Evidence That Arousal to Pedophilic Stimuli Can Change: Response to Bailey, Cantor, and Lalumière

Article *in* Archives of Sexual Behavior · December 2014

DOI: 10.1007/s10508-014-0441-9 · Source: PubMed

6 authors, including:

Paul Fedoroff
University of Ottawa
251 PUBLICATIONS **5,873** CITATIONS
SEE PROFILE

Susan Curry
Royal Ottawa Mental Health Centre
51 PUBLICATIONS **1,129** CITATIONS
SEE PROFILE

Karolina Mueller
University Hospital Regensburg
79 PUBLICATIONS **670** CITATIONS
SEE PROFILE

Rebekah C. Ranger
Royal Ottawa Mental Health Centre
34 PUBLICATIONS **237** CITATIONS
SEE PROFILE

Arch Sex Behav
DOI 10.1007/s10508-014-0441-9

COMMENTARY

Evidence That Arousal to Pedophilic Stimuli Can Change: Response to Bailey, Cantor, and Lalumière

J. Paul Fedoroff · Susan Curry · Karolina Müller · Rebekah Ranger · Peer Briken · John Bradford

We thank Drs. Bailey, Cantor, and Lalumière for their careful reviews of Müller et al. (2014), hereafter referred to as "our study" or "our article." Our study reported on the results of a retrospective analysis of a group of men assessed at the Sexual Behaviours Clinic at the Royal Ottawa Mental Health Centre between 1983 and 2011, who fulfilled the following inclusion criteria: DSM-III, IV or IV-TR (American Psychiatric Association, 1980, 1994, 2000) diagnosis of pedophilia, an initial penile plethysmography test (PPT) indicative of sexual interest in children, and a second PPT test at least 6 months later.

Because we were interested in testing the hypothesis that arousal to pedophilic stimuli can change, we selected men whose PPT at Time 1 showed a greater increase in penile circumference in response to children compared to adults. Of this group, about half showed a greater increase in penile circumference in response to adults (as compared to children) at Time 2. In the article, we noted that the men who changed PPT response profiles demonstrated both a decrease of penile circumference change in response to audiotapes describing sexual interaction with children and an increase in penile circumference change in response to audiotapes describing sexual interactions between adults.

In our article, we suggested that the demonstration of a statistically significant decrease in sexual response toward children ($p < .001$), combined with a statistically significant increase in sexual response to adults ($p < .001$), presents a challenge to the claim that pedophilic sexual interest is unchangeable.

We note that there was an error that has been formally communicated to the Editor of the *Journal of Sexual Medicine*, due to the misclassification of three of the men in the study. The removal of these cases and complete re-analysis of the data did not change the results or our conclusions.

Bailey, Cantor, and Lalumière each claimed that the notion of the immutability of pedophilic interest is not challenged by the findings of our study and attempted to offer some non-evidence-based support for their opinions. We are grateful for the opportunity to respond and do so below by reviewing each commentary in alphabetized author order.

Bailey began his commentary by equating sexual orientation with sexual arousal, "Because a man's sexual orientation/erotic interest is identical to his characteristic sexual response pattern…" We disagree because sexual orientation is different from sexual arousal. Men who have sex with men are not necessarily gay, and a gay man is still gay even if he loses his sex drive or is unable to get an erection. In our article, we speculated that the reluctance of researchers to accept there is evidence that pedophilic sexual interest can change is due to confusion between sexual orientation and sexual interest. It should be noted that our study was never designed to investigate change in sexual orientation and we explicitly stated in the article that our study does not support any recommendations aimed at changing sexual orientation. Equating orientation with interest confuses interpretation of the issue and may be clinically harmful.

Bailey wrote that "…sometimes men do not get sufficient erection during a PPT to be accurately classified." We agree. This is why we selected men who initially produced more change in penile circumference in response to child stimuli than to adults, as well as only those who demonstrated a change in

J. P. Fedoroff (✉) · S. Curry · R. Ranger · J. Bradford
Forensic Research Unit, University of Ottawa Institute of Mental Health Research, Royal Ottawa Health Care Group, 1145 Carling Ave., Ottawa, ON K1Z 7K4, Canada
e-mail: paul.fedoroff@theroyal.ca

J. P. Fedoroff · J. Bradford
Division of Forensic Psychiatry, University of Ottawa, Ottawa, ON, Canada

K. Müller · P. Briken
Institute for Sex Research and Forensic Psychiatry, University Medical Centre Hamburg-Eppendorf, Hamburg, Germany

Published online: 03 December 2014

circumference of at least 3 mm in response to at least one audio stimulus during both testing sessions. This is like the two men in Bailey's lab who produced erections only to "multiple stimuli depicting male, but no erections to female stimuli." Bailey said that follow-up "yielded evidence of previously undisclosed sexual interest in men." In other words, Bailey appears to accept that PPT can accurately measure sexual interest. We agree, but only if there is some positive response (production of some change in penile circumference in response to specific sexual stimuli). This is what we demonstrated in our article.

