination of the minimum circumference level for each person (which represents his penile circumference in a flacid state) is achieved by selecting the lowest value among his data points. The assessment of each person's maximum circumference level (representing penile circumference during full erection) is more difficult. Many subjects achieve a maximum circumference level (full erection) in response to the experimental stimuli or to the pre-assessment erotic material. When a subject's maximum circumference level to several stimuli approach the same value, and the subject reports having achieved a full erection, then the maximum raw data point for that subject can be accepted as that subject's maximum circumference level.

However, in many cases, subjects produce only a partial erection or no erection at all during the presentation of the assessment stimuli and the pre-assessment erotic material. In these cases the determination of a subject's maximum circumference level is more difficult. In some studies the solution to this problem has been to require subjects to masturbate and to achieve a full erection.

In the present study subjects were not required to do this because of a concern that such a requirement would lower the participation rate in the study.
For the current study it was decided to develop a regression equation on the basis of the minimum and maximum circumference levels of those subjects who achieved a full erection and then to use the minimum circumference levels of those who did not achieve a full erection in order to predict their maximum circumference level scores.

This approach is necessarily imprecise; therefore it was decided to determine what effects different kinds of prediction errors would have upon the subsequent data analyses.

To begin, fifteen individuals were judged to have achieved a full erection according to self report and the assessor's judgement. The assessor's judgement was sometimes required, for example, when the raw data output of an individual indicated a full erection, but the individual claimed to have only achieved 80% arousal. This kind of situation does occur since some men will report a full erection as a partial erection due to concern regarding the size of their erection.

The minimum and maximum raw data scores for each of the fifteen subjects who achieved a full erection (i.e. a maximum circumference level) are listed in Table 5. The regression equation developed from these data is \( y = 1.67529x + 98.04732 \).

The validity of this equation depends in part upon the assumption that the minimum and maximum circumference levels for those who achieved a full erection and for those who did not achieve a full erection are distributed similarly. A comparison
Table 5

Recorded minimum and recorded maximum raw data scores

<table>
<thead>
<tr>
<th>Raw Data Scores</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>107</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>154</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>124</td>
</tr>
<tr>
<td></td>
<td>33</td>
<td>180</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>149</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>135</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>114</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>108</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>108</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>118</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>150</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>107</td>
</tr>
</tbody>
</table>
of the distributions of the minimum circumference levels for those who obtained a full erection and those who did not obtain a full erection was undertaken. The mean low score for those who obtained maximum arousal was 15.13 (S.D. = 12.65), and the mean low score for those who did not obtain a maximum arousal was 15.21 (S.D. = 9.55). The lack of a significant difference between the means suggests that the circumference levels of those who did and those who did not achieve full erections are similar and that the regression equation developed from the one set of data can be used to predict the other.

Nevertheless, the error which may occur in predicting maximum circumference levels in the group using the regression equation cannot be ignored. Such error could result in either an underestimation or an overestimation of the maximum circumference level and hence underestimation or overestimations of arousal to each stimulus. In order to check on the seriousness of error introduced by predicting the maximum circumference level it was decided that two additional regression equations would be developed that deliberately produced under and overestimates in comparison to those resulting from the regression equation developed above. These additional regression equations were generated by adding and subtracting a constant (15) from the intercept, to produce the equations $y = 1.67529x + 73.04732$ and $y = 1.67529x + 113.66769$. The constant 15 is greater than the standard error of estimate of the regression equation (13.3865) and hence is a reasonable
number to use in estimating error.

In summary, three separate regression equations were used to predict maximum circumference levels. The predicted maximum circumference levels were then used to create three different estimates of percentage of arousal. The estimates of percentage of arousal are discussed in the following section.

Classification as Aroused or Non-Aroused to Children

An estimate of the percentage of arousal to each stimulus for each subject was made using the predicted maximum circumference levels generated by the three aforementioned regression equations. Table 6 details the average percent arousal to the three child stimuli, the average percent arousal to the three adult stimuli, and the Pedophilic Deviance Quotient (average child / average adult) generated by the three methods of estimation. Where the actual maximum circumference level was known, the actual percentages of arousal were used. The data are ranked ordered on the basis of the average arousal to the three child stimuli generated by the predicted estimation.

The estimates of percentage arousal were then used to classify subjects as either true positives (showed significant deviant arousal), insufficiently aroused, or false negatives. Two decision rules were used in this process. First, a subject was classified as either aroused or insufficiently aroused, depending on his average percentage arousal to adults or to children. In the literature, the cutoff points used to define
### Table 6

**Percentage arousal: Actual or predicted, under- and over-estimates**

<table>
<thead>
<tr>
<th></th>
<th>Actual or Predicted</th>
<th>Under</th>
<th>Over</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PDQ</td>
<td>Child</td>
<td>Adult</td>
</tr>
<tr>
<td>1.50</td>
<td>2.96</td>
<td>1.97</td>
<td></td>
</tr>
<tr>
<td>.93</td>
<td>3.93</td>
<td>4.24</td>
<td></td>
</tr>
<tr>
<td>1.00</td>
<td>4.27</td>
<td>4.27</td>
<td></td>
</tr>
<tr>
<td>.87</td>
<td>4.36</td>
<td>5.03</td>
<td></td>
</tr>
<tr>
<td>1.00</td>
<td>4.54</td>
<td>4.54</td>
<td></td>
</tr>
<tr>
<td>1.60</td>
<td>4.93</td>
<td>3.08</td>
<td></td>
</tr>
<tr>
<td>1.11</td>
<td>9.38</td>
<td>8.47</td>
<td></td>
</tr>
<tr>
<td>.66</td>
<td>10.61</td>
<td>16.05</td>
<td></td>
</tr>
<tr>
<td>1.95</td>
<td>10.86</td>
<td>5.58</td>
<td></td>
</tr>
<tr>
<td>1.50</td>
<td>11.31</td>
<td>7.54</td>
<td></td>
</tr>
<tr>
<td>1.32</td>
<td>12.98</td>
<td>9.80</td>
<td></td>
</tr>
<tr>
<td>1.05</td>
<td>13.64</td>
<td>13.02</td>
<td></td>
</tr>
<tr>
<td>1.24</td>
<td>13.68</td>
<td>10.99</td>
<td></td>
</tr>
<tr>
<td>.91</td>
<td>15.54</td>
<td>17.03</td>
<td></td>
</tr>
<tr>
<td>.38</td>
<td>17.61</td>
<td>46.06</td>
<td></td>
</tr>
<tr>
<td>1.30</td>
<td>25.42</td>
<td>19.60</td>
<td></td>
</tr>
<tr>
<td>1.03</td>
<td>26.27</td>
<td>25.62</td>
<td></td>
</tr>
<tr>
<td>2.71</td>
<td>32.77</td>
<td>12.11</td>
<td></td>
</tr>
<tr>
<td>3.58</td>
<td>34.91</td>
<td>9.75</td>
<td></td>
</tr>
<tr>
<td>1.07</td>
<td>39.46</td>
<td>35.81</td>
<td></td>
</tr>
<tr>
<td>1.12</td>
<td>39.91</td>
<td>35.81</td>
<td></td>
</tr>
<tr>
<td>3.35</td>
<td>43.29</td>
<td>12.93</td>
<td></td>
</tr>
<tr>
<td>2.21</td>
<td>46.97</td>
<td>21.30</td>
<td></td>
</tr>
<tr>
<td>.79</td>
<td>48.40</td>
<td>61.63</td>
<td></td>
</tr>
<tr>
<td>1.05</td>
<td>51.37</td>
<td>48.91</td>
<td></td>
</tr>
<tr>
<td>2.34</td>
<td>62.71</td>
<td>26.83</td>
<td></td>
</tr>
<tr>
<td>1.07</td>
<td>67.53</td>
<td>63.14</td>
<td></td>
</tr>
<tr>
<td>2.30</td>
<td>68.44</td>
<td>29.79</td>
<td></td>
</tr>
<tr>
<td>1.40</td>
<td>80.28</td>
<td>57.50</td>
<td></td>
</tr>
<tr>
<td>3.07</td>
<td>86.14</td>
<td>28.02</td>
<td></td>
</tr>
<tr>
<td>.92</td>
<td>89.10</td>
<td>96.57</td>
<td></td>
</tr>
<tr>
<td>1.00</td>
<td>92.52</td>
<td>92.52</td>
<td></td>
</tr>
<tr>
<td>1.02</td>
<td>97.13</td>
<td>95.34</td>
<td></td>
</tr>
<tr>
<td>1.89</td>
<td>97.51</td>
<td>51.71</td>
<td></td>
</tr>
<tr>
<td>1.48</td>
<td>97.67</td>
<td>66.00</td>
<td></td>
</tr>
<tr>
<td>1.11</td>
<td>97.86</td>
<td>88.07</td>
<td></td>
</tr>
<tr>
<td>1.01</td>
<td>98.06</td>
<td>97.50</td>
<td></td>
</tr>
<tr>
<td>1.09</td>
<td>98.35</td>
<td>90.12</td>
<td></td>
</tr>
</tbody>
</table>
significant arousal have varied from study to study and across stimuli sets. A cutoff point of greater than 20% arousal is frequently used in studies using taped stimuli (c.f. Avery-Clark & Laws, 1984) and this criterion was adopted for the present research.

