A Technical Comment on the paper:

"Sexual Orientation-Related Differences in Prepulse Inhibition of the Human Startle Response"

Qazi Rahman, Veena Kumari, and Glenn D. Wilson
Behavioral Neurocience, 2003, Vol. 177, No. 5, pp 1096-1102

This Comment by:
Halstead Harrison
University of Washington
Seattle, WA 98195-351640
206-543-4596 [voice] 206-543-0308 [FAX]
<harrison@atmos.washington.edu>

December 15, 2003

Abstract:

Data presented by Rahman *et al*. do not confidently support their finding that homosexual women exhibit a male-type startled-blink reflex.

Introduction:

In their interesting paper, Q. Rahman, V. Kumari, and G.D. Wilson [Rahman *et al*., 2003, hereafter "RKW"] report a discernable difference in prepulse inhibition of the human startled-blink response, between heterosexual and homosexual women, with the latter group's response more closely resembling that of heterosexual men. In this Comment I re-examine RKW data from a statistical perspective, concluding that no significant differences were detected.

Data:

For convenience in this discussion, the primary data from RKW are reproduced in Table I. The left four columns of the first rows of each pair of data lines in that table are pulse-startle responses, [p1, p2, ..], in arbitrary units. Each data entry reports an average of 6 replications among 15 subjects in each group [except 14 for homosexual women, FM]. Columns 5-8 in Table I contain similar averages over prepulse-inhibited startle responses of the same subjects, [pp1, pp2, ..]. The second row of each pair of lines lists standard deviations [SD] of the measurements above. For further details about these data and experimental protocols, please refer to the original paper [Rahman *et al*., 2003].

TABLE I
Upper numbers are averaged pulse [p1, ..] and prepulse [pp1, ...]
inhibited startle-blink reflexes, in arbitrary units.
Lower numbers are standard deviations [SD].

"MM" = Heterosexual men "FF" = Heterosexual women
"MF" = Homosexual men "FM" = Homosexual women

p1	p2	p3	p4	pp1	pp2	pp3	pp4	
336.80	318.42	297.71	227.84	205.17	166.95	154.08	144.83	MM
231.86	258.40	243.42	168.88	183.50	141.92	144.97	127.43	SD
355.35	310.04	282.92	275.66	291.40	302.17	238.88	258.78	FF
267.71	264.26	254.42	232.42	265.88	277.55	215.52	228.11	SD
431.95	351.75	321.48	353.04	241.62	237.95	206.40	222.84	MF
269.45	259.52	205.82	209.93	165.64	153.33	131.02	163.14	SD
489.95	359.80	346.66	291.69	268.40	227.88	196.90	221.09	FM
360.32	210.37	245.85	153.56	185.18	139.52	93.61	127.14	SD

Analysis:

RKW define a "prepulse inhibition index", PPI(i), from the data of Table I as

$$PPI(i) = 100\,[1 - pp(i) / p(i)] \quad i = 1 \text{ to } 4 \text{ for MM, FF, MF, FM} \qquad [1]$$

These PPI indices are percentage measures of the tendency of a closely preceding noise pulse to suppress a startled-blink reflex induced by a subsequent pulse. They are summarized in Table II.

TABLE II
Pre-pulse Inhibition Indexes

PPI(1)	PPI(2)	PPI(3)	PPI(4)		SE_x	$\langle\langle SE_i \rangle\rangle$	$\langle PPI \rangle \pm SE_{ppi}$	
39.08	47.57	48.24	36.43	MM	2.98	.....	**42.83 ± 18.86**	MM
19.07	17.09	17.89	20.28	SE_i	.....	18.62		
18.00	2.54	15.57	6.12	FF	3.70	.....	**10.56 ± 30.92**	FF
26.91	33.87	29.83	31.76	SE_i	.....	30.70		
44.06	32.35	35.80	36.88	MF	2.46	.....	**37.27 ± 16.58**	MF
14.38	18.38	16.05	16.51	SE_i	.....	16.39		
45.22	36.66	43.20	24.20	FM	4.74	.....	**37.32 ± 16.36**	FM
15.95	15.48	14.00	17.06	SE_i	.....	15.66		

In Table II the upper rows of the first four columns contain PPI indices derived from Table I, through equation [1]. The same four columns of the second of each pair of data rows contain the *internal* standard errors [SE_i] of the PPI indices above them, computed as

$$SE_i[PPI(i)] = (A/B)\ [(\delta A/A)^2 + (\delta B/B)^2] \ /\ [N - 2]^{1/2} \qquad [2]$$

Where $A = pp(i)$ $\delta A = SD[pp(i)]$ from Table I
$B = p(i)$ $\delta B = SD[p(i)]$ from Table I
$N = 15$ for MM, MF, and FF, or 14 for FM

Equation [2] derives from a small-error expansion that assumes Gaussian probability distribution functions [pdfs] and no correlation between δA and δB[1].

