The Specificity of Victim Count as a Diagnostic Indicator of Pedohebephilia [Letter to the

Editor]

Ray Blanchard

The original publication is available at

http://www.springerlink.com/content/c260w4213u70055h/

Corresponding author:

R. Blanchard
Law and Mental Health Program
Centre for Addiction and Mental Health
250 College Street, Toronto, Ontario M5T 1R8, Canada
e-mail: ray_blanchard@camh.net

R. Blanchard
Law and Mental Health Program,
Centre for Addiction and Mental Health,
Department of Psychiatry, University of Toronto,
Toronto, Ontario, Canada

In his recent letter to the editor of this journal, *DSM-5 proposals for paraphilias: Suggestions for reducing false positives related to use of behavioral manifestations*, First (in press) offers several criticisms and suggestions regarding the proposed diagnostic criteria for the paraphilic disorders in the upcoming fifth edition of the American Psychiatric Association's *Diagnostic and statistical manual of mental disorders* (DSM-5; American Psychiatric Association, 2010). I am writing in response to that letter. Although I am (unavoidably) writing as a member of the Paraphilias Subworkgroup of the DSM-5 Work Group for Sexual and Gender Identity Disorders, which authored the proposed criteria, I am writing to express my own conclusions and not the considered consensus of the Subworkgroup. Any errors or omissions in this reply are mine and mine alone.

The single most important point in First's letter, as I understand it, is this: An uncooperative patient cannot be diagnosed with a specific paraphilia on the sole grounds that he has offended against *x* or more victims in a way symptomatic of that paraphilia, because such a practice would produce an excessive number of false positive diagnoses. First's view is related to the notion that behavior is a more fallible indicator of paraphilia than is self-report. First further asserts that a diagnostic criterion for uncooperative subjects based on a minimum number of individual victims would produce excessive false negatives as well as false positives, because it neglects the frequency of sexual interactions with a given victim and the time period over which repeated interactions with that victim occurred.

First, as one would expect, freely acknowledges that a patient's self-report may be unreliable in the adversarial context of a forensic evaluation. His response to that situation is as follows: In the absence of reliable self-report information, a diagnosis of paraphilia is possible only if two conditions are met. Firstly, there must be a *pattern* of paraphilic arousal. The term

pattern is not defined, but it is exemplified by a large victim count plus an absence of normal sexual behavior. Secondly, other possible causes for the criminal sexual behavior must be ruled out. He includes a specific suggestion for the paraphilic disorders' diagnostic criteria to embody his second point: In order to reduce the rates of false positive diagnoses of the various paraphilias, an exclusionary criterion should be added to the criteria set of each of the paraphilias. This exclusionary criterion would list the other DSM diagnoses that should be ruled out as alternative explanations of the patient's symptoms.

First writes, in places, as if the unacceptable false positive rate that would result from diagnosing paraphilias by victim count were self-evident or already known as fact. In other places, however, he seems to entertain the possibility of treating this as an empirical question. Thus, he notes that no published study has established the minimum number of similar victims that would justify the diagnosis of a specific paraphilia in a patient who verbally denies having that paraphilia. His further statements on this head spell out not only the research design but also the precise statistical analysis that might be used to estimate the specificity and sensitivity of victim count as a diagnostic indicator of pedohebephilia. His remarks may be summarized as follows: Receiver operating characteristic (ROC) curves are typically calculated to estimate the proportions of false positive and false negative diagnoses that would result from applying a given threshold (cutting score) to a diagnostic indicator. An ROC analysis could be conducted to estimate the specificity and sensitivity of victim count as a diagnostic indicator of pedohebephilia, using a total of three or more child victims as the cutting score, and treating men who verbally admit to an erotic preference for children as the gold standard for a diagnosis of pedohebephilia.

In what follows, I will address First's various comments and criticisms, beginning with an empirical ROC analysis along the lines he suggests. Before I begin that, however, I must clear one historical matter out of the way.

First devotes a substantial proportion of his letter to explaining the vicissitudes of *behavior* in the diagnostic criteria for the paraphilias in DSM-IV (American Psychiatric Association, 1994) and DSM-IV-TR (American Psychiatric Association, 2000). The prominence of this account might leave the reader with the impression that the current proposal to use repeated sexual offenses of the same nature as one basis for diagnosing paraphilic disorders in the DSM-5 is somehow an outgrowth or consequence of the varying uses of *behavior* in DSM-IV and DSM-IV-TR. That is, from my standpoint, not the case at all.

I suggested, in my first report to the Paraphilias Subworkgroup (Blanchard, 2010a), that a patient be diagnosed as pedohebephilic if he has offended against three or more children under the age of 15, in accordance with a long tradition of clinical research and practice, whereby men who have committed sexual crimes of a similar nature against multiple victims are diagnosed with the corresponding paraphilia, whether they acknowledge that disorder or not. This tradition might not be evident in DSM-IV or DSM-IV-TR, but its influence can clearly be seen in DSM-III (American Psychiatric Association, 1980). The DSM-III diagnostic criteria for Exhibitionism (302.40), for example, read as follows: "Repetitive acts of exposing the genitals to an unsuspecting stranger for the purpose of achieving sexual excitement, with no attempt at further sexual activity with the stranger" (p. 272). There are various clinical research studies in which men are classified into paraphilic groups on the basis of their known sexual offenses (e.g., Blanchard et al., 1999; Cantor et al., 2005; Freund, 1965).

