Quantitative Reanalysis of Aggregate Data on IQ in Sexual Offenders

James M. Cantor
Centre for Addiction and Mental Health

Ray Blanchard
Centre for Addiction and Mental Health
and University of Toronto

Lori K. Robichaud
Centre for Addiction and Mental Health
and University of Windsor

Bruce K. Christensen
Centre for Addiction and Mental Health
and University of Toronto

General intelligence has been the most commonly studied neuropsychological characteristic of sexual offenders for over 70 years. Results, however, have been highly inconsistent. To assess whether sexual offenders score lower in IQ than nonsexual offenders and to explore which sexual offense characteristics relate to IQ, the authors reanalyzed all reports providing sufficient information. Data spanned 236 samples, comprising 25,146 sexual offenders and controls. The literature contained sufficient information to permit comparison of adult versus juvenile sexual offenders, offenders targeting children versus adults, offenders targeting their own versus unrelated children, and offenders targeting boys versus girls. Results confirm the association between IQ and sexual offending and suggest that previous discrepancies are attributable to how many pedophilic individuals were in each sample.

Keywords: intelligence, neuropsychology, pedophilia, sex offenders, sexual abuse

Investigations of the intellectual capacity of sexual offenders now span 8 decades, the earliest such data appearing in 1931 (i.e., Frank, 1931). Despite producing scores of studies, however, investigators have reached few, if any, reliable conclusions. As reviewed in the following, some investigators found samples of sexual offenders to have lower IQs than control samples, whereas other investigators found no such difference. Some investigators found differences between certain subtypes of sexual offenders, such as offenders who target child versus adult victims, whereas other investigators found no difference.

The persistence of interest in this topic, despite the repeated contradictions in the results, likely relates to the potential implications of such findings for the etiology of sexual offending. Brain dysfunction has been hypothesized to contribute to sexual offending since the 19th century (e.g., Krafft-Ebing, 1886/1965). Investigators have attempted to apply several neurologic and neuropsychological techniques to identify specific dysfunctions; however, their findings have been generally contradictory (for review, see Blanchard, Cantor, & Robichaud, in press). Moreover, the number of such studies has been relatively small. There has been, in contrast, a rather large number of studies providing IQ data on sexual offenders. This provides an opportunity for formally assessing the reliability of findings on the general brain function of sexual offenders. It also opens up the possibility of examining IQ differences between subtypes of sexual offenders (e.g., those who target child vs. adult victims), which might, in turn, indicate whether neurologic factors play different roles among sexual offender subtypes.

The reports of this literature have used several systems to subtype sexual offenders (see Supplemental Tables 1 and 2, which are available on the Web at http://dx.doi.org/10.1037/0033-2909.131.4.555). The most common of these has been to categorize the offenders according to the characteristics of their victims, specifically, their victims’ age, their victims’ sex, and their familial relationship with their victims (status as extrafamilial vs. intrafamilial—that is, incest). These offender characteristics are also those that appear as diagnostic categories and specifiers in the Diagnostic and Statistical Manual of Mental Disorders (4th edition, text revision [DSM–IV–TR]; American Psychiatric Association, 2000). A large proportion of the investigations, however, used no classification scheme at all, collapsing all sexual offenders into a single, heterogeneous sample. Thus, if only certain types of sexual offending behaviors relate to intellectual capacity, then the
contradictory findings thus far reported might be reflections of the varying sample compositions.

Of the characteristics explored in the literature, the distinction between sexual offenders against children and sexual offenders against adults has received the most attention, and a variety of evidence (other than IQ) confirms victim age to be a relevant grouping variable. Sexual offenders against children show greater penile tumescence in reaction to erotic verbal or visual stimuli depicting children than do sexual offenders against adults (e.g., Blanchard, Klassen, Dickey, Kuban, & Blak, 2001). Rates of non-right-handedness have been reported to be higher in men convicted of child molestation than in men convicted of no offenses (Bogaert, 2001). Similarly, levels of non-right-handedness have correlated with penile responses to stimuli depicting children and with offenders’ numbers of child sexual victims, but not with responses to stimuli depicting adults or with offenders’ numbers of adult sexual victims (Cantor et al., 2004, 2005). Finally, pedophilic men have reported greater frequencies of childhood head injuries (before age 13) than nonpedophilic men in three separate samples (Blanchard et al., 2002, 2003).

No consensus exists regarding whether sexual offenders who target intrafamilial victims represent a group distinct from sex offenders who target extrafamilial victims, and only one study has compared them on IQ (Marshall, Barbaree, & Christophe, 1986). Similarly, the field has not yet reached any conclusion regarding whether sexual offenders who target boys differ cognitively from those who target girls. Although the literature as a whole contains substantial numbers of studies both of adult and juvenile sexual offenders, no investigation has directly compared adult with juvenile sexual offenders. Indeed, no single investigation contained both types of samples. This likely results from the methods researchers used to gather their study participants—recruiting from facilities that handle adult and juvenile offenders separately.

Although one would in general anticipate contradictory findings to result from researchers’ use of different assessment instruments, sampling methods, and statistical techniques, several factors specific to sexual offender studies further complicate their analysis. Some researchers have substituted penal codes or nonspecific, common-language descriptors for behaviorally relevant grouping variables. Penal codes only inaccurately describe sexual offenses in situations where, for example, a man charged with sexual assault plea-bargains for a nonsexual charge, such as battery or common assault. (Only some reports, however, were explicit regarding this problem, such as Ford & Linney, 1995, and Vuocolo, 1969.) Such misclassifications would have the effect of reducing observed group differences. Furthermore, common language descriptors, such as the word rape, insufficiently delineate the sexual offense committed. Some studies compared men described as child molesters with men described as rapists (e.g., Henn, Herjanic, & Vanderpearl, 1976), implicitly defining rape to apply only to offenses against adult victims. Other studies, however, implicitly defined rape to apply to the use of coercion generally, regardless of the age of the victim or the actual behaviors engaged in (e.g., Awad, Saunders, & Levene, 1984). Thus, for the many studies that referred to a sample of rapists but did not provide the victims’ characteristics, it remains unclear whether the victims were adults or children (or both), whether the victims were female in all cases or female in only some cases, and so on. For greatest accuracy, any review of this literature must consider the original samples’ characteristics independently of the descriptors used by the original authors.