Second, those subjects who did demonstrate significant arousal were classified as having PDQs less than .6 (false negatives) or greater than .6 (true positives).

Application of these decision rules meant that the same individuals were being placed in the same groups regardless of the estimate of arousal used. Twenty-four subjects were identified as true positives, fourteen as insufficiently aroused and one as a false negative. Thus, even when three different estimates of percentage arousal were used, distinctions between cases of insufficient arousal, false negatives, and deviant arousal were easily made. It should be noted that not only did the cut-off of 20% arousal produce the same numbers of aroused and non-aroused subjects, but that a natural gap appeared to exist at that point in the distribution. It can be concluded that the regression equation prediction error had no effect on the classification of the subjects.

Originally, subjects who failed to exhibit sufficient arousal for interpretable results (insufficient arousal) were to be separated from those who produced sufficient arousal but did not appear deviant (false negatives). Inspection of the results,
however, revealed that only one individual could be classified as a false negative, a number insufficient to allow independent analyses. Therefore the analyses focused on the distinction between those who produced arousal to children and those who did not. The subject who produced arousal and did not appear deviant was therefore grouped with those who were not aroused for the purposes of the analyses.

Analyses of small n samples (less than 100) can produce relationships that are spurious and do not reflect true characteristics of the population. To minimize this problem, two additional definitions of the non-aroused group were created. According to the first definition, the two non-aroused subjects with the greatest arousal were redefined as aroused. The second definition reclassified the two aroused subjects with the least arousal as non-aroused. Thus, the individuals who were at the greatest risk for mis-classification were deliberately re-classified. It was expected that only robust results would emerge across all three definitions of the non-aroused group.

Section 3: Relationships Between Arousal/Non-arousal and Other Variables

In order to explore the central hypotheses of this study, the relationships between arousal, non-arousal and other variables were explored. One way analyses of variance were run with arousal/non-arousal as the grouping factor. Three
definitions of non-arousal were used. According to the first definition there were 15 subjects in the non-aroused group; the second and third definitions were conservative redefinitions of the first definition introduced to reduce the chance of spurious correlations. These redefinitions produced 13 and 17 subjects respectively in the non-aroused group. A number of significant relationships emerged. These will be discussed with regard to three factors: intelligence, Performance-Verbal IQ differences and age.

*Intelligence*

The Comprehension subtest score was the only one of the fourteen WAIS-R scores (11 sub-tests, FSIQ, VIQ, PIQ) to be significantly related to the presence or absence of arousal. Under two of the definitions, group definitions in the Comprehension subtest scores were significant at the \( p < .001 \) level, while in the third definition the group difference was significant at \( p < .0012 \). FSIQ, VIQ and PIQ means are presented in Table 7, and means of the sub-tests are presented in Table 8.

A high Comprehension score was significantly associated with membership in the non-aroused group. A low Comprehension score was associated with membership in the aroused group. This finding indicates that the outcome of plethysmographic assessment (aroused/non-aroused) is related to an aspect of cognitive ability.
<table>
<thead>
<tr>
<th>Arousal Definition</th>
<th>Intelligence Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Verbal</td>
</tr>
<tr>
<td>(1) Aroused (24)</td>
<td>94.4</td>
</tr>
<tr>
<td>Non-aroused (15)</td>
<td>106.3</td>
</tr>
<tr>
<td>(2) Aroused (26)</td>
<td>94.5*</td>
</tr>
<tr>
<td>Non-aroused (13)</td>
<td>108.0</td>
</tr>
<tr>
<td>(3) Aroused (22)</td>
<td>91.3**</td>
</tr>
<tr>
<td>Non-aroused (17)</td>
<td>109.0</td>
</tr>
</tbody>
</table>

*p < .05

**p < .01
<table>
<thead>
<tr>
<th>Arousal Definition</th>
<th>Information</th>
<th>Digit Span</th>
<th>Vocabulary</th>
<th>Arithmetic</th>
<th>Comprehension</th>
<th>Similarities</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Aroused (24)</td>
<td>8.7</td>
<td>9.5</td>
<td>9.1</td>
<td>9.3</td>
<td>8.1***</td>
<td>8.8</td>
</tr>
<tr>
<td>Non-aroused (15)</td>
<td>10.7</td>
<td>9.7</td>
<td>10.5</td>
<td>9.9</td>
<td>11.7</td>
<td>10.1</td>
</tr>
<tr>
<td>(2) Aroused (26)</td>
<td>8.7</td>
<td>9.2</td>
<td>9.0</td>
<td>9.2</td>
<td>8.3**</td>
<td>8.8</td>
</tr>
<tr>
<td>Non-aroused (13)</td>
<td>10.9</td>
<td>10.4</td>
<td>10.9</td>
<td>10.1</td>
<td>11.9</td>
<td>10.3</td>
</tr>
<tr>
<td>(3) Aroused (22)</td>
<td>8.1***</td>
<td>9.4</td>
<td>8.5*</td>
<td>8.7</td>
<td>7.7***</td>
<td>8.3</td>
</tr>
<tr>
<td>Non-aroused (17)</td>
<td>11.2</td>
<td>9.8</td>
<td>11.2</td>
<td>10.5</td>
<td>11.9</td>
<td>10.6</td>
</tr>
</tbody>
</table>

**Verbal Subtests**

<table>
<thead>
<tr>
<th>Arousal Definition</th>
<th>Picture Completion</th>
<th>Picture Arrangement</th>
<th>Block Design</th>
<th>Object Assembly</th>
<th>Digit Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Aroused (24)</td>
<td>9.5</td>
<td>9.3</td>
<td>10.3</td>
<td>10.0</td>
<td>7.9</td>
</tr>
<tr>
<td>Non-aroused (15)</td>
<td>9.3</td>
<td>8.8</td>
<td>9.7</td>
<td>8.6</td>
<td>7.9</td>
</tr>
<tr>
<td>(2) Aroused (26)</td>
<td>9.4</td>
<td>9.1</td>
<td>10.0</td>
<td>9.7</td>
<td>7.6</td>
</tr>
<tr>
<td>Non-aroused (13)</td>
<td>9.6</td>
<td>9.2</td>
<td>10.2</td>
<td>8.9</td>
<td>8.4</td>
</tr>
<tr>
<td>(3) Aroused (22)</td>
<td>9.4</td>
<td>9.0</td>
<td>10.0</td>
<td>9.7</td>
<td>7.8</td>
</tr>
<tr>
<td>Non-aroused (17)</td>
<td>9.6</td>
<td>9.4</td>
<td>10.1</td>
<td>9.2</td>
<td>8.1</td>
</tr>
</tbody>
</table>

***p < 0.01

- **p < 0.001**
Verbal IQ - Performance IQ Differences

Observation of the large number of 10 point or greater Verbal IQ - Performance IQ splits led to the incorporation of a variable reflecting these differences in the analyses. The mean VIQ - PIQ differences for the aroused/non-aroused groups are detailed in Table 9. On average, the aroused group had PIQs 7 points higher than VIQs, while the non-aroused group had mean VIQs 4 points greater than PIQs. This finding is consonant with the finding regarding the significance of the Comprehension score. A high Comprehension subtest score and a VIQ greater than a PIQ both were associated with non-arousal.

Age

Age was significantly related to group membership in two definitions, but was only significant to .056 in the third (Table 10). However, this finding is considered here and in later analyses because of theoretical reasons which suggest that age may be significantly related to patterns of arousal.