Bailey next raised the important issue of error. Because this was a retrospective study, methods to decrease error were limited. For example, using means rather than maximum response was not possible because there were not always two or more equivalent stimuli used for the assessment for each stimulus category. This is why we advocated for prospective studies to verify the phenomenon we identified, namely decrease in PPT pedophilic response and increase in PPT non-pedophilic response in some men diagnosed with pedophilia.

Concerning systemic error, in 2005 we conducted a study (unpublished) in which we compared 50 men who admitted to a sexual interest in children or to having committed a hands-on sexual offense against one or more extra-familial children, with a control group of 50 men with no known criminal offenses or sexual interest in children. Repeated measures analysis of variance was used to determine if the groups differed on their phallometric responses to the audiotape stimuli.

There was a significant interaction between Group and audiotape Scenario, $F(5.23, 507.0) = 9.96$, $p < .001$, using Greenhouse-Geyser to correct for lack of sphericity. The tests of between-subjects effects revealed that the clinical group responded significantly more than the control group to the Child Initiate audiotape, $t(97) = -3.98$, $p < .001$, and to the Child Incest audiotape, $t(97) = -2.16$, $p < .05$. The control group responded significantly more to the Adult Mutually Consenting audiotape than the clinical group, $t(97) = 4.93$, $p < .001$, and to the Non-Sexual Assault audiotape, $t(97) = 3.38$, $p < .005$. The groups did not differ on their relative responses to the Child Mutual, Child Non-Physical Coercion, Child Physical Coercion, or Child Sadistic audiotapes, all $ts(97) < 1$.

Bailey said that we "considered any man whose PPT dropped by at least 0.50 standard units to have changed in PPT-measured arousal pattern," but provided no argument in support of this supposition. In fact, in our study, we showed a statistically significant change and showed there was a change from primary arousal in response to child audiotapes to a primary arousal to adult audiotapes (accompanied by no statistically significant responses to child audiotapes).

Bailey wondered why the inclusion criteria in our study included a diagnosis of pedophilia and an initial positive pedophile index (PI), which is a calculation comparing response to children with response to adults, positive indicating a higher response to children than adults. The answer is because we wanted to test the hypothesis that men diagnosed with pedophilia with confirmatory PPT response profiles can be shown to change their PPT response profiles. We deliberately chose a group of men with pedophilia for whom PPT testing "worked" in terms of being compatible with the diagnosis of pedophilia. We note that the DSM diagnostic criteria for pedophilia do not include PPT test results and all diagnoses in our study used DSM criteria only. Bailey was correct that all the men in our study with a diagnosis of pedophilia had a positive PI at Time 1 (by study design). The fact that there was a statistically significant change at Time 2 was supportive of the study hypothesis.

Bailey wrote, "If we were to accept these results as valid, they would suggest that men with greater sexual interest in children eventually change so that their sexual interest in children eventually changes so that their sexual interest in adults and children are approximately equal." We agree but go further since the findings of our study demonstrated that some men with a diagnosis of pedophilia and greater sexual interest in children change so that their sexual interest in adults, as measured by PPT, is greater than their sexual interest in children. We agree with Bailey that this finding results in a requirement to re-think "the common view of pedophilia: that it is a persistent sexual preference for children rather than adults." We recently commented on problems with current DSM diagnostic criteria for pedophilia that presumes pedophilic disorder cannot remit (Briken, Fedoroff, & Bradford, 2014). More specifically, the data in our study demonstrated that some men with pedophilia can show a change in their phallometric profiles.