Although several studies compared offenders against children with offenders against adults, studies differed with regard to what constituted a child and an adult. Some investigators composed samples of offenders against children from men who offended against children ages 13 and younger (e.g., Cohen et al., 2002; Quinsey, Arnold, & Pruesse, 1980), whereas other investigators used men who offended against children of any age under 18 (e.g., Fisher & Howell, 1970; Rau, 1991). Conversely, some investigators defined an adult victim as one who was 18 years of age or older (e.g., Valliant, Gauthier, Pottier, & Kosyna, 2000), whereas other investigators used ages as young as 15 years as their cutoff (e.g., Quinsey et al., 1980; Rader, 1977). Thus, some studies would have classified a given offender as an offender against adults, and other studies would have classified the same offender as an offender against children. One would expect the more liberal definition of child to yield samples with fewer pedophilic men or with men more attracted to prepubescent children than to pubescent children. Thus, some of the inconsistency among investigations might be attributable to stringency in inclusion criteria.

Also complicating both narrative and quantitative review is the inaccurate and contradictory reporting in this literature. Some investigators described the same sample in more than one publication but provided different mean scores in each (cf., Ellis, 1951; Ellis & Brancale, 1956). Some researchers conducting retrospective chart reviews indicated the scope of their samples inadequately or inconsistently, preventing readers from determining the degree of overlap across publications. Several reports contained errors in their data tables (marginal totals unequal to the sum of cell entries, for example) or provided only incomplete descriptions of their methods or results (e.g., failing to report the IQ tests used; the sample or subsample sizes; the test statistics in support of the conclusions drawn; the sample variances; or, in some cases, even the samples’ mean scores). The Comments columns of Supplementary Tables 1 and 2 (http://dx.doi.org/10.1037/0033-2909.131.4.555) indicate such errors and provide the corrected information, where possible.

Although relatively many studies recorded the IQ of sexual offenders, the samples tended to be quite small. Indeed, the number of sexual offender samples being reanalyzed here exceeds the sample size in 93.6% of the sexual offender samples themselves. In this circumstance, meta-analytic techniques provide a method of reexamining group differences with far greater power than previously available.

Despite its shortcomings, this body of literature possesses two characteristics conducive to a meaningful, quantitative reanalysis. First, for a sizable proportion of the samples, investigators reported on men who were homogeneous on an individual characteristic of interest, such as the sex of the victims they chose. The use of samples that were homogeneous for a given characteristic permits those samples to be contrasted on that characteristic (e.g., samples of men sexually victimizing their own children contrasted with samples of offenders victimizing extrafamilial children). Second, the great majority of investigations reported their results in a common metric: IQ points. The use of a common metric makes it possible to use aggregate data—summary statistics, such as means,
to represent whole samples rather than individuals—and thus to calculate means for groups and subgroups of offender samples, collapsing across studies. It also obviates the need to convert scores into effect sizes, such as Cohen’s $d$, which require both a reference or control group and an estimate of the samples’ variance. In effect, metaregression compares the mean of all the experimental groups with the mean of all the control groups. This differs from the more familiar form of meta-analysis, which converts all scores to a single metric and then calculates the difference between each experimental group and its control group, comparing that overall value with zero.

Analyses of aggregate data—the basis of the meta-analyses in this study—most commonly appear in epidemiological or census-based research. In that context, the interpretation of aggregate data can pose a problem called the ecological fallacy: falsely ascribing characteristics of groups to individuals (e.g., Robinson, 1950). For example, an epidemiologist may compare the rates of a cancer with the rates of exposure to a potential carcinogen across several cities, with the rates in each city as the units of analysis. Although the epidemiologist may find a significant correlation between the rates of cancer and the carcinogen, it remains unknown whether the individual people who were diagnosed with cancer were the same people who were exposed to the potential carcinogen. This problem results from individual cases being grouped according to some characteristic other than their status on the dependent or independent variables (in this example, being grouped by city of residence rather than by having cancer or by having been exposed to the potential carcinogen). Interpretation of the following analyses, however, is not susceptible to the ecological fallacy; all members of the group were selected exactly because they exhibited the characteristic defined by the independent variable. Thus, the characteristics that define the group describe all members of that group.

To verify whether intellectual functioning relates to specifically sexual offending and to identify which sexual offense characteristics most reliably relate to intellectual capacity, we collected and reviewed all identifiable reports of the intelligence of sexual offenders, quantitatively reanalyzing those that provided sufficient data. The literature provided sufficient data to address five questions: (1) Do males committing sexual crimes differ in IQ from males committing nonsexual crimes? (2) Do sexual offenders against adults differ in IQ from sexual offenders against children? (3) Of the sexual offenders against children, do those selecting intrafamilial children differ in IQ from those selecting extrafamilial children? (4) Of the sexual offenders against children, do those selecting boys differ in IQ from those selecting girls? (5) Do juvenile sexual offenders differ in IQ from adult sexual offenders?

**Method**

The samples included in the reanalyses were identified by computer-assisted literature searches of MEDLINE, PsycINFO, and Dissertation Abstracts International and by manual review of the reference lists provided in the resultant reports. As already noted, the great majority of the available studies provided their results in a common metric; thus, a study did not need to have a control group in order to be included in the reanalyses.

**Exclusion Criteria**

Some researchers studied samples of sexual offenders that were systematically unrepresentative of sexual offenders with regard to IQ. Such studies investigated, for example, only mentally retarded offenders (e.g., Day, 1994) or physicians or priests who committed sexual offenses (e.g., Langevin, Glancy, Curnoe, & Bain, 1999). Such investigations were excluded from reanalysis. We also excluded case studies (e.g., Weiner, 1962) and the one sample of sexual offenders against animals, which consisted of 2 people.