Theoretically, age is related to arousal in that the older individual is less likely to exhibit arousal. Although, it is commonly assumed that sexual vigor decreases with age, no studies have yet demonstrated that older individuals are less likely to become aroused during plethysmographic assessment. The trend in the data here may reflect this hypothesized relationship between age and arousal.
<table>
<thead>
<tr>
<th>Arousal Definition</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Aroused (24)</td>
<td>-7.6 **</td>
<td>12.9</td>
</tr>
<tr>
<td>Non-aroused (15)</td>
<td>3.9</td>
<td>9.2</td>
</tr>
<tr>
<td>(2) Aroused (26)</td>
<td>-6.8 *</td>
<td>12.8</td>
</tr>
<tr>
<td>Non-aroused (13)</td>
<td>3.9</td>
<td>9.7</td>
</tr>
<tr>
<td>(3) Aroused (22)</td>
<td>-9.0 ***</td>
<td>12.5</td>
</tr>
<tr>
<td>Non-aroused (17)</td>
<td>4.2</td>
<td>8.9</td>
</tr>
</tbody>
</table>

*p<.05
**p<.01
***p<.001
<table>
<thead>
<tr>
<th>Arousal Definition</th>
<th>Age</th>
<th>Education</th>
<th>Sentence Length</th>
<th>Time Served</th>
<th>Youngest Victim</th>
<th>Oldest Victim</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Aroused (24)</td>
<td>40.1**</td>
<td>9.4</td>
<td>7.8</td>
<td>2.4</td>
<td>8.1</td>
<td>11.8*</td>
</tr>
<tr>
<td>Non-aroused (15)</td>
<td>49.3</td>
<td>11.9</td>
<td>7.0</td>
<td>2.3</td>
<td>9.2</td>
<td>13.3</td>
</tr>
<tr>
<td>(2) Aroused (26)</td>
<td>41.5</td>
<td>9.5</td>
<td>8.8</td>
<td>3.0</td>
<td>8.2</td>
<td>11.9</td>
</tr>
<tr>
<td>Non-aroused (13)</td>
<td>48.0</td>
<td>12.1</td>
<td>4.9</td>
<td>1.2</td>
<td>9.2</td>
<td>13.3</td>
</tr>
<tr>
<td>(3) Aroused (22)</td>
<td>39.3**</td>
<td>8.8**</td>
<td>8.2</td>
<td>2.5</td>
<td>7.9</td>
<td>11.6*</td>
</tr>
<tr>
<td>Non-aroused (17)</td>
<td>49.2</td>
<td>12.4</td>
<td>6.6</td>
<td>2.5</td>
<td>9.4</td>
<td>13.3</td>
</tr>
</tbody>
</table>

*p < .05
**p < .01
Other Variables

A number of other variables were included in the present study for theoretical reasons. Of these, neither anxiety, social desirability, or neuropsychological impairment were significantly related to arousal/non-arousal. These results are presented in Table 11.

A number of treatment variables were coded for each individual (see Table 12). Cyproterone acetate was being used by three subjects. Theoretically, this drug reduces sexual arousal. If these subjects had not produced arousal, their failure to become aroused could be attributed to the medication. However, all three individuals did produce significant amounts of arousal to children. Whether the medication did in fact reduce their arousal is not known. The effectiveness of the medication as a complete arousal suppressor must certainly be questioned.

Prior to the commencement of the study, many subjects had received some form of individual or group psychological treatment. Although a significant relationship between previous individual treatment and arousal to children appears in one of the analyses, further inspection reveals that the relationship can be attributed to two "outlying" subjects who received 50 and 65 individual sessions. No other consistent trend attributable to treatment can be observed and in fact, when one of the highly treated individuals is reclassified in the third analysis the significant relationship disappears.
<table>
<thead>
<tr>
<th>Arousal Definition</th>
<th>Crowne-Marlow</th>
<th>State Anxiety</th>
<th>Trails A</th>
<th>Trails B</th>
<th>Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Aroused (24)</td>
<td>14.4</td>
<td>43.0</td>
<td>31.0</td>
<td>89.6</td>
<td>.63</td>
</tr>
<tr>
<td></td>
<td>Non-aroused (15)</td>
<td>18.7</td>
<td>39.2</td>
<td>37.3</td>
<td>92.4</td>
</tr>
<tr>
<td>(2) Aroused (26)</td>
<td>15.3</td>
<td>42.4</td>
<td>33.4</td>
<td>94.3</td>
<td>.62</td>
</tr>
<tr>
<td></td>
<td>Non-aroused (13)</td>
<td>17.7</td>
<td>39.9</td>
<td>33.5</td>
<td>82.8</td>
</tr>
<tr>
<td>(3) Aroused (22)</td>
<td>14.2</td>
<td>44.4</td>
<td>31.6</td>
<td>90.5</td>
<td>.68</td>
</tr>
<tr>
<td></td>
<td>Non-aroused (17)</td>
<td>18.3</td>
<td>37.9</td>
<td>35.8</td>
<td>90.9</td>
</tr>
</tbody>
</table>
### Table 12

**Arousal/non-arousal: Treatment**

| Arousal Definition | Individual | | | | | | | | Group | | | | | | | | Medication | | | | |
|-------------------|------------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
|                   | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| (1) Aroused (24)   | 5.5  | 12.5 | 6.2  | 8.4 | 2.4  | 6.8 |       |    |       |    |       |    |       |    |       |    |       |    |
| Non-aroused (15)   | 9.1  | 16.4 | 3.7  | 9.3 | 0.0  | 0.0 |       |    |       |    |       |    |       |    |       |    |       |    |
| (2) Aroused (26)   | 7.3  | 16.2 | 5.7  | 8.2 | 2.2  | 6.6 |       |    |       |    |       |    |       |    |       |    |       |    |
| Non-aroused (13)   | 5.8  | 8.9  | 4.3  | 9.9 | 0.0  | 0.0 |       |    |       |    |       |    |       |    |       |    |       |    |
| (3) Aroused (22)   | 2.2  | 5.6  | 6.3  | 8.7 | 1.3  | 3.6 |       |    |       |    |       |    |       |    |       |    |       |    |
| Non-aroused (17)   | 12.8 | 18.9 | 3.8  | 8.8 | 1.8  | 7.3 |       |    |       |    |       |    |       |    |       |    |       |    |
In summary, two variables, Comprehension and Verbal-Performance differences had robust relationships with the dependent variable, arousal/non-arousal. An additional variable, age, was less robustly, but still significantly related to arousal.

Classification of Subjects into Aroused/Non Aroused Groups by Means of Logistic Regression

Given that three variables were significantly related to the dependent variable, arousal/non-arousal, it is important to identify the relative importance of these variables. Is one variable more important than the others? Or is one combination of these variables superior to other combinations? In order to answer these questions, the three variables, Comprehension scale score, Verbal IQ - Performance IQ differences, and age were used as predictor variables in a series of logistic regressions. Then, each variable's comparative usefulness in classifying the dependent variable was examined.

Before the results of these analyses are presented, however, it should be noted that, while logistic regression is an ideal method for investigating these data, selecting variables on the basis of the strength of their individual relationship with the dependent variable is a poor strategy. Ideally, one would wish to include all of the variables measured in the study, because it is conceivable that a variable which has no independently
significant relationship with the dependent variable may be highly significant in a classification sense as a suppressor variable. Unfortunately, such analyses would have required considerably more subjects than were studied. Given available data, then, the method chosen to analyse the data was most appropriate.

The statistical significance of variables included in a logistic regression can be determined in at least two ways: by inspection of the improvement chi-square, and by inspection of the p-value of the last added term. Additionally the practical significance of a group of predictors can be assessed by inspection of the predicted probabilities plot using the developed regression equation. The latter method for assessing the value of a particular group of predictors was applied in the current study. This method essentially means comparing the relative number of errors made in prediction when each combination of variables is employed.

Comprehension

As can be seen in Table 13, it is clear that the most reliable way to classify subjects is to group them according to Comprehension subtest scores. Comprehension scores classify subjects well across all three definitions of non-arousal.

