

## Purported Changes in Pedophilia as Statistical Artefacts: Comment on Müller et al. (2014)

James M. Cantor

Published online: 22 July 2014  
© Springer Science+Business Media New York 2014

Müller et al. (2014) claimed to have observed large changes in sexual arousal patterns among pedophiles, including both decreases in pedophiles' responses to stimuli representing children and increases in those to adults. Although such a finding could be of great import, that team's result is also attributable to—and more likely the result of—their application of an unsound research design and misinterpreted statistics.

The consequence of the first error is readily seen through concrete example: Assume that the phallometric test they employed detects pedophilia only at chance, with test-takers having a 50–50 chance of being labelled pedophilic. Among 100 men undergoing the test, (an average of) 25 men would be labelled pedophilic both times, 25 would be labelled non-pedophilic both times, 25 would first be labelled pedophilic and subsequently as non-pedophilic, and the remaining 25 would first be labelled non-pedophilic and subsequently as pedophilic. The Müller et al. method retained only the cases labelled pedophilic in test session one: That method would exclude the 25 who came out non-pedophilic both times and the 25 who were non-pedophilic-then-pedophilic, while retaining the 25 who came out pedophilic both times and the 25 who were pedophilic-then-non-pedophilic. Thus, the method would artificially select a sample of 50 men labelled pedophilic, 25 (50 %) of whom would score non-pedophilic in test session two. Müller et al.'s reported result ("change" in 48.9 % of their sample of 43 men) is indistinguishable

from what would emerge from entirely random data. Although they claimed to be presenting evidence of genuine change among pedophiles, Müller et al. instead presented evidence suggesting complete invalidity of their clinic's phallometric test protocol. Remarkably, although Müller et al. refer to having unpublished data that suggest some accuracy for detecting pedophilia among men who already admit to being pedophilic (thus being rather moot), they neither provided nor cited any data supporting the validity of their test with men who deny pedophilia (the pertinent task).

The Müller et al. team also erred in interpreting their statistics to be showing an increase in sexual arousal towards adults over and above any decreases in response towards children. Müller et al. appear unaware that, after z-score transformation, all scores become entirely dependent: Any seeming increases in response to any subset of stimuli are merely the mathematical reflection of decreases in the remaining subset. This is again best seen by concrete example:

Table 1 provides a hypothetical case, showing his absolute responses (mm of change in penile circumference) and transformed scores (ipsatized z-scores). In test session one (upper panel), the subject shows obviously greater arousal to children, as reflected by both absolute and transformed scores. In test session two, the subject shows exactly the same amount of penile response to adults (in mm) as he did in the first session while selectively suppressing his responses to children. The transformed scores, however, greatly increased in numeric value, despite that the actual arousal levels were entirely unchanged. Although one could assert that the subject showed a greater *relative* response to adults than when he selectively suppressed his arousal to children in test session two, it is an error to believe that the subject manifested an actually greater response to (or interest in) adults than previously.

Thus, despite their claims, Müller et al. are reporting only a statistical tautology reflecting an unsound research design,

---

J. M. Cantor (✉)  
Complex Mental Illness Program, Centre for Addiction and Mental Health, 250 College St., Toronto, ON M5T 1R8, Canada  
e-mail: James.Cantor@camh.ca

J. M. Cantor  
Department of Psychiatry, University of Toronto, Toronto, ON, Canada

**Table 1** Hypothetical case of a pedophile undergoing two phallometric test sessions, one before learning to suppress penile responses to children (upper panel) and one after (lower panel)

	Responses to child stimuli			Responses to adult stimuli			M (SD)
Session one							
Absolute score (mm)	14.00	16.00	18.00	10.00	10.00	10.00	13.00 (3.52)
Ipsatized z-score	0.28	0.85	1.42	−0.85	−0.85	−0.85	0.00 (1.00)
Session two							
Absolute score (mm)	4.00	6.00	8.00	10.00	10.00	10.00	8.00 (2.53)
Ipsatized z-score	−1.58	−0.79	0.00	0.79	0.79	0.79	0.00 (1.00)

misinterpreted statistics, and, very plausibly, an entirely invalid phallometric test protocol.

**Acknowledgments** Müller et al. (2014) was published in the *Journal of Sexual Medicine*, so this commentary was originally submitted there as a Letter to the Editor. Because that journal declined to publish it, it was submitted instead to *Archives*.

## Reference

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.