Ad Hoc Analysis—The Specificity of Victim Count

Because First's central concern is false positive diagnoses of pedohebephilia, the most important methodological aspect of an empirical investigation is assembling a gold standard group of teleiophiles (men with an erotic preference for persons between the ages of physical maturity and physical decline). This can not, in reality, be perfectly achieved. A teleiophilic preference is the desirable preference, in the clinic, in the courts, and in general society; and many men will falsify their self-report and manipulate their phallometric test results,[1] if they can, to obtain that diagnosis.

The closest thing to a gold standard group of teleiophiles would be a group of men whose known sexual offenses, self-report, and phallometric test results all indicated teleiophilia. In that case, however, there would be no diagnostic indicator left to test; one would have used up all one's degrees of freedom, so to speak, by employing all possible indicators for the purpose of group selection. For the following study, therefore, I selected a "silver standard" group of teleiophiles on the basis of self-report and phallometric test results. The primary limitation of such a silver standard group is that some of the men "falsely" diagnosed as pedohebephilic according to their victim count might, in fact, have been correctly diagnosed. Thus, the obtained estimate of specificity should be regarded as a minimum estimate.

Method

Subjects

Between November 1995 and October 2009, 3,166 male patients were administered the same phallometric test for erotic object (gender and age) preferences at the Kurt Freund

---

[1] Phallometric testing is a psychophysiological technique for assessing erotic interests in males. In phallometric tests for erotic gender- and age-orientation, the individual's penile blood volume is monitored while he is presented with a standardized set of laboratory stimuli depicting male and female children, pubescents, and adults. Increases in the subject's penile blood volume (i.e., degrees of penile erection) are taken as an index of his relative attraction to different classes of persons.

Laboratory of the Centre for Addiction and Mental Health (Toronto, Ontario, Canada). The sources of the clinical referrals included parole and probation officers, prisons, defense lawyers, various institutions (ranging from group homes for mentally retarded persons to regulatory bodies for health or educational professionals), and physicians in private practice. As would be expected from the preponderance of criminal justice sources, the majority of patients had one or more sexual offenses against children, adults, or both. Men who had no involvement with the criminal justice system and who initiated referrals through their physicians included patients who were unsure about their sexual orientation, patients concerned about hypersexuality or "sex addiction," patients experiencing difficulties because of their excessive use of telephone sex lines or massage parlors, clinically obsessional patients with intrusive thoughts about unacceptable sexual behavior, and patients with paraphilic behaviors like masochism, fetishism, and transvestism. Subsets of these patients have been analyzed in two previous studies (Blanchard, 2010b; Blanchard et al., in press).

There were 2,725 patients who had given informed consent for their assessment data to be used for research purposes, and whose sexual history data were complete and cross-checked at the time these data were retrieved (Blanchard et al., in press). Self-report data on sexual attractions were available for 2,715 of these. From this pool, I selected 998 subjects for the present study, using criteria explained later. The selected men had a mean age of 34.93 years (*SD* = 11.98) and a median education of high school graduation.

Materials and Measures

Sexual History

A standardized form, described in detail by Blanchard et al. (2009), was used to record the patient's history of sexual offenses. Most of that information came from objective documents

that accompanied his referral, for example, reports from probation and parole officers. The offense-history data were cross-checked against, and supplemented by, other information provided by the patient himself, including the number and nature of any additional sexual offenses that were admitted by the patient but for which he was never charged. For the present study, the total number of different children under the age of 15 with whom the patient had interacted sexually was called the *victim count*. It did not matter, for the purpose of computing this variable, if the child was male or female, consenting or coerced, related or unrelated to the patient, approached physically (e.g., fondled) or non-physically (e.g., exposed to), approached once or approached on multiple occasions.

The recorded sexual offenses included those related to the possession or (rarely) manufacture of child pornography. Detailed information on the number of images involved was often not available. This variable was therefore recorded dichotomously, as present or not present. For purposes of this study, child pornography was scored as present if the patient had charges of this nature or if he admitted to the use of child pornography (or both).

The patient was also asked to rate his sexual attraction to persons in 12 gender–age categories (e.g., females aged 17 years or older, males aged 17 years or older, females aged 15–16 years, males aged 15–16 years, and so on) using a five-point scale. The patient's information was solicited by the laboratory manager in a structured sexual history interview, which the manager conducted the same day he administered the phallometric test.