The fifth data column in Table II lists the *external* standard errors, SE_x, of the replicated measurements listed in the first four columns. These SE_x were estimated with three disposable degrees of freedom, as

$$SE_x(i) = \{<(PPI(i,j) - <PPI(i,j)>)^2>^{1/2} \ /\ [J - 1]\}^{1/2} \qquad [3]$$

where $J = 4$, and the notation $<x^k>$ refers to $(1/J)\Sigma\ x^k$ $j=1$ to 4

These external SE_x may be thought of as best-current estimates of the parameters, σ_x, of Gaussian pdfs that might be achieved if many more than four replications of PPI(i,j) had been measured on the *same* set of test subjects. Note in Table II that these external SE_x are commendably less than the internal SE_i, which are precision-limiting with this experiment.

The sixth data column of Table II, $<<SE_i>>$, contains the root-mean-squares of the *internal* standard errors, SE_i. Note that it is not proper at this step to attempt a divisor by $[J - 1]^{1/2}$, because the separate test groups [MM, FF, etc.] have intrinsic, "real" variance that may not be diminished by repeated measurements.

---

[1] With $F = 1 - A/B$, $SE[F(A,B)] = [(\partial F/\partial A\ \delta A)^2 + (\partial F/\partial B\ \delta B)^2]^{1/2} \ /\ [N-2]^{1/2}$. Square both sides, assume $<\delta A\ \delta B> = 0$, and take the square-root again. A Gaussian pdf of ξ, a continuous variable with zero mean, is $P(\xi)d\xi = [2\pi\sigma^2]^{-1/2} \exp[-1/2\ (\xi/\sigma)^2]d\xi$.

The last two, bold-faced data columns of Table II contain the linear averages, <PPI(i)>, of the four PPI(i,j) columns, with j=1 to 4, and their standard errors, $SE_{ppi} = [(SE_x]^2 + <<SE_i>>^2]^{1/2}$. As with the previous example, these latter may be thought of as the best current estimates, σ_{ppi}, of Gaussian pdfs that might be measured with similar protocols applied to many new and independent sets of 15 [or 14, for FM] subjects, if they were available.

From the last two columns of Table II it is straightforward to compute PPI differences,[$\Delta(k) = <PPI(j)> - <PPI(i)>$, i = 1 to 3, j = i + 1 to 4], and their standard errors, $SE(k) = [SE_{ppi}(i)^2 + SE_{ppi}(j)^2]^{1/2}$, between the six possible pairings of <PPI>, as in Table III.

TABLE IIIa
Pre-Pulse Inhibition Differences

k	Δ(k)		SE(k)	Label
1	32.28	±	35.91	MM - FF
2	5.56	±	24.81	MM - MF
3	5.51	±	24.33	MM - FM
4	-26.72	±	34.80	FF - MF
5	**-26.77**	**±**	**34.46**	**FF - FM**
6	-0.05	±	22.67	MF - FM

Discussion:

None of the PPI differences in Table III displays a magnitude exceeding its standard error, SE. In particular, the bold-faced line reveals no significant difference in prepulse startle-reflex inhibition between heterosexual [FF] and homosexual [FM] women. Alert readers will have guessed this result from the large standard deviations reported by RKW in their Table I, and from the small test-subject counts in each group.

A few further remarks suffice to conclude this Comment.

Always apprehensive of statistical subtleties, I have tested these calculations with Monte-Carlo simulations of separate numerical experiments upon 15- [or 14-] member groups of each class [MM,FF, ..], using in each case 10,000 Gaussian random numbers generated to have the same means and standard deviations reported in Table I.

After checking to confirm that these simulations reproduced the entries Table I to within ± 0.02, it was not a surprise that subsequent manipulations reproduced Table III with negligible differences. But if I calculated a PPI index immediately after "testing" each group, and then assembled statistics directly from these PPI without first grand-averaging into p(i) and pp(i)], the resulting SEs were significantly greater than those of Table III. This effect results from the large SDs of Table I, which with aggregations too small to approach the Central Limit Theorem admit many large and unphysical PPI cases, of both signs, that result from near zero [and negative!] p(i) in the denominator of equation [1]. That is, the true pdfs of the PPI, measured in small sets, are certainly *not* Gaussian, and the small-error approximation of equation [2] is only marginally valid. Correcting for this effect would require explicit knowledge of the experimental pdfs, which were not reported by RKW.

Additional concerns can be raised about effects of test-subject training and of correlations between p(i) and pp(i). Both would tend further to increase SEs and to diminish confidence.

This Comment should not be construed as falsifying the hypothesis that homosexual and heterosexual women display different prepulse startle-inhibition reflexes. That conjecture may turn out to be so, but the present data do not confidently support it.

Reference:

Rahman, Q, Kumari, V, & Wilson, G.B. (2003)
Sexual Orientation-Related Differences in Prepulse Inhibition of the Human Startle Reflex *Behavioral Neuroscience, 117*,1096-1102