

June 14, 2007

Dear Reader:

Professor Bushman and Anderson's response¹ to my and Dr. Crain's letter² is disturbing.

In their often referenced article, Bushman and Anderson tried to show that various illnesses and social problems (ex., smoking and lung cancer) only marginally correlate with known risk factors. They then tried to link "aggression," using comparable statistics, with the exposure to violence in the media. By finding similar correlations, the authors hoped to prove that viewing violence on TV or in videogames is a significant risk factor.

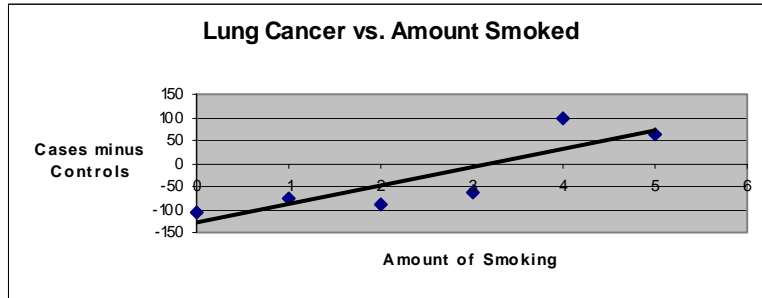
Unfortunately, it is now revealed that Bushman and Anderson never calculated actual correlations. Rather, they used *estimates*. Nevermind that they never disclosed this in their article; one should now ask how accurate their estimates were? Was 2+2 equal to "about" 4, or something vastly different?

When scientists use approximants, they describe the exact assumptions that make the estimate more or less trustworthy. They also calculate a numerical measure of their accuracy. Quantifiable measures, such as confidence intervals for estimates, are a necessary step before using such methods in research. Otherwise, the estimates become little more than guesses. Bushman and Anderson's article, even in its recently amended form, provides neither a list of their assumptions nor any measure of the error introduced by their manipulation of the data. This is inexcusable. Moreover, the problems are compounded by Bonnett's analysis. Bonnett described several different ways to *estimate* a correlation, depending on the specific data.³ Why and how did Bushman and Anderson chose one of these methods and not a different one? The authors never addressed this.

There are more problems. Bushman and Anderson converted statistical Relative Risks into correlations. They do not describe or justify the method. The estimating formulae outlined by Bonett use Odds Ratios, not Relative Risks. However, as Bonett discussed, the Relative Risk can, under very special circumstances, approach equaling the Odds Ratio - i.e., when one is researching an infrequent outcome. Did Bushman and Anderson assume this was the case? With this data, was that assumption valid? How much more error did this procedure introduce?

Taking an example from the Bushman and Anderson's article, I tried to explore these issues. Using the Wynder and Graham data on Smoking and Lung Cancer⁴ (see Table 1 from the 2007 Bushman and Anderson reply), we compared the "smoking status" variable to the "difference" variable and calculated a Pearson's correlation of 0.86 (p<.001). For unknown reasons, in their reply Bushman and Anderson were unable to compute this degree of correlation. I stand by my number. For those who are curious, I refer you to the supporting references, showing my calculations in detail.

Smoking Status (coded)	Number of Patients With Lung Cancer (A)	Number of Patients Without Lung Cancer (B)	Difference (A-B)
Nonsmoker (0)	8	114	-106
Light (1)	14	90	-76
Moderate (2)	61	148	-87
Heavy (3)	213	278	-65
Excessive (4)	187	90	97
Chain (5)	123	59	64



I then collapsed the data from a 6x2 table into a 2x2 table, losing much information in the process. One would not normally do this but it allowed me to empirically look at the accuracy of the estimates.

	No Lung Cancer	Lung Cancer
"None" to "Moderate" Smoking	352	83
"Heavy" to "Chain" Smoking	427	523