Duplicate information was excluded as follows: In the case of complete overlap in reports, we retained only the report that provided the most detailed breakdown of the data (cf. Ellis, 1951; Ellis & Brancal, 1956).

When a sample was reported and then later expanded and rereported in a subsequent document, we retained only the report with the larger sample (cf. Firestone, Bradford, Greenberg, & Larose, 1998; Firestone, Bradford, Greenberg, & Nunes, 2000). We then excluded reports for which an individual person could simultaneously appear in multiple samples; for example, when a sex offender had multiple types of victims, some researchers would classify him both as an offender against children and as an offender against adults, violating the independence of the samples’ scores (e.g., Gebhard, Gagnon, Pomeroy, & Christenson, 1965).

One prolific research group, Langevin, Hucker, and colleagues, published eight reports between 1985 and 1989, reporting mean IQs in relation to different aspects of a single, but continuously expanding, database of offenders being assessed at their facility. Although those authors indicated in some of the reports that the sample overlapped with those in some other reports, the reports included very large numbers of independent variables, and the sample sizes used in the statistical analyses varied greatly. It was therefore difficult to discern the degrees of overlap specifically in the IQ data among those reports. Thus, we included the single largest sample of each independent type of offender across the set of reports from these researchers.

Next, we excluded individual samples from within a report when it was unclear whether the sample actually comprised sexual offenders (despite the authors’ description of the sample as such). Ambiguity in a sample’s status occurred when (a) the behaviors constituting the offense would not reasonably be considered offenses by contemporary standards (e.g., engaging in consensual homosexual sexual contact, disseminating “obscene” material), (b) investigators’ operational definition of a group would have combined offenses with victims and offenses without victims (e.g., a sample convicted of “sodomy,” which could have included both coerced sexual contact and consensual sexual activity between two adults), and (c) the study did not define its terms sufficiently to determine what the behaviors were (e.g., “debauchery”).

Some offenses, such as solicitation and bigamy, constitute sexual offenses legally but do not necessarily involve a victim or suggest any psychiatric condition in themselves (polygamy is still normative in some communities, including small, isolated communities in North America). In order to focus the reanalyses on behaviors most likely to represent an underlying disorder (typically, one of the paraphilias), we did not include prostitution, solicitation, bigamy, polygamy, or sexual harassment as offenses. Such determinations appear in Supplemental Tables 1 and 2 (http://dx.doi.org/10.1037/0033-2909.131.4.555).

In total, the following reanalyses encompassed 75 reports, describing 236 independent samples (140 sexual offender samples plus 96 comparison samples), representing 25,146 individual people (7,045 sexual offenders plus 18,101 comparison participants). Supplemental Tables 1 and 2 (http://dx.doi.org/10.1037/0033-2909.131.4.555) indicate the samples included in the reanalyses with an asterisk (*) in the sample size columns.

**Variables Recorded**

**Full-Scale IQ (FSIQ).** Most reports provided the mean IQs of their samples directly. Several reports, however, provided frequency distribu-
tions instead of sample means. For each of those reports, we calculated the sample’s mean score to be the weighted average of the approximate midpoints of each category in the frequency distribution. A dichotomous variable recorded whether the original report provided its data as mean IQ scores or as frequency distributions, to allow for subsequent control of any mean difference associated with the alternate methods of acquiring IQ data points.

Some reports used verbal descriptors of categories of intelligence, rather than numeric ranges; that is, the reports divided their samples into “average,” “dull normal,” and “moron,” and so forth. Nonetheless, the numeric equivalents we used for these descriptors appear in the Comments columns of Supplemental Tables 1 and 2 (http://dx.doi.org/10.1037/0033-2909.131.4.555). When a report provided the sample’s mean Verbal IQ (VIQ) or Performance IQ (PIQ) subscale score, but not the FSIQ score, the subscale score (or their average, when both were available) substituted for the FSIQ score. A categorical variable recorded whether the report provided the FSIQ or whether VIQ, PIQ, or their mean substituted for FSIQ.

**Publication status.** Also recorded was whether a sample came from a published or unpublished report. Government reports were coded as published.

**IQ test used.** As noted already, some variance in samples’ mean IQ scores could be expected as a result of different samples receiving different IQ tests. To permit the analysis of any such differences, we recorded the IQ test used for assessing each sample. Specifically, we recorded which samples received the Wechsler-Bellevue Scales (6 samples; Wechsler, 1939), the Wechsler Adult Intelligence Scale (36 samples; Wechsler, 1955), the Wechsler Adult Intelligence Scale—Revised (36 samples; Wechsler, 1981), the Wechsler Intelligence Scale for Children (WISC, 8 samples; Wechsler, 1949), the Wechsler Intelligence Scale for Children—Revised (WISC–R, 20 samples; Wechsler, 1974), the Wechsler Intelligence Scale for Children—Third Edition (WISC–III, 9 samples; Wechsler, 1991), a short-form of any of the Wechsler Intelligence Tests (19 samples); the Stanford–Binet Intelligence Test (1 sample; Terman & Merrill, 1937), any of the Army General Classification Tests (22 samples; e.g., Science Research Associates, 1947), the Shipley Institute of Living Scales (3 samples; Shipley, 1946), any of the Raven’s Matrices tests (9 samples; e.g., Raven, 1960), and “other” (46 samples). One additional code indicated that the original authors did not report which IQ test they used (79 samples). The total number of samples receiving each intelligence test exceeds the total number of samples in the reanalyses because some investigators used more than one test to assess their samples. This most frequently occurred with studies analyzing clinical data archived in institutions that changed their clinical instruments over time (e.g., they first used the WISC, then the WISC–R, then the WISC–III, etc.).