Phallometric Measurement

The Kurt Freund Laboratory is equipped for volumetric phallometry, that is, the apparatus measures penile blood volume change rather than penile circumference change. The volumetric method measures penile tumescence more accurately at low levels of response

(Kuban, Barbaree, & Blanchard, 1999). A photograph and schematic drawing of the volumetric

apparatus are given in Freund, Sedlacek, and Knob (1965). The major components include a

glass cylinder that fits over the penis and an inflatable cuff that surrounds the base of the penis

and isolates the air inside the cylinder from the outside atmosphere. A rubber tube attached to the

cylinder leads to a pressure transducer, which converts air pressure changes into voltage output

changes. Increases in penile volume compress the air inside the cylinder and thus produce an

output signal from the transducer. The apparatus is calibrated so that known quantities of volume

displacement in the cylinder correspond to known changes in transducer voltage output.

The specific test used in this study has been described in detail by Blanchard, Klassen,

Dickey, Kuban, and Blak (2001). The test stimuli were audiotaped narratives presented through

headphones and accompanied by slides. There were seven categories of narratives, which

described sexual interactions with prepubescent girls, pubescent girls, adult women,

prepubescent boys, pubescent boys, and adult men, and also solitary, nonsexual activities

("neutral" stimuli). All narratives were written in the second person and present tense and were

approximately 100 words long. The narratives depicted fantasy situations in which sexual

interaction with children would be relatively plausible or sexual interaction with adults would be

relatively plausible as well as somatic and social attributes indicating the physical maturity of the

imaginary target. Samples narratives are given in Blanchard et al. (2007). The narratives

describing heterosexual interactions were recorded with a woman's voice, and those describing

homosexual interactions, with a man's. Neutral stimuli were recorded with both.

Each test trial consisted of one narrative, accompanied by photographic slides on the

three adjacent projection screens, which simultaneously showed the full-length front view, full-

length rear view, and close-up genital region of a nude model who corresponded in age and

gender to the topic of the narrative. Figure 1 in Blanchard et al. (2007) illustrates the standard pose used for the full-length front views. Each trial included three nude models, each presented for 18 sec. Therefore the total duration of a trial was 54 sec, during which the examinee viewed a total of nine slides, three at a time. Neutral narratives were similarly accompanied by slides of landscapes.

The full test consisted of four blocks of seven trials, with each block including one trial of each type in fixed, pseudorandom order. Although the trial length was fixed, the intertrial interval was variable, lasting as long as necessary for penile blood volume to return to baseline. The time required to complete the test was usually about 1 hour.

Two kinds of scores, representing the amount of responding and the direction of responding, were calculated from the raw phallometric data. The amount of responding was quantified with a standard measure in the Kurt Freund Laboratory, the *Output Index* or *OI* (Freund, 1967). This is the average of the three greatest responses to any stimulus category except "neutral," where penile response is expressed in cubic centimeters (cc) of blood volume increase from the start of a trial.

The assessment of response direction in the Kurt Freund Laboratory involves several steps. This process, whose calculations and rationale have been presented at length by Blanchard et al. (2001), may be briefly explained as follows. During the stimulus trials, penile blood volume change is sampled four times per second and recorded as a curve of blood volume change over time. The examinee's response during a given trial is measured in two ways: (a) as the maximum deflection of the curve (i.e., the greatest departure from initial value occurring during the 54 sec of the trial), and (b) as the area under the curve. Each examinee's 28 deflection scores are converted into standard scores, based on his own deflection data (in other words, they

are converted into ipsative *z* scores), and the same operation is carried out on his area scores. Next, for each examinee, the standardized deflection and standardized area scores are averaged to yield a separate composite score for each of the 28 trials. Finally, the data are reduced to seven scores for each examinee by averaging his four composite scores in each of the seven stimulus categories. These seven *category scores* are taken as measures of the examinee's relative erotic interest in adult women, pubescent girls, prepubescent girls, and so on.

Final Gating Criteria and Assignment to Groups

Subjects were considered eligible for this study if their OI's were greater than or equal to 2.50 cc. This is slightly higher than my laboratory's customary threshold of 1.00 cc for considering a phallometric test result clinically interpretable (e.g., Blanchard et al., 2001), but it is consistent with the higher response criteria used in most other phallometric laboratories (see review in Lykins et al., 2010). An increase in penile blood volume of 2.50 cc in the Kurt Freund Laboratory corresponds roughly to 10% of a full erection for the average adult male. I chose the higher value because recent research has suggested that the reliability of phallometric diagnosis is significantly related to the magnitude of the subject's OI (Lykins et al., 2010).

A subject was selected for the Teleiophilic Group if he met both of the following criteria: (a) phallometric test results showed that the higher of his two category scores for adults (adult women and adult men) was over 0.25 *z* score units greater than the highest of his four category scores for children (prepubescent girls, pubescent girls, prepubescent boys, and pubescent boys), and (b) he verbally reported, in his structured sexual history interview, that his strongest sexual attractions were to women or men age 17 or older, and that he felt zero sexual attraction to girls or boys under the age of 15.