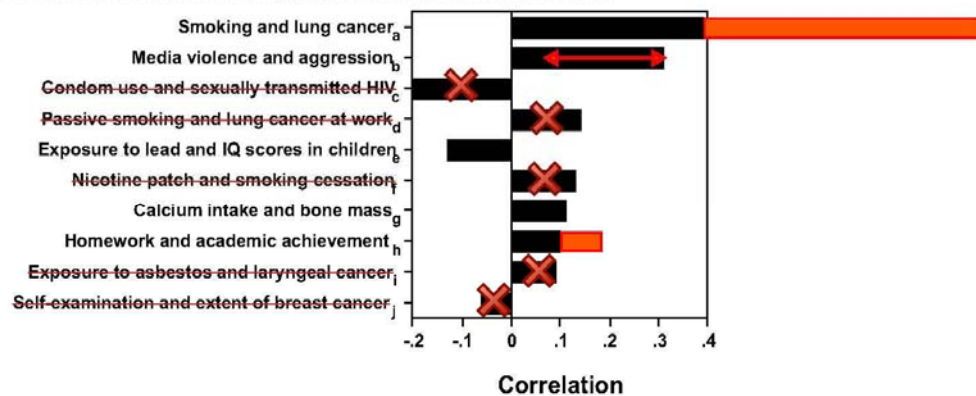
In the 2x2 table, the calculated correlation (the phi coefficient, as per Bonett) is seen to drop to 0.34 ($p < .001$). Using this reduced dataset, the Odds Ratio for developing lung cancer among the heavy smokers was calculated to be 5.19. The Relative Risk was 1.80. Finally, I calculated⁵ the *estimated* correlation, using the Digby's tetrachoric approximation that Bushman and Anderson used (see Bonett).

	Value
Pearson correlation (6x2 table)	0.86
Reduced (2x2 table) phi correlation	0.34
Estimated correlation based on Relative Risk	0.22 (35.3% underestimated)
Estimated correlation based on Odds Ratio	0.55 (61.8% overestimated)

While admittedly this is just one example, it is alarming. These are huge errors - especially the overestimate derived from the Odds Ratio, which should be the more accurate of the two. If we believed Bushman and Anderson's estimates, 2+2 equals about 6.5.

In my analysis, Bushman and Anderson's Figure 2 might appear something like the chart that follows. In the diagram, the changes in red indicate my findings upon trying to replicate their calculations. The lines noted with an "X" represent data that are estimates and should be excluded. While other parts of the article were also flawed and irreproducible, Figure 2, alone, is enough to indicate serious problems.

Figure 2
Comparison of the Effect of Violent Media on Aggression With Effects From Other Domains



Note. All correlations are significantly different from zero. a = the effect of smoking tobacco on lung cancer, as estimated by pooling the data from Figures 1 and 3 in Wynder and Graham's (1950) classic article. The remaining effects were estimated from meta-analyses: b = Paik and Comstock (1994); c = Weller (1993); d = Wells (1998); e = Needleman and Gatsonis (1990); f = Fiore, Smith, Jorenby, and Baker (1994); g = Weiten, Kemper, Past, and van Staveren (1995); h = Cooper (1989); i = Smith, Handley, and Wood (1990); j = Hill, White, Jolley, and Mapperson (1988).

My conclusion from this mess is that we should be very wary of meta-studies and, more specifically, those studies that purport to change statistical measures, like Odds Ratios or Relative Risk, into correlation estimates. If researchers must use such estimates, they need to do so cautiously. At a minimum, they must disclose what they are doing, reference what formula is being used, justify its use, describe what assumptions were made about the data, and provide the reader with a confidence interval for the estimate.

Bushman and Anderson's paper does none of these and, regrettably, we should question its validity, at least until it is completely reanalyzed. We should also wonder what other studies in the medical literature have used such methods to estimate correlation, besides the one we have identified⁶. One might also question the quality of a peer-review system that allows such statistically flawed articles, such as Bushman and Anderson's, to be published.

Jerald Block, MD

¹ Bushman BJ, Anderson CA (2007). Measuring the Strength of the Effect of Violent Media on Aggression. *American Psychologist*.

² Block JJ, Crain BR. (2007). Omissions and Errors in "Media Violence and the American Public". *American Psychologist*.

³ Bonett DG (2007). Transforming Odds Ratios Into Correlations for Meta-Analytic Research. *American Psychologist*.

⁴ Wynder EL, Graham EA (1950). Tobacco smoking as a possible etiological factor in bronchiogenic carcinoma. *JAMA*. 143:329-336.

⁵ Block JB (2007). Smoking and Cancer Statistics. Last accessed March 9, 2007. www.jeraldjblock.medem.com.

⁶ Bushman BJ, Anderson CA (2001). Media violence and the American public: Scientific facts versus media misinformation. *American Psychologist*. 56: 477-489.