**Offense status.** In order to answer Question 1, we identified samples as being composed of (a) sexual offenders or (b) nonsexual offenders, that is, men convicted of crimes other than those of a sexual nature. Nonsexual offenders represent an ideal comparison population for sexual offenders men convicted of crimes other than those of a sexual nature. Nonsexual offenders against children plus extrafamilial adults. To avoid confounding any effects of victim age with any effects of familiality, these factors themselves are not fully crossed—the literature includes no samples of sexual offenders against extrafamilial children. Although the results presented by this variable should be interpreted with the broader definition of extrafamilial offenses. This variable subclassified the sexual offenders into “child molesters” and “rapists.” In such cases, despite that the original authors did not provide the actual age range of the rapists’ victims, the victims could be reasonably presumed to be adults. The summaries made appear in Supplemental Tables 1 and 2 (available on the Web at http://dx.doi.org/10.1037/0033-2909.131.4.555).

Also recorded was the definition of child. In describing the maximum age of the victims, some authors used inclusive age ranges (such as victims ages 15 or under) and others used exclusive ranges (such as victims under 15 years of age). To homogenize the definitions, we recorded all the age cut-offs as exclusive ranges. Thus, a sample composed of offenders against children under 15 received a value of 15, whereas a sample composed of offenders against children aged 15 or under received a value of 16.

**Sexual offenders against intrafamilial children versus sexual offenders against extrafamilial children.** This variable subclassified the sexual offenders against children into (a) samples representing sexual offenders against extrafamilial children and (b) samples representing sexual offenders against intrafamilial children. Analogous to the coding of victim age group, this variable included levels to capture samples whose familiality was unclassifiable: (c) samples that had no child victims, (d) samples with victims whose familial relationships with their victims were either mixed (i.e., the sample consisted of some combination of men who offended against intrafamilial and extrafamilial children or both) or unknown (i.e., not provided in the original report), and (e) samples for which it is unknown whether there were child victims. As was true for victim age group, victims’ familiality could sometimes be inferred from context when the original report did not provide it. Such inferences appear in Supplemental Tables 1 and 2 (available on the Web at http://dx.doi.org/10.1037/0033-2909.131.4.555).

For emphasis, this variable represents the familiality of victim children, not the familiality of victims regardless of victims’ age. Although it could be of interest to evaluate the effects of victims’ age independently of the effects of familiality, these factors themselves are not fully crossed—the literature includes no samples of sexual offenders against intrafamilial adults. Thus, comparing offspring with extrafamilial offenders would, in practice, compare samples of sexual offenders against intrafamilial children with samples of sexual offenders against extrafamilial children plus extrafamilial adults. To avoid confounding any effects of victim age with any effects of victim familiality, we coded this variable to represent a proper subset of the samples of sexual offenders against children.

Unfortunately, several authors failed to indicate whether offenses against step-children constituted intrafamilial offenses. Other authors indicated specifically that offenses against biological relatives and against step-relatives both represented intrafamilial offenses. Thus, the results presented by this variable should be interpreted with the broader definition of intrafamilial familiality.

**Sexual offenders against girls versus sexual offenders against boys.** The coding for the sex of offenders’ child victims paralleled that for the familiality of offenders’ child victims, reclassifying the samples into (a) offenders against boys and (b) offenders against girls. The groups representing unclassifiable samples included (c) samples that had no child
victims, (d) samples with victims whose sexes were mixed or unknown, and (e) samples with victims whose ages were mixed or unknown. Samples composed of individuals receiving formal diagnoses of pedophilia were coded according to the sex of the children in which they were most interested.

**Offender age group.** A dichotomous variable encoded which samples represented adult members and which represented juvenile members. Not all authors used 18 years as the cut-off age to distinguish adult from juvenile offenders; however, because too few authors provided their cut-off for defining the sample, we retained only the dichotomous description of the sample as provided by the original authors. Other demographic variables, such as ethnicity and socioeconomic status, were not provided by a sufficient number of reports for analysis.

**Sample size.** The size of each sample was recorded simply as the number of individuals making up the sample for which the mean IQ appeared. Some investigations were missing IQ data on some of their study participants. This typically occurred when the study set out to examine some characteristic (or characteristics) other than IQ but provided IQ data in order to describe the general characteristics of their sample. In such situations, the sample sizes appearing in Supplemental Tables 1 and 2 (available on the Web at http://dx.doi.org/10.1037/a0033-2909.131.4.xxx) represent the number of individuals included in the mean IQ score and not necessarily the size of the entire sample participating in that study. Samples for which the sample size could not be identified or derived were dropped. In one case, this occurred because the report contained a typographical error (Vera, Barnard, & Holzer, 1979), obscuring the actual sample size. In other cases, the authors simply did not provide the sizes of the samples or subsamples (Peters, 1976; Wormith, Bradford, Pawlak, Borzeki, & Zobar, 1988).

**Statistical Analyses**

The five aforementioned research questions were addressed primarily by analysis of covariance (ANCOVA), with multiple regression used for supplementary analyses. For each analysis, the dependent variable was sample IQ, and the independent variables represented the five classifications by sample characteristic: offense status, victim age group, familiarity of child victims, sex of child victims, and sample age group. In each ANCOVA, only the comparisons between the classifiable samples were of interest. In order to retain the contributions of the unclassifiable samples to the variance and covariance estimates, however, we included in each ANCOVA both the classified and the unclassifiable samples and evaluated the groups of interest with two-tailed planned comparisons.

All analyses included the same two covariates: the year of the publication of each report and the format of the IQ data in the report (i.e., mean IQ score vs. frequency distribution of IQ scores). We chose year of publication to control for what has been called the Flynn effect, that is, the demonstrated increase in population IQ scores over time (Flynn, 1984, 1987; Kanaya, Scullin, & Ceci, 2003), because the samples realanalyzed here span most of the 20th century. We chose the data format as a covariate to compensate for differences that might have been inadvertently introduced by estimating mean scores from frequency distributions.

All analyses were weighted by sample size. One report (Beattie, 1951) contributed several disproportionately large samples. In order to determine whether one investigation unduly influenced the final results, we repeated the data analyses, excluding that investigation. Doing so did not change which analyses were and were not statistically significant; therefore, that study was retained.