A subject was selected for the Pedohebephilic Group if he met essentially the obverse criteria: (a) the highest of his four category scores for children (prepubescent girls, pubescent girls, prepubescent boys, and pubescent boys) was over 0.25 z score units greater than the higher of his two category scores for adults (adult women and adult men),[2] and (b) he verbally reported that his strongest sexual attractions were not to women or men age 17 or older, and that he felt more than zero sexual attraction to girls or boys under the age of 15.

Using these criteria, 817 men were selected for the Teleiophilic Group, and 181 men were selected for the Pedohebephilic Group.

Results

Table 1 shows, for each group, the number of subjects who sexually offended against 0, 1, 2, 3, 4, 5, and 6 or more children under the age of 15 years. This table can be used to estimate the sensitivity and specificity of victim count as a diagnostic indicator of pedohebephilia, using different numbers of victims as the cutting score. I will use, as an illustrative example, a cutting score of 3 victims: Every man with fewer than 3 victims is classified as a teleiophile, and every man with 3 or more victims is classified as a pedohebephile. With that cutting score, the specificity of victim count is $(335 + 289 + 121)/817 = 91\%$, and the sensitivity is $(25 + 10 + 9 + 53)/181 = 54\%$.

Insert Table 1 about here

---

[2] This rule has been used for the clinical diagnosis of pedohebephilia in the Kurt Freund Laboratory for over 20 years (see Blanchard et al., 2001).

The trade-off between sensitivity and specificity may be graphically represented in a receiver operating characteristic (ROC) curve. Figure 1 shows the ROC curve for the data presented in Table 1.

Insert Figure 1 about here

One measure of the adequacy of a diagnostic test is the area under the ROC curve (AUC). There is no universal agreement on what AUC values indicate a poor, fair, good, or excellent clinical test—indeed, such labeling depends partly on context—but a value around .70 might be considered "fair," and a value around .80, "good." In the present case, the AUC was .72, 95% CI [.67, .77].

There is a problem with the foregoing analysis (apart from the obvious problem that one or more of the men in the Teleiophilic Group who had offenses against children might truly have been pedohebephiles rather than "false positives," even though they claimed an erotic preference for adults and produced phallometric results consistent with that claim). The more subtle problem is that the expectation that *none* of the men in the Teleiophilic Group should have offenses against children is not matched by an equal and opposite expectation that *all* of the men in the Pedohebephilic Group should have offenses against children. A man with no criminal charges of any kind could have ended up in the Pedohebephilic Group because he came forward and requested clinical help (or because concerned care-givers requested help on his behalf), and a man with no real-life child victims could have ended up in that group via criminal charges for child pornography. Such cases would have the effect of lowering the computed sensitivity and the AUC of victim count as a diagnostic indicator.

I conducted two further analyses to demonstrate the problem in regarding the Pedohebephilic Group as a gold standard. Both analyses used the easily manipulated variable of child pornography. In the first analysis, I simply excluded from the whole sample any man who had been charged with, or admitted to, the use of child pornography. This reduced the Teleiophilic Group to 772 subjects, and the Pedohebephilic Group to 106 subjects.

Table 2 shows, for each group, the number of subjects who sexually offended against different numbers of children under the age of 15 years. For the same cutting score of 3 victims, the sensitivity increased to 66%. The specificity remained virtually unchanged, at 91%.

Insert Table 2 about here

The manipulation also produced the expected effect on the ROC curve (Figure 2). The AUC increased to .79, 95% CI [.73, .85].

Insert Figure 2 about here

My second method for demonstrating the effect of non-molesting routes into the Pedohebephilic Group was to treat child pornography (for diagnostic purposes) as the equivalent of child molestation. Seto, Cantor, and Blanchard (2006) found that child pornography, treated as a dichotomous variable, had a diagnostic value equal to real-life offenses against multiple victims. I therefore created a new variable, the *victim equivalency count*, by adding 2 "victims" to the total score for any man who had been charged with, or admitted to, the use of child pornography. In this analysis, I returned to the use of the full sample.

Table 3 shows, for each group, the number of subjects with different scores on the victim equivalency count. I used a cutting score of 3 victims/victim-equivalents. Remarkably, the estimates for sensitivity and specificity were almost the same as those in the previous analysis: 66% and 89%, respectively.

Insert Table 3 about here

The ROC curve was affected somewhat more (Figure 3). The AUC increased to .83, 95% CI [.80, .87].

Insert Figure 3 about here

Discussion

First's helpful suggestion that the specificity of victim count—more specifically, the threshold score of three individual victims—should be demonstrated rather than assumed is addressed by the foregoing study. The specificity appears to be around 90%. For reasons already explained, this should be regarded as a minimum estimate of specificity. Although this study was limited to the diagnosis of pedohebephilia—the disorder for which I had the greatest quantity and highest quality of data—there is no obvious reason to assume that a cutting score in the neighborhood of three victims would produce markedly different false positive rates for other paraphilic disorders.