Bailey argued that the results of our study were "almost certainly the effects of measurement error rather than true change." He based his opinion on the belief that some of the men in our sample did not actually have pedophilia and that their PPT responses to children at Time 1 represented a combination of two errors: falsely high responses to children and falsely low responses to adults. Presumably, he relied on the unfortunate combination of the same two errors but in the reverse direction at Time 2. He noted with a figure in his commentary that the variance of PI scores at Time 1 was less than the variance of PI scores at Time 2. He said that he cannot imagine any explanation for this other than random error and regression toward the mean.

However, an arguably more plausible explanation is that a true change did occur but not all men's PPT results changed the same amount, a fact we highlighted in our article (only about half of the men in our study showed significant change). Bailey wrote that his "argument does not prove that Müller et al.'s assessment protocol is essentially random or that this hypothetical situation fully explains the result." We agree. He adds that it was "unfortunate for Müller et al. that the mean of the subsequent PI was so close to zero…" In fact, Müller et al. were "fortunate" because the mean PI in the men who showed a change in PPT profiles was −1.70 (z score). While not included in the article, we calculated change using raw scores and PIs with the same results. Table 1 shows the analyses for the two groups using raw scores. We want

to emphasize that we do not consider ourselves "fortunate" (or not), based on the results of our study. The fact is that we set out to test a hypothesis and were simply reporting the results.

Bailey expressed concern with the "poor" test–retest correlation of PI scores but this assumes that PI scores do not change. He admits that the results of our study suggest either that the common notion of pedophilia as a moderately to very stable phenomenon is false and that pedophilic interests are much less stable than personality traits or that the measure of sexual interest employed by Müller et al. was rife with measurement error. We generally agree with the first part of Bailey's statement that the idea of pedophilic interest being changeable is supported. Given that Bailey has been unable to prove that the measure was "rife" with error, we maintain that our study does raise questions about the immutability of pedophilic interest. While we disagree with the conceptualization of pedophilia as a "personality trait," our study was not designed to test that novel conjecture. As a side note, the stability of personality traits as well as personality disorders has been questioned and changes have recently been reported for those conditions that previously also were considered to be unchangeable (e.g., Cooper, Balsas, & Oltmanns, 2014).

Bailey thinks it is a problem that the data failed to demonstrate a statistically significant relationship between whether the subjects changed and the interval between Time 1 and Time 2. He thinks this is because "…one would expect the amount of change to increase systematically with the time available in which to have changed." However, demonstration of the notion that change must occur in the way Bailey imagines is not required to challenge the hypothesis that change cannot occur.

Bailey concluded with some suggestions to improve the study. He recommended collecting self-report data (while simultaneously stating he doubts they would be reliable). We agree that self-report data are important, especially in making the diagnosis, which by definition requires behavioral and/or subjective data. He recommended "an intervention to alter sexual interest" and a control group that does not receive the "intervention." He is proposing a case controlled treatment study. Leaving aside the ethical issues of not treating people who are seeking treatment, we agree. However, as stated in our article, the current study was retrospective and not designed as a treatment study. We think the first task is to show change of sexual interest is possible. The next step(s) will be to determine what interventions facilitate change in which specific people.

Table 1 Within-group comparisons of sexual arousal pattern via *t* tests for paired samples using raw scores

Group	Initial phallometric assessment		Subsequent phallometric assessment		
	M	SD	*M*	SD	*p*
Interest changer (IC) ($n = 18$)					
Pedophilic index (raw scores)	6.51	5.38	−7.32	8.09	.001
Peripheral sexual arousal towards child stimuli (raw scores)	12.70	8.47	6.27	5.68	.009
Peripheral sexual arousal towards adult stimuli (raw scores)	6.19	6.03	13.59	9.34	.002
Interest non-changers (NC) ($n = 22$)					
Pedophilic index (raw scores)	7.83	6.74	6.10	4.74	ns
Peripheral sexual arousal towards child stimuli (raw scores)	12.95	8.71	10.07	5.03	ns
Peripheral sexual arousal towards adult stimuli (raw scores)	5.12	5.52	3.97	2.62	ns

Bailey quoted Freund who argued that positive phallometric test results are more significant (presumably meaning more likely to be true) when it contradicts a person's claim of favorable change than when it confirms such a change. This may be right but cannot be proven empirically in a real life setting. However, we disagree that data demonstrating a "favorable change" can be ignored, especially when the "favorable change" occurs in two opposite but complimentary directions. We agree with Bailey's recommendation that "… scientists should always be open to the possibility that their views are incorrect." He thinks "degree of openness will, and should, reflect available evidence." We think the data should trump any preconceived opinion or bias. Bailey concluded with reference to unpublished data derived from an Internet survey of respondents who rated attraction to children higher than their attraction to adults. He claims the data support "…very strong preferences that have been stable for many years and, as such, provide a strong challenge to the validity of Müller et al.'s results." We contend that even if the study was published and peer reviewed on the basis of the information provided, it would pose no challenge to the findings of our study.