For both empirical and theoretical reasons, the juvenile samples \( n = 71 \) samples) could neither be meaningfully divided into the same subtypes as the adult samples \( n = 165 \) samples) nor collapsed together. The empirical reason is that authors describing juvenile offenders did not as frequently subdivide their samples. The literature contained no samples of juvenile sexual offenders against boys only, against intrafamilial children only, or against extrafamilial children only. The theoretical reason is that the subclassifications used with the adult sexual offender samples do not necessarily have the same meaning when applied to juvenile sexual offenders. For example, the victim of an adult intrafamilial offender is typically the offspring of the offender; the victim of a juvenile intrafamilial offender, however, is typically a sibling. Second, authors investigating juvenile sexual offenders referred to offenders with “peer aged” victims. If a 44-year-old sexual offender contacts a 14-year-old victim, the offender would easily be classified as an offender against children; when a 15-year-old sexual offender contacts a 14-year-old victim, however, it remains unclear whether to classify the offender as an offender against children (because the victim is technically a minor) or as the counterpart of an offender against adults (because sexual interest in a 14-year-old is generally considered age-appropriate for a 15-year-old). Thus, indiscriminately collapsing juvenile and adult sexual offenders could confound other, pre-existing group differences. Thus, the juvenile sexual offender samples received separate analyses from the adult sexual offender samples.

**Results**

**Associations Between Sample IQ and Potential Covariates**

Three potentially relevant covariates were analyzed for any association with sample IQ over and above the two covariates already included. Each comparison included the same two preselected covariates. ANCOVA of sample IQ revealed no significant group differences among samples that provided FSIQ directly \( n = 137 \) adult samples and 65 juvenile samples) and those for which VIQ \( n = 11 \) and 0, respectively), PIQ \( n = 10 \) and 3, respectively), or their mean \( n = 7 \) and 3, respectively) substituted for FSIQ, among either the adult samples, \( F(3, 159) = 1.54, p = .21 \), or the juvenile samples, \( F(2, 66) = 0.05, p = .95 \). ANCOVA of the juvenile samples has one fewer numerator degrees of freedom because no juvenile samples used VIQ to substitute for FSIQ, leaving an empty cell.) ANCOVA similarly revealed no significant group difference between published and unpublished reports of either adult samples, \( F(1, 161) = 2.86, p = .09 \), or juvenile samples, \( F(1, 67) = 0.01, p = .94 \). Testing for differences in IQ between samples that received different IQ tests was complicated by some samples having received more than one test. ANCOVA of only those samples that received a single test \( n = 147 \) adult samples and 40 juvenile samples) revealed no overall group differences among either the adult samples, \( F(8, 136) = 1.00, p = .48 \), or the juvenile samples, \( F(6, 31) = 2.25, p = .06 \). The fewer numerator degrees of freedom for the juvenile samples’ analysis reflect that fewer IQ tests were used with the juvenile samples than with the adult samples.

**Adult Offenders**

Figure 1 shows the means for the 165 adult samples, classified by their offense type or status as nonoffenders and adjusted for the two covariates. ANCOVA indicated significant omnibus differences in IQ among the three groups, \( F(2, 160) = 15.29, \text{partial } \eta^2 = .16, p < .0001 \). The single planned comparison confirmed specifically that the sexual offender samples scored significantly lower in IQ than the nonsexual offender samples, \( F(1, 160) = 23.85, \text{partial } \eta^2 = .13, p < .0001 \). Thus, considered in its entirety, the literature supports the claim that men who commit sexual offenses score lower in IQ than men who commit nonsexual offenses.
Figure 2 shows the adjusted means for the same 165 samples, this time subclassifying the samples of sexual offenders into sexual offenders against children and sexual offenders against adults. ANCOVA indicated significant omnibus group differences in IQ among the five groups, $F(4, 158) = 7.74$, partial $\eta^2 = .16$, $p < .0001$, but only the sexual offenders against children scored significantly below the nonsexual offenders, $F(1, 158) = 18.39$, partial $\eta^2 = .10$, $p < .0001$. The sexual offenders against adults scored intermediately between the sexual offenders against children and the nonsexual offenders, but not significantly differently from either, $F(1, 158) = 1.80$, partial $\eta^2 = .01$, $p = .18$, and $F(1, 158) = 0.13$, partial $\eta^2 = .00$, $p = .72$, respectively. These results provide some support for the hypothesis that sexual offenders’ IQs relate to the age of the victims they target.

Sufficient information existed in this literature to permit an additional approach to confirm that adult sexual offenders’ IQs relate to the ages of their victims. As already indicated, prior investigators varied in how young a victim had to be before they would classify the offender specifically as an offender against children. If there does exist an association between offenders’ IQs and their victims’ ages, then one would hypothesize that the samples with the youngest victims would have the lowest mean IQs. Thus, we regressed IQ onto the age cut-off for those studies for which that cut-off was known ($n = 31$ samples), forcing the two covariates into the regression equation. This revealed a highly significant association between sample IQ and the victim age cut-off applied in forming the sample, $t(27) = 5.17$, $\beta = 0.64$, $p <$
The unstandardized regression coefficient, B = 2.21, indicated that for each decrease of 1 year in the definition of child, the samples’ mean IQ dropped by approximately 2 IQ points; a sample composed of sexual offenders against children ages 12 and under should be expected to show a mean IQ more than 10 points lower than a sample composed of sexual offenders against children ages 17 and under. This result suggests that the contradictory conclusions of the literature appear to be related not only to whether the victims were children, but also to how young the children were.

To determine whether the sexual offenders against intrafamilial children differed in IQ from sexual offenders against extrafamilial children, the samples of offenders against children were further subclassified, grouping them according to the familiality categories already described. ANCOVA of the sample IQ scores, categorized by familiality, indicated significant omnibus differences among the five groups, F(4, 158) = 7.28, partial η² = .16, p < .0001; however, the IQ of extrafamilial offender samples did not differ from that of the intrafamilial ones, F(1, 158) = 0.00, partial η² = .00, p = .97 (see Figure 4). Reclassifying the offenders against children, this time by the sex of child victims, also yielded an omnibus difference among the five groups, F(4, 158) = 6.79, partial η² = .15, p < .0001; however, no significant difference between sexual offenders against girls and sexual offenders against boys appeared, F(1, 158) = 0.06, partial η² = .00, p = .82 (see Figure 5).