Using Comprehension scores only, 6 errors are made in classifying 39 subjects, which yields an 85% correct classification rate. Simply predicting the base rate of arousal
Table 13

Logistic regression results as aroused/non-aroused: Number of predictor errors

<table>
<thead>
<tr>
<th>Predictor Set</th>
<th>Number not Aroused</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>15</td>
</tr>
<tr>
<td>Comprehension (C)</td>
<td>6</td>
</tr>
<tr>
<td>Age (A)</td>
<td>10</td>
</tr>
<tr>
<td>Differences (D)</td>
<td>9</td>
</tr>
<tr>
<td>C x A</td>
<td>8</td>
</tr>
<tr>
<td>C x D</td>
<td>6</td>
</tr>
<tr>
<td>A x D</td>
<td>5</td>
</tr>
<tr>
<td>C x A x D</td>
<td>7</td>
</tr>
</tbody>
</table>
in this study, 24 of 39 subjects would be correctly classified, a hit rate of 59%.

As the logistic regressions demonstrate, the Comprehension subtest score can be used to determine whether a given subject belongs to the aroused or the non-aroused group. In order to determine which score of the Comprehension subtest should be used as a cutoff for designating group membership, it is necessary to work backwards from the scatter plots of predicted probabilities produced by BMDP PLR. The classification error table (Table 13), is derived from a process in which the optimum cutoff for each individual probability plot is selected. However, in order to apply a cutoff based on the current sample to future samples, a single "best" cutoff must be specified.

The ideal cutoff for the current sample was selected through visual analysis of the three scatter plots generated by the BMDP PLR when the Comprehension subtest score was the predictor variable. Particular attention was paid to the scatter plot generated on the basis of the primary definition of the aroused/non-aroused groups. A probability cutoff of .45 was selected, which resulted in no change in the classification errors for the primary and the secondary definitions of the groups. The classification errors based on the third definition of aroused/non-aroused increased from six to ten, as an optimum cutoff was no longer being employed with that probability plot.
The Comprehension subtest score corresponding to a probability of .45 was obtained in the following manner. The probability \( p \) is a function of a variable \( t \), which can be expressed as the natural log of \( p/(1-p) \). The Comprehension score is found by solving the equation \( C=(t-a)/b \), where \( a \) and \( b \) are coefficients generated by the BMDP PLR program. Using the primary definition of aroused/non-aroused \( a=-4.6902 \) and \( b=.42109 \). When the equation is solved using these values, the result is a score of 10.93. In whole numbers, a scaled score of 11 or greater defines the non-aroused group while a scale score of 10 or below defines the aroused group.

**Pedophilic Deviance Quotients**

The Pedophilic Deviance Quotient (PDQ) is a ratio of degree of arousal to children divided by degree of arousal to adults. PDQs are customarily computed in studies of pedophiles, and a cutoff score of .6 or .8 is usually used to designate a profile as pedophilic or not. In the current study, all aroused subjects have PDQs above .6 and all but two have PDQs above .8. Pedophilic Deviance Quotients are presented in Table 7. For the purposes of the following analyses, only those subjects who were classified as aroused were included in the analyses, since a PDQ is meaningless without meaningful levels of arousal.
Pedophilic Deviance Quotients and Intelligence: Correlations

Marshall et al. (1986) reported a correlation of -.33 between Pedophilic Deviance Quotients and intelligence as measured by Raven's Progressive Matrices Test. The relationships between PDQs and intelligence as measured by the WAIS-R were therefore examined in the present study. Table 14 lists the correlations between FSIQ, VIQ and PIQ and PDQs. Again multiple definitions of non-arousal were used. None of the intelligence measures had a significant relationship with PDQ. Therefore this study does not replicate the Marshall et al. (1986) finding that PDQs and intelligence per se are correlated.

Pedophilic Deviance Quotients and Intelligence: Cutoff Scores

Marshall et al. (1986) also examined the relationship between PDQs and intelligence from another perspective by using cutoff scores to group subjects as having IQs below 80, between 80-100, and above 100. A similar approach was tried in the current study. The cut-offs of below 90, 90-110, and above 110 were also tried.

Examination of Table 15 shows that no significant relationships were found between the measures of intelligence and PDQs.
Table 14

Pedophilic deviance quotients and intelligence: Correlations

<table>
<thead>
<tr>
<th>Intelligence Measure</th>
<th>Number</th>
<th>Aroused</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Scale</td>
<td>-.05</td>
<td>-.06</td>
</tr>
<tr>
<td>Verbal</td>
<td>-.01</td>
<td>-.05</td>
</tr>
<tr>
<td>Performance</td>
<td>-.09</td>
<td>-.06</td>
</tr>
<tr>
<td>Information</td>
<td>-.06</td>
<td>-.09</td>
</tr>
<tr>
<td>Digit Span</td>
<td>.10</td>
<td>.05</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>-.02</td>
<td>-.10</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>-.06</td>
<td>-.05</td>
</tr>
<tr>
<td>Comprehension</td>
<td>-.14</td>
<td>.04</td>
</tr>
<tr>
<td>Similarities</td>
<td>-.01</td>
<td>-.13</td>
</tr>
<tr>
<td>Picture Completion</td>
<td>-.12</td>
<td>.20</td>
</tr>
<tr>
<td>Picture Arrangement</td>
<td>-.23</td>
<td>.09</td>
</tr>
<tr>
<td>Block Design</td>
<td>-.22</td>
<td>.12</td>
</tr>
<tr>
<td>Object Assembly</td>
<td>-.23</td>
<td>.07</td>
</tr>
<tr>
<td>Digit Symbol</td>
<td>.09</td>
<td>.01</td>
</tr>
<tr>
<td>IQ Type</td>
<td>Cut-off Scores</td>
<td>Number</td>
</tr>
<tr>
<td>---------</td>
<td>----------------</td>
<td>--------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>24</td>
</tr>
<tr>
<td>&lt;80</td>
<td>1.4 (6)</td>
<td>1.4 (6)</td>
</tr>
<tr>
<td>80-100</td>
<td>1.8 (10)</td>
<td>1.8 (10)</td>
</tr>
<tr>
<td>&gt;100</td>
<td>1.5 (8)</td>
<td>1.6 (6)</td>
</tr>
<tr>
<td>Verbal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;90</td>
<td>1.7 (10)</td>
<td>1.7 (10)</td>
</tr>
<tr>
<td>90-110</td>
<td>1.7 (10)</td>
<td>1.7 (10)</td>
</tr>
<tr>
<td>&gt;110</td>
<td>1.2 (4)</td>
<td>1.3 (2)</td>
</tr>
<tr>
<td>&lt;80</td>
<td>1.7 (2)</td>
<td>1.7 (2)</td>
</tr>
<tr>
<td>80-100</td>
<td>1.8 (9)</td>
<td>1.8 (9)</td>
</tr>
<tr>
<td>&gt;100</td>
<td>1.5 (13)</td>
<td>1.5 (11)</td>
</tr>
<tr>
<td>Performance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;90</td>
<td>1.9 (7)</td>
<td>1.9 (7)</td>
</tr>
<tr>
<td>90-110</td>
<td>1.4 (9)</td>
<td>1.4 (9)</td>
</tr>
<tr>
<td>&gt;110</td>
<td>1.6 (8)</td>
<td>1.7 (6)</td>
</tr>
<tr>
<td>&lt;80</td>
<td>1.5 (3)</td>
<td>1.5 (3)</td>
</tr>
<tr>
<td>80-100</td>
<td>1.7 (12)</td>
<td>1.7 (12)</td>
</tr>
<tr>
<td>&gt;100</td>
<td>1.6 (9)</td>
<td>1.7 (7)</td>
</tr>
<tr>
<td>Full Scale</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;90</td>
<td>1.5 (8)</td>
<td>1.5 (8)</td>
</tr>
<tr>
<td>90-110</td>
<td>1.6 (10)</td>
<td>1.6 (10)</td>
</tr>
<tr>
<td>&gt;110</td>
<td>1.7 (6)</td>
<td>1.9 (4)</td>
</tr>
</tbody>
</table>
Generalization Data

Applied research must especially concern itself with
generalizability. Strictly speaking, the results obtained in
this study may not be generalizable to the entire population of
incarcerated pedophiles, since only volunteer incarcerated
pedophiles were studied. It is conceivable that volunteer
incarcerated pedophiles differ in important ways from
non-volunteer incarcerated pedophiles. Therefore comparisons were
made between the two groups on intelligence, age, length of
sentence and time served.