Cantor candidly disclosed that his commentary was declined as a Letter to the Editor by *Journal of Sexual Medicine* and his arguments were based primarily on hypothetical data he made up to prove several points. He began by inventing an experiment in which a random group of men who might have pedophilia were tested in a PPT machine with a "chance" degree of accuracy. He claimed the "pertinent" task would be to accurately identify men "who deny pedophilia." We disagree. The study we conducted began with men already diagnosed with pedophilia, with 100 % showing PI scores indicative of greater sexual arousal to children than to adults.

Next, Cantor argued that ipsatized *z*-scores mean that increases in PPT scores in response to adult stimuli mathematically force a decrease in the PPT scores in response to child stimuli. However, we also calculated the results using raw scores and they were the same. Table 1 in Cantor's commentary, which consisted of made-up data with different characteristics than the data in our article, is therefore not applicable. This is because Cantor's invented data consist of data in which the arousal to adults did not change. It is also surprising to see Cantor appearing to argue that absolute change is more significant than relative change since, in another article, Blanchard et al. (2009) concluded the results of their study "…dramatically demonstrate the utility—or perhaps necessity—

of relative ascertainment in the laboratory assessment of erotic age-preference" (p. 431).

The final commentary was by Lalumière. In the interest of full disclosure, Lalumière has recently joined the research lab where our study was conducted but was not involved in our study. He began with some general statements that deserve clarification. He wrote that "Phallometry involves the measurement of erections while examinees are exposed to a variety of sexual stimuli in the laboratory." It is more correct to state that PPT involves measurement of changes in penile circumference since most men do not achieve full erection during PPT (however, we note that two did in Bailey's lab). Lalumière cited a statement made by Fedoroff on a national radio show in Canada. While we do not dispute the statement, it is more correct to indicate that Fedoroff is on record as questioning the popular claim that the hypothesis that pedophilic sexual interest is unchangeable has been proven (e.g., Fedoroff, 2009). Fedoroff has never said all people with pedophilia can change. In fact, our study was designed to investigate whether *any* man with pedophilia can be shown to have changed his interest as measured by PPT.

Lalumière quoted Seto (another member of the lab where the study was conducted but, who like Lalumière, was not involved in the study that is under debate here). He says there is a "suggestion that pedophilia can be thought of as a sexual orientation…akin to heterosexuality and homosexuality." We absolutely disagree with equating homosexual orientation and pedophilia, and in our article suggested that the reason why the results of our study arouse skepticism is due to confusion between sexual orientation and sexual interest. Gay men can show positive responses to stimuli involving men on PPT but, as Bailey pointed out, a lack of penile tumescence does not mean lack of interest. We go further to state that lack of sexual interest does not preclude homosexuality. Whether or not readers agree is not necessary to evaluate the significance of our study, which was not a study of sexual orientation but rather a study of whether or not some men with the diagnosis of pedophilia show changes in a proxy for sexual interest, which is PPT.

Lalumière next reviewed Cantor's hypothetical (and, in our view, invalid) example in detail. We agree with Lalumière's comment that "The flip of a coin possibility could be discounted if the phallometric assessment used in the Müller et al. (2014) study had known sensitivity and specificity." The sensitivity and specificity of PPT assessment from a study conducted in 1997 was provided on p. 1223 of our article (sensitivity = 77.8 %; specificity = 76.3 %). He wonders if the diagnoses were made "independently of the phallometric results." The answer is "yes" since the DSM criteria are independent of PPT results, as indicated in our article.

Next, Lalumière essentially reproduced the hypothetical experiment suggested by Cantor. He stated, "There is no way to tell how many of the 43 men included in the study were truly pedophilic and how many were not, even if we accept the sensitivity and specificity values reported by Müller et al." We suppose it is possible that men sometimes fake having pedophilia and are able to fake pedophilic PPT results. In fact, Fedoroff published an article on that topic (Fedoroff, Hanson, Mcguire, Malin, & Berlin, 1992). However, to use Bailey's term, we think it is "implausible." Even if the study sample was infiltrated by pedophile pretenders, it does not explain why they would decide to stop faking at Time 2. Lalumière does raise the interesting question of what a "true" pedophile is. For the purpose of our study, the term was defined according to DSM criteria and supported by PPT testing at Time 1.