As a confirmatory analytical technique, we conducted the following multiple regression, simultaneously entering all of the preceding independent variables and covariates. The reanalysis of the previously collected sample statistics, multiply regressing them onto independent variables of interest and covariates, constitutes a metaregression (Stanley & Jarrell, 1989). Although metaregressions typically examine effect size statistics as their dependent variables, the common unit of the present literature permits sample means, rather than differences between sample means, to be examined. Because one can reformulate ANCOVA designs as multiple regressions, the preceding analyses could also have been conducted as metaregressions. We chose to present them in ANCOVA format, however, to facilitate direct and intuitive comparison with existing reports in the literature. Unlike regression coefficients, the means of groups of samples retain IQ-point units.

For this analysis, the variables representing categorical or dichotomous values were expressed with dummy coding and effects coding (Cohen & Cohen, 1983, pp. 183–204), as appropriate. One set of four dummy-coded variables reflected whether a sample represented nonoffenders, sexual offenders against adults, sexual offenders against children, or sexual offenders against victims of mixed or unknown ages. As in the ANCOVA-based analyses, nonsexual offenders were chosen as the reference category because they share with sexual offenders the histories of having committed, having been apprehended for, and having been convicted of a crime. A value of +1 in one of the aforementioned dummy variables indicated that the sample was of the type indicated by that variable, and a value of 0 in all four dummy variables indicated that the sample consisted of nonsexual offenders. Thus, after regressing IQ onto this set of variables, a significant contribution by a dummy variable would indicate that the group represented by that variable differed significantly from nonsexual offenders.

The familiality and the sex of child victims were both effects coded. Specifically, samples composed of sexual offenders against extrafamilial children received a code of +1, and samples composed of sexual offenders against intrafamilial children received a code of −1, whereas all other samples received a code of 0 on this variable. Thus, after regressing IQ onto this variable, a significant regression coefficient would indicate a significant difference in IQ between these two subtypes of sexual offenders against children, over and above any other characteristics already in the regression equation (and minimizing any effects associated with samples coded 0; Cohen & Cohen, 1983). Furthermore, with these codes, a positive regression coefficient would indicate that samples of sexual offenders against extrafamilial children have higher IQs than samples of sexual offenders against intrafamilial children. The analogous strategy coded for sex, with the values of +1 indicating samples of sexual offenders against girls, −1 indicating samples of sexual offenders against boys, and 0 indicating all other samples. Thus, a positive regression coefficient would indicate that sexual offenders against girls had higher IQ scores than sexual offenders against boys.

The metaregression results (see Table 1) confirmed the prior analyses. Samples composed of men who committed sexual offenses against children scored significantly lower in IQ than samples composed of men who committed only nonsexual offenses. Although samples of sexual offenders against adults scored approximately 4.5 IQ points lower than samples of nonsexual offenders, the effect was not statistically significant.

![Figure 3. Mean IQ scores of samples of adult sexual offenders against children by the age cut-off used to define victims as child victims (n = 31 samples). Age cut-offs were recorded as exclusive ranges. Thus a sample made up of offenders against children under 15 received a value of 15, whereas a sample made up of offenders against children aged 15 or younger received a value of 16. See text for results of partial correlation analysis.](image)
IQ did not relate significantly to the effects code representing the comparison of offenders against boys with offenders against girls. Nor did it relate to the code representing the comparison of offenders against intrafamilial children with offenders against extrafamilial children.

Juvenile Offenders

Analysis of the juvenile samples proceeded as follows: The juvenile samples were classified into sexual offenders and nonsexual offenders, with an additional code again used to capture the remaining juvenile samples (i.e., nonoffenders), to allow retention of their variance and covariance in the analyses as before. This formed a 2 × 3 ANCOVA with sample age group (adult or juvenile) and sample offense status (sexual offender, nonsexual offender, or nonoffender) as factors. The same two covariates, year of report and data format in each report, were included in each analysis. Because only the comparisons involving the offender samples were of interest, only the comparisons involving them were planned. The omnibus test statistics are provided here for completeness. ANCOVA indicated significant main effects by sample age group, F(1, 228) = 10.74, partial $\eta^2 = .05, p = .001$, and by sample offense status, $F(2, 228) = 4.75$, partial $\eta^2 = .04, p = .01$, but no significant interaction between them, $F(2, 228) = 0.69$, partial $\eta^2 = .01, p = .50$ (see Figure 6).

The juvenile samples did not follow the same pattern as did the adult samples; the juvenile sexual offender samples did not differ significantly from the juvenile nonsexual offender samples, $F(1, 228) = 1.34$, partial $\eta^2 = .01, p = .25$. Moreover, both the juvenile sexual offenders and the juvenile nonsexual offenders scored significantly lower in IQ than their adult counterparts, $F(1, 228) = 13.67$, partial $\eta^2 = .06, p < .0001$, and $F(1, 228) = 27.04$, partial $\eta^2 = .11, p < .0001$, respectively.

Subclassifying the juvenile sexual offenders by the ages of their victims also revealed no omnibus group differences, $F(4, 64) = 1.84$, partial $\eta^2 = .10, p = .13$, and neither the juvenile sexual offenders against adults ($n = 5$ samples, adjusted mean IQ = 82.4) nor the juvenile offenders against children ($n = 18$ samples, adjusted mean IQ = 93.8) differed significantly from the juvenile samples of nonsexual offenders ($n = 28$ samples, adjusted mean IQ = 91.5), $F(1, 64) = 2.84$, partial $\eta^2 = .04, p = .10$, and $F(1, 64) = 0.96$, partial $\eta^2 = .02, p = .33$, respectively. It is difficult to ascertain at this point whether this reflects (a) that juvenile sexual offender IQs differ in their associations with their offense characteristics or (b) that the literature provided too few samples homogeneous on these characteristics for reliable analyses.