Responses to First's Other Major Points

Patterns vs. Counts

First lays great emphasis on the notion that the diagnosis of paraphilia in uncooperative subjects must be based on a *pattern* of behaviors rather than a *count* of behaviors. In fact, he uses the word *pattern* 16 times in his letter. One would expect, given the importance of this concept to him, that he would supply two or three examples of diagnosable patterns. In fact, he supplies one: "involvement of a large number of victims over time coupled with the absence of sexual behavior involving non-paraphilic arousal stimuli."

One would also expect that if First is going to advance only one example, he would put forward his best one. That makes the limitations of this example all the more striking. Half of this "pattern" is simply victim count, with the vague term *large* substituted for a suggested minimum number of individual victims. The other half of the "pattern" is based on self-report, whose veracity is precisely the problem with uncooperative subjects in the first place. This requires some translation to be seen clearly.

In plain language, the phrase "involvement of a large number of victims over time coupled with the absence of sexual behavior involving non-paraphilic arousal stimuli" means that the patient has been performing a lot of deviant sexual activities and no normal sexual activities. If the patient has been performing some normal sexual activities then he should not, according to First, be diagnosed as paraphilic.

Now, a man's total number of normal sexual experiences (e.g., his total number of consenting adult sexual partners) is not a matter of public record in the way that his number of known sexual-offense victims is. The only way a clinician can know how many normal sexual experiences a man has had is by asking him. Thus, all an uncooperative patient has to do to avoid a diagnosis of paraphilia is to claim that he had, during the time frame in question, a number of normal sexual experiences with one or more consenting adult sexual partners, and that during all

of them his mind was filled with thoughts of the situation at hand and not of fantasies that his partners were 11 years old or that he was strangling them rather than making love to them.

First's preference for patterns over counts may be based on a feeling that patterns are in the spirit of the DSM whereas counts are not. Yet, it is difficult to see how *victim counts* and *symptom counts* are completely different in form or function. Both approaches recommend a quantitative threshold for making a categorical decision. The DSM-IV-TR includes various examples of diagnoses based on a patient's showing *x* or more symptoms. This is especially true of the personality disorders. The core diagnostic criterion for Antisocial Personality Disorder (301.7), for example, requires the patient to have manifested three or more symptoms (p. 706). The sole diagnostic criterion for Borderline Personality Disorder (301.83) requires the patient to have shown five or more symptoms (p. 710).

Multiple Sexual Offenses Involving the Same Victim

First asserts as, established fact, that diagnosing pedohebephilia in uncooperative subjects from a minimum number of three victims would result in "a significant false negative problem." This follows from his premise that "pedophilia often involves multiple sexual offenses against one or two child victims who are known to the perpetrator." He cites no research to support this statement, nor does he attempt to define "often" as a proportion, computed on the total population of men assessed for pedohebephilia, of those men who offended against only one or two children, but on multiple occasions. In my experience, the men who get the opportunity to molest the same child on multiple occasions tend to be incest offenders, that is, men who have offended against their daughters or stepdaughters. The available evidence suggests that father-daughter incest offenders are *less* likely to be pedohebephilic than extrafamilial offenders (see original research and review by Blanchard et al., 2006).

There is a bigger problem, quite apart from the extra uncertainty in diagnosing father-daughter incest offenders, with attempting to diagnose pedohebephilia from the number of sexual interactions rather than the number of sexual victims: There is, to my knowledge, no published evidence that a man who has offended against a child on multiple occasions is more likely to be pedohebephilic than a man who has offended against a child on one occasion. There is therefore no justification for modifying the proposed criteria in order to diagnose pedohebephilia in uncooperative subjects on the grounds that they have interacted with a child on x or more occasions, or for y or more months. There is certainly no reason to drop the minimum victim criterion in order to accommodate such plausible but untested potential diagnostic signs. First himself reminds the reader that "one of the central requirements of the DSM-5 revision process is for recommendations to be grounded in empirical evidence."

Need for Formal Exclusionary Criteria

Clinicians who specialize in the assessment of sexual offenders are quite likely to see patients who, in addition to the sexual crimes that prompted the referral, present with mental retardation, antisocial personality disorder, alcohol or substance abuse, schizophrenia, (occasional) neurological disorders, and a miscellany of other conditions. The diagnostic question is essentially the same as in patients who present without obvious psychiatric disorders: Is the patient's sexual behavior probably the result of a paraphilia or of something else? The main difference between this group of patients and others is that the candidates for "something else" are more obvious.

First's comments on this head are quite brief, and it would be unfair to infer his complete views from them. It is noteworthy, however, that he does not mention the many cases in which a genuine paraphilia and some other psychiatric disorder are present in the same patient. Thus, he

discusses paraphilia and schizophrenia (for example) as competing explanations for some

observed sexual behavior, not as potentially comorbid conditions in which the patient's

paraphilia—not his schizophrenia—may be the motivating cause of his criminal sexual behavior.

In line with this perspective, he suggests that an exclusionary criterion should be added to all of

the paraphilias. This exclusionary criterion would list the other DSM diagnoses that should be

ruled out as alternative explanations of the patient's symptoms.