Intelligence

Since a major concern of this study was intelligence, an
attempt was made to use the information available on the
intelligence of those who were recruited, but did not
participate in the study. The Shipley-Hartford, a "quick"
measure of (primarily) verbal intelligence was available for 21
individuals who did not participate in the study and for 14 of
those who did. These groups were compared. Those participating
in the study had significantly higher Verbal, Abstract, and
Combined scores, but the WAIS-R equivalent scores were not
significantly higher (see Table 16). Thus, it appears that
non-participators may possess less verbal intelligence than
those who participated. However, these differences in verbal
intelligence do not appear to be large.
Table 16

Participators/non-participators: Shipley Hartfield scores

<table>
<thead>
<tr>
<th>Group</th>
<th>Verbal</th>
<th>Abstract</th>
<th>Combined</th>
<th>WAIS Equivalent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participators</td>
<td>126.2*</td>
<td>112.6*</td>
<td>119.5*</td>
<td>109.2</td>
</tr>
<tr>
<td>N=14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-participators</td>
<td>110.7</td>
<td>98.4</td>
<td>103.0</td>
<td>101.1</td>
</tr>
<tr>
<td>N=21</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p<.05
**Other Variables**

File data on three variables were collected in reference to a total of 50 pedophiles who were approached regarding the study but who decided not to participate. These variables were age, length of sentence, and time served (see Table 17).

As can be observed by examination of Table 17, no significant differences between participators and non-participators were found on any of these variables.

In summary, given the available file data, attempts were made to compare those who participated in the study with those who did not. As far as can be determined, the populations differed only in that the participators were more verbally intelligent than the non-participators.
Table 17

Participators/non-participators: Other variables

<table>
<thead>
<tr>
<th>Group</th>
<th>Age</th>
<th>Sentence Length</th>
<th>Time Served</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participators</td>
<td>43.6</td>
<td>7.5</td>
<td>2.4</td>
</tr>
<tr>
<td>N=39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-participators</td>
<td>42.9</td>
<td>10.2</td>
<td>4.2</td>
</tr>
<tr>
<td>N=50</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER IV
DISCUSSION

Review of Hypotheses

Four hypotheses were proposed. The primary hypothesis was that intelligence is related to the outcome of plethysmographic assessment. The other three hypotheses concerned the comparability of the current sample of pedophiles to other samples of pedophiles described previously in the literature. Before results relating to the major hypothesis are explored the other three hypotheses will be discussed.

In order to determine the representativeness of the current sample, three descriptive criteria were investigated: average IQ, the number or percent insufficiently aroused, and the number or percent of false negatives. The first descriptive variable, average IQ of the present sample, was compared to mean IQs reported in a series of previous studies on pedophiles' intelligence. The remaining two of these descriptive criteria from the present sample, percent of subjects insufficiently aroused and percent of false negatives, were primarily compared to data reported by Marshall et al. (1986). The sample studied in Marshall et al. (1986) was selected as a comparative pedophile sample, first, because Marshall et al. are among the few who have reported this important information, and second, because their sample represents a typical pedophilic sample.
Comparisons of the current group with other samples of pedophiles provide information about the representativeness of the current sample and allow for informed inferences about the external validity of the present study.

**Hypothesis 1 - Average IQ**

The first hypothesis predicted that the mean intelligence scores of the current sample of pedophiles would be similar to those reported for other samples in the literature. The results of many of these studies were reviewed in the Introduction (for a summary see Table 1). In general, previous studies employing either the WAIS or the WAIS-R have reported the mean Full Scale IQ of the sample to fall within an IQ range of 90 - 100. The average full scale IQ of the current sample was 100.6 IQ points. This finding is consistent with results in earlier studies and indicates that the current sample of pedophiles is not different from other pedophilic samples on the key experimental variable of intelligence.

**Hypothesis 2 - Percent Insufficiently Aroused**

The second hypothesis was that between 20% to 40% of pedophiles would produce insufficient arousal to allow their results to be interpreted. Marshall et al. (1986) reported that 11 of 51, (21.6%) of pedophilic subjects developed insufficient arousal. Similar figures were reported by Avery-Clark & Laws (1984) with 11 of 50, (22%) of subjects being insufficiently aroused. However, informal clinical reports suggest that larger
percentages of insufficiently aroused subjects are common. Consistent with some clinical impressions, 14 of 39, or 36% of subjects were insufficiently aroused in the current study. This percentage is somewhat higher than that reported in earlier studies, yet within the range of results reported by other clinicians. It is considered that the current pedophilic sample is not markedly different from previous pedophile samples reported in the literature in terms of the percent of subjects who were insufficiently aroused to allow their plethysmographic results to be interpreted.

**Hypothesis 3 - Percent False Negatives**

The third hypothesis was that approximately 40% of those pedophiles who produced sufficient arousal to allow their results to be interpreted would produce a normal, non-deviant sexual arousal pattern. These subjects would be false negatives, because they would not be identified as pedophiles on the basis of their sexual arousal pattern. The figure of 40% false negatives was drawn from the Marshall et al. (1986) study in which a Pedophilic Deviance Quotient (PDQ) of .6 resulted in the classification of 40% (16 of 40) subjects as normal. In the current study a PDQ of .6 resulted in the classification of only one subject as normal. That is, only one subject produced a normal pattern of arousal to adults and minimal arousal to children. Because the Marshall et al. (1986) study is the only report known in which the information on the number of false negatives was published, comparisons with other studies cannot
be made. Therefore it is not possible to state which result better typifies pedophiles in general.

One explanation for the different findings may relate to the relative severity of pedophilic behavior in each sample. Using type of sentence as an indicator of crime severity, which would be based on such factors as the number of victims, type of sexual acts and degree of coercion used, the present sample is judged to be more seriously pedophilic than the sample of Marshall et al. (1986). Subjects in the current study have all been sentenced to time in a federal prison (greater than or equal to 2 years). In contrast, 65% of the Marshall et al. (1986) sample were at "some stage of judicial process or investigation" (p. 427, 1986). Thus it is reasonable to suppose that not all of the subjects who took part in Marshall et al. (1986) study would have received federal sentences. Subjects may have been acquitted, received fines or probation, or received provincial sentences (up to two years less a day). Thus the different nature of the samples may explain why fewer false negatives than predicted were found in the current sample.

In summary, it was hypothesized that the current sample of pedophiles was similar to other previously reported samples of pedophiles with respect to average IQ, percent insufficiently aroused, and percent false negatives. It was determined that the current sample is similar in intelligence to previously reported samples, and produced a similar number of insufficiently aroused subjects. However, the current sample differed markedly from the
Marshall et al. (1986) sample in terms of the percent of false negative results. The current sample has far fewer false negatives. This result may be best explained by reference to the more seriously pedophilic nature of the current sample.

_Hypothesis 4 - Cognitive Ability and Plethysmographic Assessment_

The fourth hypothesis was the main experimental hypothesis.

It was predicted that pedophiles' intelligence would be related to the outcome of plethysmographic assessment. Based on the findings of Marshall et al. (1986), it was expected that more intelligent pedophiles would be less likely to exhibit deviant arousal, that is, that these subjects would be false negatives; while less intelligent pedophiles would be more likely to exhibit deviant arousal and would emerge as true positives.

However, the finding that only one subject produced a false negative profile prevented a between group comparison of the intelligence of those pedophiles who appeared deviant (true positives) and those pedophiles who appeared normal (false negatives). Thus, no conclusions could be drawn about the problem of false negatives. It was possible, however, to consider the relationship between insufficient arousal and intelligence. Therefore, the data analyses examined the relationship of intelligence to the outcome data of two groups: those who became aroused (the aroused group); and those who did not become aroused and the single false negative subject (the
non-aroused group). The discussion that follows, then, focuses on the between group analyses using the aroused/non-aroused groups. The relationship between intelligence and relative level of deviance (as defined by the Pedophilic Deviance Quotient) within the aroused group is considered at a later point.

The principal finding of this study is that two specific aspects of intelligence were related to the aroused/nonaroused grouping: Comprehension subtest scores and Verbal IQ - Performance IQ differences. Subjects who failed to demonstrate arousal had significantly higher Comprehension subtest scores and also had greater Verbal IQs than Performance IQs (positive Verbal IQ - Performance IQ differences). Conversely, those subjects who became aroused had lower Comprehension subtest scores and had Performance IQs greater than Verbal IQs (negative Verbal IQ - Performance IQ differences).

No other aspects of intelligence measured in this study were related to arousal/nonarousal. Because the Comprehension and Picture Arrangement subtests are both generally considered to measure social understanding on the WAIS-R, it might be reasonable to expect that Picture Arrangement scores would also be related to arousal. However, Picture Arrangement scores were not significantly related to the aroused versus non-aroused groupings. Thus, it appears that verbally mediated social comprehension, rather than a more globally mediated social comprehension, is the important variable. This conclusion is supported by the fact that the non-aroused group also had the
higher Verbal than Performance IQs.