Lalumière goes further and stated "…we cannot tell the proportion of true pedophiles among the 43 used in the Müller et al. (2014) study. We only know that it was less than 100 %." How do we know? In fact, we suggest that the men in our study were exactly the sorts of men who skeptics claim are unchangeable. The argument that any man with pedophilia who changes into being a man with no symptoms of pedophilia must not have had pedophilia in the first place is a good example of what Cantor refers to as a tautology. It may explain why in the DSM-5 diagnosis of pedophilic disorder is the only paraphilia that cannot be designated as in remission (Briken et al., 2014). It is of interest that even using Lalumière's hypothetical data and analysis, he predicted "10 would be expected to not be detected…on the second assessment, based on the sensitivity value reported absent any actual changes in sexual interests." We found 18.

Lalumière next repeated Cantor's concerns about *z*-scores to which we have responded above. There are some parts of his discussion that bear correction. Lalumière suggested that men in our study may have learned from other offenders how to fake the PPT (e.g., by thinking about a child when they see an image of an adult). In the current study, all stimuli were auditory. We agree it is possible that men in our study learned to suppress responses to children (in fact, that is an aim of treatment). However, to our knowledge, there are no previous reports of successful acquisition of greater response to adult stimuli compared to child stimuli (both in absolute and relative terms) as measured by PPT. He suggests that we analyze the raw scores. As indicated, this was done with the same results.

Lalumière also offers "regression to the mean" as an explanation of our study's results. However, regression to the mean does not explain the increase in sexual interest in adults above the mean to the point where the men who changed interest would be classified as showing non-pedophilic PPT response profiles if only their Time 2 data were analyzed. Also, because this was a retrospective study, we did not have a control group or multiple baseline measurements to account for regression to the mean.

Lalumière made several suggestions to improve the study, including selection of a group based on the relevant diagnosis and statistical corrections. These recommendations were followed in our study. For future studies, Lalumière also recommended re-assessing all study participants regardless of concerns about

dangerousness, etc., a randomized treatment control design and using multiple indicators. We agree, provided ethical concerns can be met.

We hope we have answered the questions posed and stimulated readers to reconsider the possibility that sexual interest can and does change. In our opinion, the best way to confirm or refute the findings of our study is by attempted replication. We challenge our critics (and "supporters") to attempt to replicate or empirically refute our study and to publish their results. In the meantime, we plan to do the same. We are especially interested in testing the hypothesis that sexual orientation is indeed different from sexual interest.

References

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

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

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

Blanchard, R., Kuban, M., Blak, T., Cantor, J., Klassen, P., & Dickey, R. (2009). Absolute versus relative ascertainment of pedophilia in men. *Sexual Abuse: A Journal of Research and Treatment, 21*, 431–441.

Briken, P., Fedoroff, J. P., & Bradford, J. W. (2014). Why can't pedophilic disorder remit? *Archives of Sexual Behavior, 43*, 1237–1239.

Cooper, L. D., Balsis, S., & Oltmanns, T. F. (2014). Aging: Empirical contribution. A longitudinal analysis of personality disorder dimensions and personality traits in a community sample of older adults: Perspectives from selves and informants. *Journal of Personality Disorders, 28*, 151–165.

Fedoroff, J. P. (2009). The paraphilias. In M. G. Gelder, N. C. Andreasen, J. J. Lopez-Iber, & J. R. Beddes (Eds.), *The new Oxford textbook of psychiatry* (2nd ed., pp. 832–842). Oxford: Oxford University Press.

Fedoroff, J. P., Hanson, A., Mcguire, M., Malin, H. M., & Berlin, F. (1992). Simulated paraphilias: A preliminary study of patients who imitate or exaggerate paraphilic symptoms and behaviours. *Journal of Forensic Sciences, 37*, 902–911.

Müller, K., Curry, S., Ranger, R., Briken, P., Bradford, J., & Fedoroff, J. P. (2014). Changes in sexual arousal as measured by penile plethysmography in men with pedophilic sexual interest. *Journal of Sexual Medicine, 11*, 1221–1229.