First justifies this recommendation, in part, with an example in which differential

diagnosis involves disorders with relatively confusable symptoms (specific phobia vs. OCD,

PTSD, or separation anxiety disorder). It is worth noting that such models do not put the problem

of diagnosing paraphilia in people with mental disorders in the most natural context: The

symptoms of schizophrenia are not similar to the symptoms of exhibitionism, and the symptoms

of mental retardation are not similar to the symptoms of pedohebephilia. Whether because he

draws on such models or for some other reason, First's recommendations for the assessment of

uncooperative patients emphasize differential diagnoses and ignore comorbid diagnoses, and he

accordingly stresses that clinicians assessing an uncooperative patient must rule out other

possible causes for sexual offending before making a diagnosis of paraphilia.

That is not my view at all. I think that the presence or absence of a paraphilia must, as

much as possible, be investigated independently of any other, obvious psychiatric problem. That

follows from my experience that paraphilias can be, and sometimes are, comorbid with other

DSM conditions. One cannot assume that a schizophrenic patient who exposed himself to a

strange woman did so *because* he has schizophrenia; he might have exhibitionism as well as

schizophrenia. One certainly cannot assume that a mentally retarded patient who molested a

child did so *because* he has mental retardation; there is evidence suggesting significant

comorbidity of pedohebephilia and mental retardation (e.g., Blanchard et al., 1999; Rice, Harris, Lang, & Chaplin, 2008). In my opinion, adding a boilerplate exclusionary criterion to each of the paraphilia criterion sets is likely to do more harm than good, in that it would steer clinicians away from the possibility of comorbid diagnoses.

Conclusion

I stressed the importance of avoiding false positive diagnoses in my first report to the Paraphilias Subworkgroup (Blanchard, 2010a), and I explicitly suggesting biasing diagnostic criteria toward making false negative diagnoses rather than false positive diagnoses. I have even been criticized for this (O'Donohue, 2010). I therefore believe that First and I are on the same side of this fundamental issue. I also believe, however, that his concerns about false positive diagnoses of paraphilia are exaggerated or misplaced, and that the changes he suggests for the proposed DSM-5 diagnostic criteria for the paraphilias are unnecessary or even undesirable.

References

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders*
(3rd ed.). Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders*
(4th ed.). Washington, DC: Author.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders*
(4th ed., text revision). Washington, DC: Author.

American Psychiatric Association (2010). Proposed draft revisions to DSM disorders and
criteria. http://www.dsm5.org Retrieved July 2, 2010.

Blanchard, R. (2010a). The DSM diagnostic criteria for Pedophilia. *Archives of Sexual Behavior,
39,* 304–316.

Blanchard, R. (2010b). The fertility of hebephiles and the adaptationist argument against
including hebephilia in DSM-5. *Archives of Sexual Behavior, 39,* 817–818.

Blanchard, R., Klassen, P., Dickey, R., Kuban, M. E., & Blak, T. (2001). Sensitivity and
specificity of the phallometric test for pedophilia in nonadmitting sex offenders.
Psychological Assessment, 13, 118-126.

Blanchard, R., Kolla, N. J., Cantor, J. M., Klassen, P. E., Dickey, R., Kuban, M. E., & Blak, T.
(2007). IQ, handedness, and pedophilia in adult male patients stratified by referral source.
Sexual Abuse: A Journal of Research and Treatment, 19, 285–309.

Blanchard, R., Kuban, M. E., Blak, T., Cantor, J. M., Klassen, P., & Dickey, R. (2006).
Phallometric comparison of pedophilic interest in nonadmitting sexual offenders against
stepdaughters, biological daughters, other biologically related girls, and unrelated girls.
Sexual Abuse: A Journal of Research and Treatment, 18, 1–14.

Blanchard, R., Kuban, M. E., Blak, T., Klassen, P. E., Dickey, R., & Cantor, J. M. (in press). Sexual attraction to others: A comparison of two models of alloerotic responding in men. *Archives of Sexual Behavior*.

Blanchard, R., Lykins, A. D., Wherrett, D., Kuban, M. E., Cantor, J. M., Blak, T., . . . Klassen, P. E. (2009). Pedophilia, hebephilia, and the DSM-V. *Archives of Sexual Behavior, 38,* 335–350.

Blanchard, R., Watson, M. S., Choy, A., Dickey, R., Klassen, P., Kuban, M., & Ferren, D. J. (1999). Pedophiles: Mental retardation, maternal age, and sexual orientation. *Archives of Sexual Behavior, 28,* 111–127.

Cantor, J. M., Klassen, P. E., Dickey, R., Christensen, B. K., Kuban, M. E., Blak, T., . . . Blanchard, R. (2005). Handedness in pedophilia and hebephilia. *Archives of Sexual Behavior, 34,* 447–459.

First, M. B. (in press). DSM-5 proposals for paraphilias: Suggestions for reducing false positives related to use of behavioral manifestations. *Archives of Sexual Behavior.*

Freund, K. (1965). Diagnosing heterosexual pedophilia by means of a test for sexual interest. *Behaviour Research and Therapy, 3,* 229–234.