These results strongly suggest that Comprehension and Verbal IQ - Performance IQ differences are important in plethysmographic assessment and therefore deserve further consideration. The Comprehension subtest, which had the strongest statistical relationship with the aroused/non-aroused grouping, is considered first, and discussion of the Verbal IQ - Performance IQ differences follows.

Arousal/Non-Arousal and the Comprehension Subtest

The Comprehension subtest of the WAIS-R measures:

"an individual's ability to understand common social customs and situations and to provide for appropriate activity as well as the reasons behind such activities. It assesses one's awareness of social norms and practices, and one's ability to abstract the reasons behind them." (Golden, 1979, page 16).

Essentially, Golden's interpretation means that the Comprehension subtest measures social understanding. Thus it appears that the social understanding of those who responded to plethysmographic assessment (the aroused group) and those who did not (the non-aroused group) differs substantially. Subjects with a high level of social understanding were not, as a group, identifiable as pedophiles on the basis of their plethysmographic assessment. In most cases, subjects with a high level of social understanding did not produce sufficient arousal
for their plethysmographic assessment results to be interpreted. In comparison, subjects with a poorer level of social understanding were, as a group, readily identifiable as pedophiles on the basis of plethysmographic assessment.

The mean Comprehension subtest score of the aroused group was 8.1. Although this score is below average for the general population, it is sufficiently high that the aroused group knew that deviant sexuality was being assessed and were able to distinguish between appropriate and inappropriate sexual cues. Therefore, the differences between groups cannot be attributed to one group's inability to distinguish between deviant and appropriate stimuli.

How then can the difference between the groups on the Comprehension subtest be explained? The simplest explanation for this finding is that the more socially aware subjects deliberately faked their results. In this regard, it has been demonstrated that some motivated subjects can inhibit their arousal. The present results might suggest, then, that more socially aware pedophiles are better able to develop and implement strategies that are effective in inhibiting their arousal and that these subjects faked non-arousal in order to make a good impression.

An alternative explanation suggests that the failure to become aroused is the product of a greater sensitivity to the context within which the sexual stimuli are presented. From this
perspective, the more socially aware subjects are not thought to be more motivated to control their arousal than less socially aware subjects. In certain circumstances it would be sensible for a pedophile to attempt to inhibit his arousal. One circumstance when controlling deviant arousal would be advantageous would be prior to sentencing, when appearing less deviant might result in a shorter prison term. However, the circumstances under which the current study was conducted were such that it was not directly advantageous for any subject to inhibit their arousal. Each participant in the study was not only a convicted and incarcerated pedophile but a self-confessed one. Therefore, exhibiting deviant arousal to children should not have been alarming to any of the subjects. Additionally, it was made very clear to each subject that no benefits or punishments would be accrued as a result of participation in the study, and that the results would not be communicated to any corrections staff. Subjects knew that no negative consequence, such as a longer sentence as a result of exhibiting deviant arousal, would ensue from the plethysmographic assessment. Furthermore, subjects knew that no benefits, such as positive recommendations for parole, would result if they did not exhibit deviant arousal. Therefore it is argued that the circumstances were such that the subjects should not have been highly motivated to inhibit their arousal.

At the same time, it is a commonly acknowledged fact that inmates within the prison system are often suspicious or
cynical. They do not trust what is told to them. This mistrust develops both as a function of their denial of their "deviance" and also as a result of actual experiences within the correctional system. Yet, it seems probable that the more intelligent pedophiles would be more likely to appreciate that taking part in the study would not have recriminating effects. The more socially aware pedophiles should be more likely to understand and believe that the researcher was bound by a complex series of ethical and procedural guidelines that would in fact increase the likelihood that the researcher would not betray an inmate's trust. On the other hand, less socially aware pedophiles might be less familiar with research conventions such as ethics committees and professional standards and be more distrustful. Hence they should be more likely to fake their profiles.

In general, the evidence supporting the explanation that individuals were motivated to suppress arousal for personal gain does not seem very plausible. The fact that higher comprehension pedophiles were less likely to exhibit arousal (deviant or otherwise) is not explained adequately by the notion that they were more motivated to suppress arousal. On the contrary, it is argued that higher comprehension pedophiles were more likely to realize that faking was unnecessary within a research setting where confidentiality is guaranteed.

The question remains: why are subjects with a high Comprehension score less likely to exhibit arousal and why are
subjects with a low Comprehension score more likely to exhibit arousal? The second explanation of this finding relies on a definition of the Comprehension subtest as a measure of past verbal learning about social functioning. This verbally based memory for social functioning permits a more complex cognitive filtration of external events. Events become understood not only in terms of their immediate stimulus qualities but also in terms of past verbally mediated learning about social functioning. A high social comprehension score may reflect a greater ability to place stimulus events into a cognitive framework. On the other hand, a low social comprehension score may reflect more direct and less cognitively mediated experiencing.

It is argued, then, that high Comprehension score subjects pay attention to a variety of internally generated cognitive factors which mediate the perception of the sexual stimulus in such a way that the subjects fail to become aroused. The exact nature of the cognitive mediation process is difficult to determine. However, an anecdotal example may indicate the manner in which cognitive mediation interacts with arousal. One subject with a high Comprehension score spontaneously remarked during debriefing that knowing his victims was extremely important to him. He stated that it was his knowledge of them that enabled him to become sexual with them and that listening to tapes about people he did not know in the assessment setting was not arousing. Thus, he was expressing how an internally generated verbal memory was critical to his becoming aroused, whereas the
assessment setting and the cognitive associations that were generated by it failed to arouse him. This subject's remarks are considered illustrative of the process whereby pedophiles with high Comprehension scores fail to become aroused within the assessment situation.

A more skeptical interpretation of the non-arousal data would hold that whether or not positive or negative contingencies were to follow plethysmographic assessment, subjects would nevertheless wish to manifest less arousal to the deviant stimuli. A prisoner who understands that his behavior in the research setting could have no official influence on his future might still wish to practice the inhibition of arousal to deviant stimuli, whether in the presence of a researcher or not. Accordingly, it is possible that the subjects in this study did, indeed, seek to inhibit their own arousal. However, accepting this view, it is also true that those who did so were those with higher Comprehension scores. Either way, the practical conclusion remains that plethysmographic assessment of offenders with higher social comprehension is likely to be a less fruitful undertaking.

Arousal/Non-Arousal and Verbal IQ - Performance IQ Differences

The significant differences in Verbal IQ - Performance IQ splits between the aroused and non-aroused groups may further support the hypothesis that subjects' sexual responses are affected by cognitive factors as well as by stimulus qualities.
Verbal IQ measures verbal ability, especially prior verbal learning, whereas Performance IQ primarily measures visual-spatial ability, especially visual perception. When PIQ was greater than VIQ, arousal to plethysmographic assessment was more likely. In contrast, when VIQ was greater than PIQ, no arousal was likely. Thus, the stronger a subject's visual perceptual skills, which involve paying attention to the immediate environment, were in comparison to that subject's verbal abilities, which involved prior learning, the more likely he was to become aroused. On the other hand, the stronger a subject's verbal-memory skills were in comparison to that subject's visual perceptual skills, the less likely he was to become aroused. It is supposed that these results support the previously formulated interpretation of the Comprehension score differences. Both sets of results point to a cognitive mediation of immediate perceptual processes which serves to reduce the sexual response to sexual stimuli within the assessment situation.

In summary, two factors related to intelligence were found to differentiate the aroused and non-aroused groups significantly: Comprehension subtest score and Verbal IQ - Performance IQ differences. The comprehension subtest measures social understanding, although the exact mechanism whereby social understanding results in an absence of arousal is not known. Two explanations for this finding were considered. The first explanation is that higher Comprehension pedophiles are
more able and more motivated to inhibit their arousal. This explanation considers the failure to produce arousal as deliberate. The second explanation is that higher comprehension subjects are not more motivated to inhibit their arousal since no advantage was to be gained for doing so. According to the second explanation, arousal is inhibited because greater attention is paid to the social context within which sexual stimuli arise. Such cognitive awareness mediates and diffuses the sexual response to the sexual stimuli. From this point of view, the failure to inhibit arousal is not considered to be deliberate. The first explanation remains the most parsimonious.