Unfortunately, as already mentioned, the literature did not provide sufficient detail regarding the juvenile sexual offenders to permit a full ANCOVA.
against children to permit their subclassification by the familiality or sex of their child victims.

Discussion

These reanalyses confirm, first, that adult males who commit sexual offenses score lower in IQ than do adult males who commit nonsexual offenses (see Figure 1). This difference cannot be attributed easily to disparities in the IQ tests administered or to whether a report was published—neither characteristic related significantly to IQ scores. Because the year of each report and the format of the data in each report (i.e., sample means vs. frequency distributions) were entered as covariates in all analyses, the IQ differences cannot be attributed to these factors either. Moreover, these results also confirm that for adult men, IQ differences between sexual and nonsexual offenders do not occur uniformly across sexual offender subtypes—the younger the victim age cutoff used to define a sample of offenders against children, the lower the sample’s mean IQ (see Figure 3).

The failure of sexual offenders against adults to differ significantly from nonsexual offenders (see Figure 2) may reflect one or more of several possible situations. (a) The null hypothesis is true; sexual offenders against adults have IQs no different from nonsexual offenders, and the only sexual offender characteristic that relates to IQ is the propensity to offend against children. (b) Offenders against adults actually do score below nonsexual of-

Table 1
Simultaneous Metaregression of IQ Onto Sexual Offense Characteristics of Adult Samples (N = 165)

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE</th>
<th>β</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariates</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year of publication</td>
<td>0.04</td>
<td>0.03</td>
<td>0.107</td>
<td>0.24</td>
</tr>
<tr>
<td>Frequency-distributed data</td>
<td>-3.88</td>
<td>0.85</td>
<td>-0.336</td>
<td>0.0001</td>
</tr>
<tr>
<td>Dummy-coded sample types</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonoffender</td>
<td>3.18</td>
<td>2.99</td>
<td>0.081</td>
<td>0.29</td>
</tr>
<tr>
<td>Sexual offender against children</td>
<td>-6.12</td>
<td>1.46</td>
<td>-0.396</td>
<td>0.0005</td>
</tr>
<tr>
<td>Sexual offender against adults</td>
<td>-4.48</td>
<td>3.30</td>
<td>-0.096</td>
<td>0.18</td>
</tr>
<tr>
<td>Sexual offender, victim ages mixed or unknown</td>
<td>-4.44</td>
<td>1.34</td>
<td>-0.243</td>
<td>0.001</td>
</tr>
<tr>
<td>Effects-coded factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Familiality of child victim(s)</td>
<td>-1.42</td>
<td>1.89</td>
<td>0.059</td>
<td>0.46</td>
</tr>
<tr>
<td>Sex of child victim(s)</td>
<td>-0.55</td>
<td>1.65</td>
<td>-0.024</td>
<td>0.74</td>
</tr>
</tbody>
</table>

Note. $R^2 = .236$. The 165 samples in the analysis represent 22,301 individual cases. The four dummy-coded variables indicate differences relative to nonsexual offenders, which received a value of 0 on all four dummy variables. For the effects coding of familiality, negative regression coefficients indicate extrafamilial offenders against children scoring more poorly in IQ than intrafamilial offenders against children. For the effects coding of sex, negative regression coefficients indicate sexual offenders against female children scoring more poorly in IQ than sexual offenders against male children.
fenders, but the small number of observations provided insufficient power for differences to reach statistical significance. (c) Sexual offenders against adults (like sexual offenders against children) consist of subtypes, only some of which show IQs lower than those of nonsexual offenders, and the heterogeneity within the samples obscured underlying differences. Also worth noting is that because researchers varied in their definitions of what constituted an adult victim of a sexual offense, some samples that were categorized as offenders against adults actually included an unknown proportion of individuals who might reasonably be categorized as offenders against children. This would serve to decrease the mean IQ of the samples of offenders against adults. Unfortunately, too few reports provided their research definitions of adult victims to permit an evaluation or control of this potential effect.

Taken together, these results suggest that IQ relates primarily to the presence of pedophilia among sexual offenders and that differences in samples’ scores reflect differing proportions of genuinely pedophilic offenders occurring within those samples. As previously stated, investigations using phalometric assessment techniques have confirmed that sexual offenders against children show greater penile responses to stimuli involving nude children than do sexual offenders against adults (e.g., Blanchard et al., 2001). Moreover, samples composed of men who offended against children ages 13 and younger would be reasonably expected to contain a larger proportion of genuinely pedophilic offenders than would samples composed of men who offended against children ages 18 and younger.

Subclassification of samples of offenders against children by the sex of their victims and by their familial relationship to their victims revealed no group differences in IQ. This result may be interpreted as those characteristics being unrelated to IQ, at least among the samples thus far reported. It bears emphasizing, however, that these analyses comprised relatively few samples; that investigators varied in their classification of step-children as intrafamilial or extrafamilial victims; and that the familial status and sex of victims are not, in practice, independent characteristics, because the great majority of intrafamilial victims are female and the literature contains no samples of offenders against intrafamilial males.

Noteworthy is that the mean IQ of the nonsexual offender samples (adjusted, weighted $M = 98.9$, raw $M = 98.6$) was essentially equal to the population mean of 100; one might expect these offenders to score more substantially below average. This may be because the least intelligent offenders—the sexual offenders—were separated from the other types, thus elevating the mean of the remaining (nonsexual) offenders. Because sexual offenders represent a substantial proportion of incarcerated men, their presence would reduce the mean IQ scores of randomly selected samples of the incarcerated.

No prior studies compared adult and juvenile sexual offenders with regard to IQ, and their comparison here yielded two unexpected results, namely, that juvenile offenders have lower IQs than adult offenders and that the IQ difference between sexual offenders and nonsexual offenders occurs only among the samples of adult offenders (see Figure 6). Because adult sexual offenders had frequently shown their sexually inappropriate behaviors during their own childhoods and because adult nonsexual offenders often begin their criminal careers as juveniles, one might expect that the adult and juvenile samples each represented the same populations, merely sampled at different points in their lives. The IQ differences between the adult samples and juvenile samples of each offender type, however, indicate that this is not the case. For example, less intelligent offenders (sexual or nonsexual) might come to the attention of legal and medical authorities at younger ages because they are unable to elude capture and conviction for as long as can more intelligent offenders.