Freund, K. (1967). Diagnosing homo- or heterosexuality and erotic age-preference by means of a psychophysiological test. *Behaviour Research and Therapy, 5,* 209–228.

Freund, K., Sedlacek, F., & Knob, K. (1965). A simple transducer for mechanical plethysmography of the male genital. *Journal of the Experimental Analysis of Behavior, 8,* 169–170.

Kuban, M., Barbaree, H. E., & Blanchard, R. (1999). A comparison of volume and

    circumference phallometry: Response magnitude and method agreement. *Archives of*

    Sexual Behavior, 28, 345–359.

Lykins, A. D., Cantor, J. M., Kuban, M. E., Blak, T., Dickey, R., Klassen, P. E., & Blanchard, R.

    (2010). The relation between peak response magnitudes and agreement in diagnoses

    obtained from two different phallometric tests for pedophilia. *Sexual Abuse: A Journal of*

    Research and Treatment, 22, 42–57.

O'Donohue, W. (2010). A critique of the proposed DSM-V diagnosis of Pedophilia. *Archives of*

    Sexual Behavior, 39, 587–590.

Rice, M. E., Harris, G. T., Lang, C., & Chaplin, T. C. (2008). Sexual preferences and recidivism

    of sex offenders with mental retardation. *Sexual Abuse: A Journal of Research and*

    Treatment, 20, 409–425.

Seto, M. C., Cantor, J. M., & Blanchard, R. (2006). Child pornography offenses are a valid

    diagnostic indicator of pedophilia. *Journal of Abnormal Psychology, 115,* 610–615.

Table 1. Frequency distribution of victim count, computed on all subjects

	Group			
	Teleiophiles		Pedohebephiles	
Victim count	n	%	n	%
0	335	41.0%	43	23.8%
1	289	35.4%	32	17.7%
2	121	14.8%	9	5.0%
3	42	5.1%	25	13.8%
4	14	1.7%	10	5.5%
5	9	1.1%	9	5.0%
≥ 6	7	.9%	53	29.3%
Total	817	100.0%	181	100.0%

Table 2. Frequency distribution of victim count, computed on subjects lacking charges or self-admissions regarding child pornography use

	Group			
	Teleiophiles		Pedohebephiles	
Victim count	*n*	%	*n*	%
0	306	39.6%	18	17.0%
1	280	36.3%	13	12.3%
2	116	15.0%	5	4.7%
3	40	5.2%	20	18.9%
4	14	1.8%	7	6.6%
5	9	1.2%	4	3.8%
≥ 6	7	.9%	39	36.8%
Total	772	100.0%	106	100.0%

Table 3. Frequency distribution of victim equivalency count, computed on all subjects

Victim equivalency count	Group			
	Teleiophiles		Pedohebephiles	
	n	%	*n*	%
0	306	37.5%	18	9.9%
1	280	34.3%	13	7.2%
2	145	17.7%	30	16.6%
3	49	6.0%	39	21.5%
4	19	2.3%	11	6.1%
5	11	1.3%	9	5.0%
6	2	.2%	13	7.2%
7	1	.1%	7	3.9%
≥ 8	4	.5%	41	22.7%
Total	817	100.0%	181	100.0%

Figure Captions

Figure 1. ROC curve for victim count, computed on all subjects.

Figure 2. ROC curve for victim count, computed on subjects lacking charges or self-admissions regarding child pornography use.