**Other Findings**

*Arousal/Non-Arousal and Age*

One other relationship was found to be significant under two definitions of non-arousal, and marginally significant under the third definition of non-arousal (p=.056). Age was related to the outcome of plethysmographic assessment, in that older subjects were less likely to produce arousal during the assessment than were younger subjects. This finding appears to be an empirical validation of the generally understood notion that sexual vigor declines with age. No subject under the age of 38 failed to become aroused, while 17 of 25 subjects over the age of 40 were non-aroused. At first glance, this finding might suggest that age is a predictive variable which could reliably indicate which subjects were more likely to become aroused during the
assessment. However, examination of the Comprehension subtest scores of the aroused and non-aroused subjects over 40 indicates that it is not the age which is crucial but their score on the Comprehension subtest (see Table 17). It seems that most of those with higher Comprehension subtest scores were also older. Regardless, it seems likely that an interactional effect is present, with older subjects being less likely to demonstrate arousal.

Logistic regression analyses were undertaken in order to explore the relative efficiency with which the three variables, Comprehension subtest score, Verbal IQ - Performance IQ differences and age, could classify the current sample into aroused and non-aroused subjects.

The results of the logistic regressions clearly showed that Comprehension was the strongest variable and had the most predictive utility of the three variables.

Arousal/Non-Arousal and Other Psychometric Variables

In addition to Intelligence, a number of other variables were included in the study as potential explanations for any group differences in arousal that might occur. These variables were social desirability, anxiety and neuropsychological impairment.

Social desirability reflects the degree to which a subject answers questions in a socially desirable direction. It was
<table>
<thead>
<tr>
<th>Age</th>
<th>Arousal Definition</th>
<th>15</th>
<th>13</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 - 38</td>
<td>Aroused</td>
<td>8.3 (14)</td>
<td>8.3 (14)</td>
<td>8.3 (14)</td>
</tr>
<tr>
<td></td>
<td>Non-aroused</td>
<td>(0)</td>
<td>(0)</td>
<td>(0)</td>
</tr>
<tr>
<td>40 - 65</td>
<td>Aroused</td>
<td>7.9 (10)</td>
<td>8.3 (12)</td>
<td>6.6 (8)</td>
</tr>
<tr>
<td></td>
<td>Non-aroused</td>
<td>11.7 (15)</td>
<td>11.9 (13)</td>
<td>11.8 (17)</td>
</tr>
</tbody>
</table>

( ) = Number of subjects
thought that individuals with a higher score on the Crowne-Marlowe Social Desirability Scale might be less likely to demonstrate deviant arousal. These subjects would suppress arousal knowing that it was a socially undesirable response to the stimuli. However, the aroused and non-aroused groups did not differ in social desirability. Therefore a subject's tendency to answer items in a socially desirable direction was unrelated to whether he did or did not produce arousal in this study.

Anxiety has also been recognized as a variable which affects a subject's ability to become aroused within the assessment situation. Since the work of Masters and Johnson (1971) anxiety and sexual arousal have been thought to be negatively correlated, with higher levels of anxiety being associated with lower arousal. Anxiety was measured by means of the State Anxiety Inventory, which was administered immediately after the plethysmographic testing and assessed a subject's anxiety level at that moment. The aroused and non-aroused groups did not differ in state anxiety. Therefore a subject's anxiety level was unrelated to whether he did or did not produce arousal in this study.

Recently, a link between neuropsychological impairment and pedophilia has been reported. For example, Hucker et al. (1986) recently concluded that pedophiles have a left temporal-parietal dysfunction. However, as Hucker et al. (1986) did not conduct plethysmographic assessments with their sample, the relationship between neuropsychological impairment and the outcome of
plethysmographic assessment is unknown. In this regard it has been suggested that such neuropsychological impairment may lead to poorer impulse control and to perseverative actions. As such, pedophilia would be characterized as perseverative behavior coupled with poor impulse control. In the current study, a measure of frontal lobe dysfunction, the Trail Making Test, was administered. However, the aroused and non-aroused groups did not differ in terms of cognitive deficits as measured by the Trail Making Test. Therefore a subject's cognitive impairment, at least as assessed by the Trail Making Test, was unrelated to whether he did or did not produce arousal.

In summary, no relationships were found between scores on the above measures of social desirability, anxiety, or cognitive impairment (as measured by the Trail Making test) and the outcome of plethysmographic assessment.

**Pedophilic Deviance Quotients**

A Pedophilic Deviance Quotient (PDQ) is derived from a ratio of the average arousal to children divided by the average arousal to adults. Marshall et al. (1986) had reported a correlation of -.33 between intelligence, as measured by Raven's Progressive Matrices and PDQs. Analyses were undertaken to examine the correlation between intelligence as measured by the WAIS-R and PDQs in the current sample. No significant relationships were found. Therefore, the relationship between
PDQs and intelligence reported by Marshall et al. (1986) was not replicated.

Marshall et al. (1986) also examined the relationship between intelligence and PDQs by examining the average PDQs for three levels of intelligence: <80, 80-100 and >100. They found subjects with IQs lower than 80 to have significantly higher PDQs, indicating greater ratios of arousal to children compared to adults. Similar analyses were undertaken with the current sample (see Table 16). However, there were no significant relationships between levels of intelligence and PDQs, even though three different IQ measures were used (FSIQ, VIQ, PIQ) and two different classifications of intelligence (<80, 80-100, >100; <90, 90-110, >110).

The lack of association between intelligence and PDQ is important since decisions as to whether or not an individual is likely to be a pedophile are based at least partially upon the PDQ. If the PDQ was a function of intelligence as well as sexual arousal, it would be difficult to know whether a particular PDQ was reflecting intelligence or arousal. Thus, the finding of no relationship between intelligence and PDQ establishes greater confidence in the PDQ as a measure of arousal, something which the Marshall et al. (1986) finding had cast into doubt.

Along this line, the writer has reservations about the concept of PDQs inasmuch as they involve no quantification of the actual amplitude of a response. Thus, a subject with arousal
to children of 35% and arousal to adults of 30% has a PDQ of 1.17. A subject with a far larger amplitudinous response of 70% to children and 60% to adults, likewise has a PDQ of 1.17. Simply reporting the PDQs misses potentially valuable information. However, as a shorthand descriptive statistic it seems to be unrivaled.

Implications

Theoretical Implications: Support for Model 3, S - Cognition - R

In the Introduction, three theoretical models which have guided clinical practice and research investigation with the plethysmograph were discussed. Model 1 was a simple stimulus-response model which was considered untenable given a number of studies which had demonstrated state effects on the outcome of plethysmographic assessment. Model 2 was a stimulus-state-response model which accounted for the effects of subjects' internal states upon sexual arousal. However, this model was unable to account for the finding of Marshall et al. (1986) that intelligence was related to the outcome of plethysmographic assessment. Therefore, Model 3 was proposed, in which cognitive features were thought to systematically modify sexual responses to sexual stimuli.

The current study found that the level of a cognitive factor, as measured by the Comprehension subtest, was highly postdictive of whether a subject would produce arousal or not.
Since social comprehension is understood to be a trait variable, that is, to change little over time and across situations, this result was taken as support for the view that cognitive factors are important in understanding the plethysmographic assessment procedure. Within the assessment situation, many subjects routinely do not become aroused, a fact which has long frustrated clinicians who have sought a more perfect assessment tool. The results of this study suggest that, at least with the present sample, failure to produce arousal was not a random occurrence but rather occurred in accordance with identifiable correlates. Pending replication, it is postulated that a model of plethysmographic assessment should routinely consider cognitive mediational factors.

Clinical Implications

The principal finding of this study is that the level of a cognitive factor, social comprehension, has a strong bearing on whether or not pedophiles produce arousal within the plethysmographic assessment situation. It is suggested that pedophiles with a higher Comprehension subtest score possess more verbally mediated cognitive frameworks within which the sexual stimuli are processed. These higher Comprehension subjects are therefore reacting to their internally generated frameworks as well as to the actual qualities of the stimulus. As a result, their responses to the sexual stimuli are less primitive and more cognitively diffuse. This finding has at least three clinical implications.
The first implication is that a subject's failure to become aroused, if they have a high Comprehension subtest score, may not be the result of deliberate faking but may be due to cognitive style. Consequently, failure to become aroused during assessment for parole or other judicial purposes is not necessarily a sign of uncooperativeness or deliberate deception.