Also unexpected was the finding that the juvenile sexual offender samples differed only nonsignificantly from the juvenile nonsexual offender samples. Because the direction of the group difference that was present was in the expected direction (and analogous to that between the adult offender groups), the failure to obtain statistical significance may be, in part, attributable to the fewer observations available or to a floor effect imposed by the juvenile samples’ lower IQ scores overall.

In sum, the present reanalyses indicate that the inconsistencies among the previous investigations are due in significant part to differences in the characteristics of their samples; the present data indicated that these are the sexual offenders’ age group and their victims’ age group. Although we explored or controlled as many between-studies factors as the literature permitted, it would be worthwhile to verify these conclusions within a single study.
Taken together, the present results suggest the field has accumulated sufficient justification to move beyond investigating whether sexual offenders, particularly those with pedophilia, show low IQs and increase attention to why they do. One explanation is that sexual offenders do not, in fact, have lower IQs than other types of offenders, but only appear to because of an ascertainment bias, that is, because less intelligent sexual offenders more frequently become apprehended and less frequently have the financial resources to assist in acquittal. This explanation would imply that ascertainment bias does not apply to nonsexual offenders (whose IQs, in this study, approximated the population average). There is, however, no compelling reason to theorize that the chances of apprehension for a man who molests a stranger’s child are influenced by his IQ, but that the chances of apprehension for a man who breaks into a stranger’s home, for example, are not.

A simpler explanation is that the IQ differences are genuine and reflect an underlying deficiency of brain function. This situation could be explained by two different causal associations, one direct and one indirect. Authors arguing for the direct association (e.g., Galski, Thornton, & Shumsky, 1990; Stone & Thompson, 2001) maintain that poor cognitive functioning reflects disinhibited decision making or a failure to comprehend consequences, yielding sexual offending behavior. This theory, however, fails to account for the pattern of results demonstrated here: Dysexecutive and disinhibition theories would predict all types of offenders to manifest low IQs, whereas our most fine-grained analysis of victim age and offender IQ (see Figure 3) suggests that low IQ is a characteristic particular to pedophilia. Finally, the direct theory is difficult to reconcile with the IQs of the sexual offender samples falling in the normal range of IQ (albeit significantly lower than comparison samples) and with the relatively rare occurrence of pedophilia.

In contrast, the present data are consistent with the notion we have previously offered of an indirect association between IQ and pedophilia (i.e., Blanchard et al., 2002; Cantor et al., 2004). Specifically, we suggest that a third variable—a perturbation of prenatal or childhood brain development—produces both pedophilia and low IQ. That is, low IQ correlates significantly with pedophilia, not because it causes pedophilia, but because both characteristics result from some common etiological factor.

Throughout the 20th century, investigators have strongly favored nonbiological over biological theories in the development of pedophilia. When considered with the other reliable correlates of pedophilia, however, the present data suggest a biological influence. Lower mean IQs, greater frequencies of childhood head injuries (Blanchard et al., 2002, 2003), and elevated levels of non-right-handedness (Bogaert, 2001; Cantor et al., 2004, 2005) can be interpreted relatively easily as reflections of a perturbation of brain development. Although these correlates do not rule out psychosocial influences, they do suggest that psychosocial theories are incomplete. In addition to their impact for etiological theories, the present data have implications for the classification schemes used in research. Despite that many prior studies treated sexual offenders as a single population, pedophilic and nonpedophilic sexual offenders represent distinct categories; a common etiological pathway should not be assumed, and contemporary investigators who indiscriminately collapse them together risk obscuring underlying patterns.

Future research will hopefully attempt to ascertain whether sexual offenders demonstrate general or more selective cognitive differences, using tests associated with more specific cognitive functions. Some such attempts have already been reported but have not reliably identified any specific deficits in executive functioning, memory, motor, or achievement tests (for a review, see Blanchard et al., in press). Unfortunately, most of these attempts used small and highly heterogeneous samples of sexual offenders. The present results suggest that examination of heterogeneous samples can offer little. The present results furthermore indicate that samples must be, in general, much larger than those typically used. In the present analyses, samples of adult sexual and nonsexual offenders differed by approximately 5 IQ points, equivalent to an effect size of 0.33 (or moderate to small; Cohen, 1988). To detect differences of this magnitude with a power of 0.80 requires experimental and control samples with 141 subjects each. Of the 100 samples of adult sexual offenders reanalyzed here, only 9 contained that many. If we are to learn any methodological lesson from the IQ literature on sex offenders, it is to attend to these deficiencies before another 8 decades pass.

References

* References marked with an asterisk indicate studies included in the meta-analysis.


*Benoiit, J. L. (1990). Incarcerated adolescent child molesters, violent offenders, and nonviolent offenders: An analysis of abuse history, per-


Glueck, B. C., Jr. (1955). Final report: Research project for the study and


*Hambridge, J. A. (1994). Pedophiles’ ratings of adult and child photo-


Henn, F. A., Herjanic, M., & Vanderpearl, R. H. (1976). Forensic psychi-


*Jersild, J. (1967).


*Lei, L. K., & Starzynski, J. (1990). Adolescent rapists and adolescent


*Hocker, S., Langevin, R., Wortzman, G., Bain, J., Handy, L., Chambers, J.,

*Hocker, S., Langevin, R., Wortzman, G., Dickey, R., Bain, J., Handy, L.,

*James, L. C. (1987). MMPI profiles, I.Q. scores, and demographic data


Robinson, W. S. (1950). Ecological correlations and the behavior of


Rader, C. M. (1977). MMPI profile types of exposers, rapists, and assault-


Army General Classification Test


Rader, C. M. (1977). MMPI profile types of exposer, rapists, and assault-


*Yeudall, L. T. (1977). Neuropsychological assessment of forensic disord-


Received February 16, 2004

Revision received November 15, 2004

Accepted January 27, 2005