Figure 3. ROC curve for victim equivalency count (child pornography use counted as the equivalent of two victims and added to number of real-life victims), computed on all subjects.
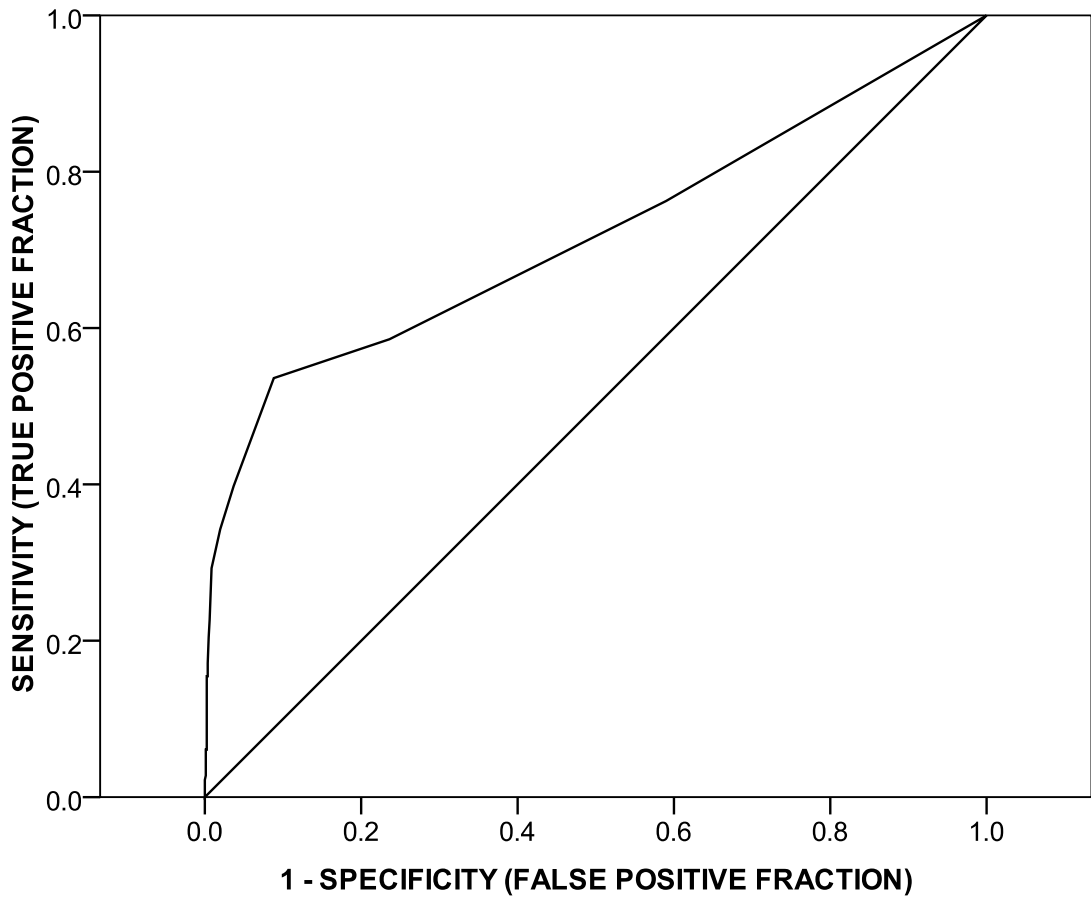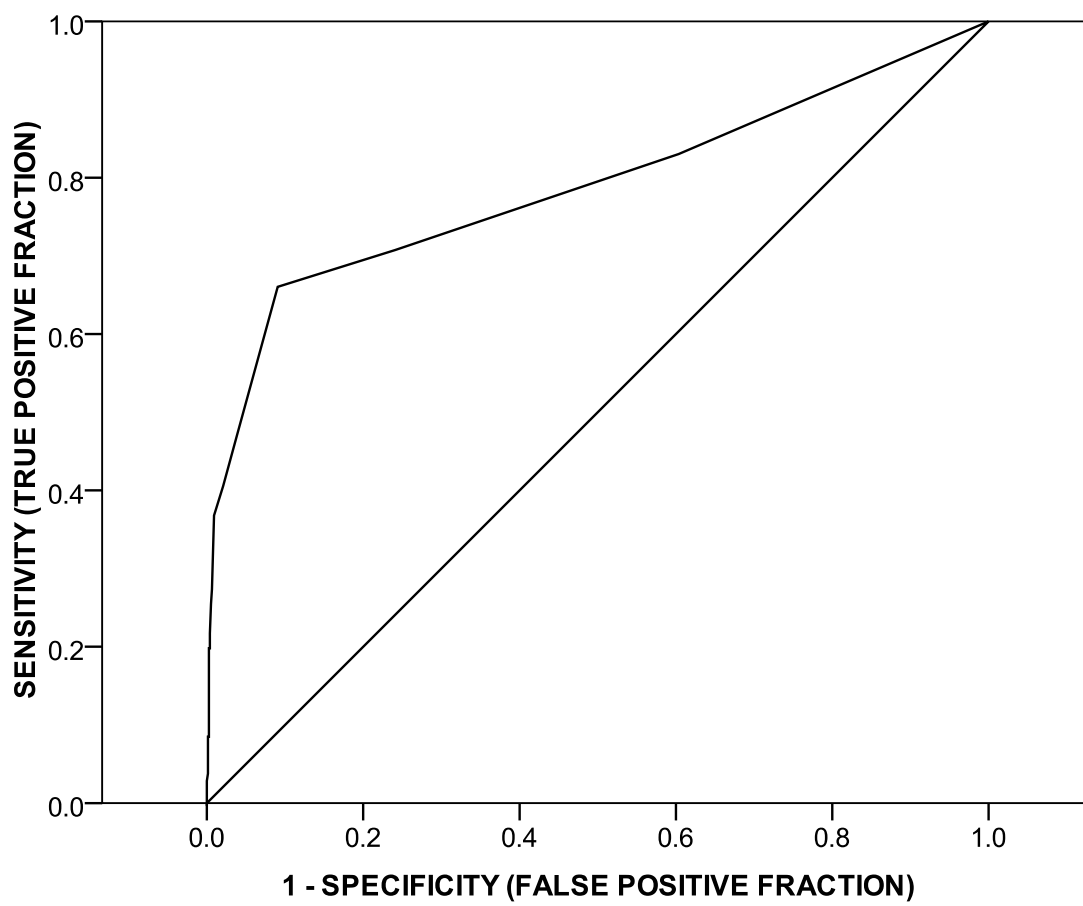
Figure 1

Figure 2

Figure 3