The second implication is that a subject's Comprehension subtest score can be used to predict the probability that he will demonstrate arousal during a plethysmographic assessment. While attempts can, of course, still be made to assess an individual with a high Comprehension score, the plethysmographic testing can probably be terminated sooner, when early signs of non-arousal confirm the postulated expectation of non-arousal based upon the Comprehension subtest score.

The third implication of the current findings is speculative, but may be the most important. Different treatment strategies may be appropriate for the low Comprehension aroused group and the high Comprehension non-aroused group. The low Comprehension group's sexual arousal was uninhibited by the assessment situation. This group may be best treated by aversion conditioning techniques which would reduce sexual arousal without extensive cognitive processing. However, in the high Comprehension group, a cognitive approach aimed at modifying schemas that support inappropriate arousal may be appropriate.
A fourth implication derives from the finding that in this study, intelligence was unrelated to Pedophilic Deviance Quotients. The current study would indicate that once aroused, the resultant assessment profiles reflect sexual preferences and not intelligence.

In order for assessors to profit from the first three of these clinical implications, the Comprehension subtest would need to be administered prior to plethysmographic assessment. Fortunately, the Comprehension subtest is quickly and inexpensively administered. Subjects who answer more of the questions correctly are less likely to become aroused during plethysmographic assessment. Their plethysmographic assessments will therefore take less time as an assessment is terminated when it is clear no arousal is forthcoming. Thus administering the test will add little time to the assessment procedure.

As a guide to subject classification, the Comprehension cutoff score developed on the current study can be used. A score of 11 or higher predicts non-arousal, a score of 10 and below predicts arousal.

**Generalizability**

It is reasonable to ask about the generalizability of the present findings. It is understood that the more stringently a sample is defined in the first place, the fewer are the samples to which the results apply.
By the most stringent standards, the results of the current study should be applied only to federally incarcerated, self-confessed pedophiles who volunteer under conditions of confidentiality to undertake a plethysmographic assessment. However, only 40% of pedophiles approached agreed to participate in the study. Therefore, there are limits to the generalizability of the results. Recognizing the limitations of such a precisely defined sample, attempts were made to extend the results to include all incarcerated pedophiles whether they volunteered or not. This was done by comparing some file data collected on those who participated and those who were approached and declined. Analysis of this data indicated that those who participated and those who did not participate did not differ on the variables of age, current sentence length, or time served. However, those who volunteered possessed more verbal intelligence as measured by the Shipley-Hartford test.

The real limit on generalizability probably relates to how "pedophilic" a sample is being considered. All of the subjects in this study had received federal time, that is sentences equal to or greater than 2 years. The same results might not obtain in a sample that is "less seriously pedophilic". That the current population was "seriously pedophilic" can be demonstrated not only by the fact that they are serving federal time but by reference to the large number of their victims. The victim estimates were made from a cursory examination of the file material available. Since it is acknowledged that inmates are
charged with a fraction of their offenses this sample can be assumed to have had even more victims than were estimated.

Another limit to the generalizability of the results from this study is the age of the population studied. As noted previously, this sample was older than most samples of pedophiles reported upon recently. One implication of the age of the sample is that they are likely to be more seriously pedophilic, and are likely to have had on average more sexual contacts with children.

Future Research

A number of avenues for further research are suggested by the results of the current study.

One further kind of study would examine the influence of age upon responsiveness to sexual stimuli.

A second kind of investigation would study a sample of lower IQ offenders, not convicted of sexual crimes. Marshall et al. (1986) stated that "low intelligence men may readily respond sexually to young females at assessment although this may not necessarily predict deviant sexual acts." The current study, while not examining such offenders, did find a relationship between social comprehension and the production of arousal in pedophiles during the plethysmographic assessment. It is possible that low Comprehension score subjects may show an
increased likelihood of arousal to children, regardless of whether they have actually, or will actually commit deviant sexual acts.

A third type of future research would study a mixed group of individuals being assessed before their appearance in court. The relationship between Comprehension subtest score and degree of arousal could be explored in this sample in order to evaluate how widely the relationship holds. Does Comprehension predict arousal only in pedophiles, or might it predict behavior in other populations as well.

A fourth avenue for future research is one which the current study had hoped to explore, but could not because only one false negative subject was obtained. The question remains, do pedophiles who appear normal upon plethysmographic assessment differ on a measure of intelligence from those who appear deviant.

A fifth kind of future investigation would replicate and extend the current finding concerning the relationship between the production of arousal and social comprehension. For example, individualized assessment tapes—which pay particular attention to the contextual cues that appear to have been important in a pedophile's offence pattern may prove more successful in eliciting arousal.
Summary

Although plethysmographic assessment has been used for research and clinical purposes for over 25 years, little has been discovered about the relationship between plethysmographic assessment and intelligence. The first published account documenting such a relationship did so incidentally, and used the Ravens Progressive Matrices. The current study undertook specifically to examine an hypothesized relationship between the outcome of plethysmographic assessment and intelligence. To this end the WAIS-R, a more widely used intelligence test than the Ravens Progressive Matrices was employed. Results indicated that both the Comprehension subtest score and Verbal IQ - Performance IQ differences were significantly related (p=.001) to the presence or absence of arousal during plethysmographic assessment. However, for those subjects who became aroused, results also indicated that intelligence was unrelated to the amount of deviancy exhibited (PDQ). These results suggest that cognitive mediational factors should be considered for both a theoretical and a clinical understanding of plethysmographic assessment of pedophilia.
APPENDIX A

Memo from the Corrections Service of Canada

To: R.P.C. Patients
From: Doctor Luke Glancy
Subject: Research Study

Mr. Conrad Bowden is asking you to participate in a research study.

The Correctional Service of Canada would like you to be aware of the following:

1. This study is being undertaken independently of the Correctional Service. Your participation in the study is of no concern to the Correctional Service and will not affect your standing within the correctional system. Do not feel under any pressure to participate in this study if you do not wish to.

2. This study is not a therapy program, and no therapeutic benefits are likely to result from participation in the study.

3. Mr. Bowden will answer any further questions or concerns you may have when he meets you.

Dr. Luke Glancy
Clinical Director
APPENDIX B

DESCRIPTION OF THE STUDY

This experiment is designed to assess my sexual arousal to a variety of sexual episodes. As well, the experiment assesses my responses to a number of questionnaires. The purpose of the experiment is to explore the relationship between scores on the questionnaires and sexual arousal.

My sexual arousal will be measured while I listen to verbal descriptions of a variety of sexual episodes. While I listen to these sexual episodes my arousal will be measured by a small penile transducer, an apparatus which I will place around my penis while I am alone in the privacy of the laboratory. This device will only record changes in the circumference of my penis and cannot hurt me physically. A male experimenter, Conrad Bowden, will record the data. However, I will be seated in privacy throughout the experiment.

During any stage of the experiment, I may voluntarily cease participating without penalty of any kind.

The recordings of my sexual arousal and any other information I disclose to the experimenter will be totally confidential. I will be assigned a number which I must remember throughout my involvement with the experiment. Recordings of my sexual arousal and the information I provide in any of my interviews will be coded according to this number and no
government department, nor any agency or individual will be able to trace these records to me.

The overall results of this experiment will be communicated to me at the end of the experiment and I may be appraised of my own results if I so request.

If at any time during my participation in this experiment I become dissatisfied with my treatment, I understand that I am at liberty to take my complaints to Dr. Roger Blackman, Chairman of the Department of Psychology, Simon Fraser University. I will be given a stamped and addressed envelope for this purpose if I desire.

I agree of my own free will to participate in this experiment. This consent and release is given freely and I affirm that I am not acting under fraud, duress, or menace of any person whomsoever.

I authorize the use of all records and personal data derived from these experiments for research purposes and publications provided there is no disclosure of the identity of the undersigned.

I acknowledge reading a memo issued by the Correctional Service stating: 1) my standing within the Correctional Service will not be affected by my decision to participate or not participate in the study; and, 2) the study is not a therapy program.
I also acknowledge having been allowed 48 hours within which to consider my decision to participate in this study.

Date:_________ Name:_________

Witness:_________ Signature:_________
Screening Questionnaire

1. Subject number

2. Age

3. Reasons for incarceration

4. Relation of victim to offender

5. Sex of preferred child partner

6. Medication in the last 2 months

7. Drugs in last 36 hours

8. Difficulty in getting erections

9. Previous plethysmographic assessments
REFERENCES


