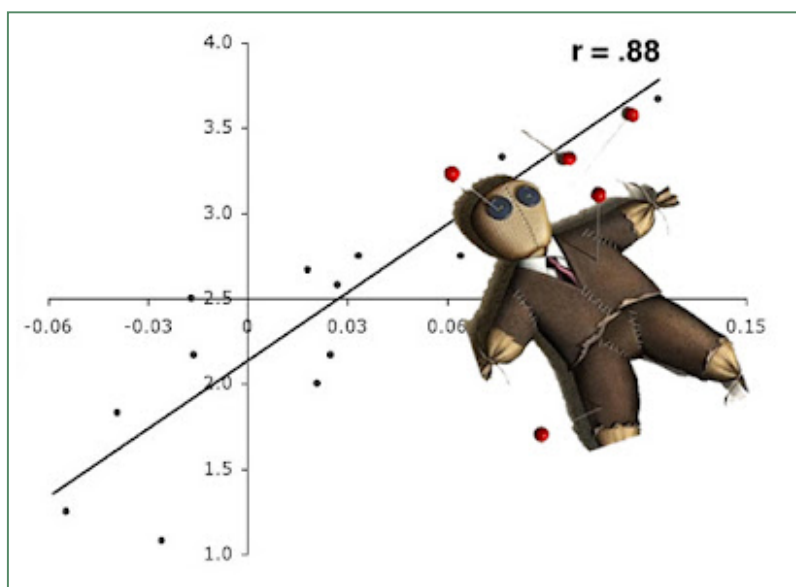


The Neurocritic

Deconstructing the most sensationalistic recent findings in Human Brain Imaging, Cognitive Neuroscience, and Psychopharmacology

MONDAY, JANUARY 05, 2009

Voodoo Correlations in Social Neuroscience



The end of 2008 brought us the tabloid headline, [Scan Scandal Hits Social Neuroscience](#). As initially reported by [Mind Hacks](#), a new "bombshell of a paper" (Vul et al., 2009) questioned the implausibly high correlations observed in some fMRI studies in [Social Neuroscience](#). A new look at the analytic methods revealed that over half of the sampled papers used faulty techniques to obtain their results.

[Edward Vul](#), the first author, deserves a tremendous amount of credit (and a [round of applause](#)) for writing and publishing such a critical paper under his own name [unlike all those cowardly pseudonymous bloggers who shall go unnamed here]. He's a graduate student in [Nancy Kanwisher's Lab](#) at [MIT](#). Dr. Kanwisher¹ is best known for her work on the [fusiform face area](#).

Credit (of course) is also due to the other authors of the paper ([Christine Harris](#), [Piotr Winkielman](#), and [Harold Pashler](#)), who are at the University of California, San Diego. So without further ado, let us begin.

A Puzzle: Remarkably High Correlations in Social Neuroscience

Vul et al. start with the observation that the new field of Social Neuroscience (or Social Cognitive Neuroscience) has garnered a great deal of attention and funding in its brief existence. Many high-profile neuroimaging articles have been published in *Science*, *Nature*, and

About Me



Name:
The Neurocritic

Born in West Virginia in 1980, The Neurocritic embarked upon a roadtrip across America at the age of thirteen with his mother. She abandoned him when they reached San Francisco and The Neurocritic descended into a spiral of drug abuse and prostitution. At fifteen, The Neurocritic's psychiatrist encouraged him to start writing as a form of therapy.

[View my complete profile](#)

Previous Posts

[Scan Scandal Hits Social Neuroscience](#)

[Deal, No Deal, or Dots?](#)

[Fainting In The Name Of](#)

[Dangerous Sandwiches](#)

[Snotty or Nice?](#)

[More Mindreading](#)

[Can You Reread My Mind?](#)

[Crime, Punishment, and Jerry Springer](#)

[However, Jerry Springer says he has no talent.](#)

[The Jerry Springer of academia?](#)



Follow

Neuron, and have received widespread coverage in the popular press. However, all may not be rosy in paradise:²

Eisenberger, Lieberman, and Williams (2003), writing in *Science*, described a game they created to expose individuals to social rejection in the laboratory. The authors measured the brain activity in 13 individuals at the same time as the actual rejection took place, and later obtained a self-report measure of how much distress the subject had experienced. Distress was correlated at $r=.88$ with activity in the anterior cingulate cortex (ACC).

In another *Science* paper, Singer et al. (2004) found that the magnitude of differential activation within the ACC and left insula induced by an empathy-related manipulation was correlated between .52 and .72 with two scales of emotional empathy (the Empathic Concern Scale of Davis, and the Balanced Emotional Empathy Scale of Mehrabian).

Why is a correlation of $r=.88$ with 13 subjects considered "remarkably high"? For starters, it exceeds the reliability of the hemodynamic and behavioral (social, emotional, personality) measurements:

The problem is this: It is a statistical fact... that the strength of the correlation observed between measures A and B reflects not only the strength of the relationship between the traits underlying A and B), but also the reliability of the measures of A and B.

Evidence from the existing literature suggests the test-retest reliability of personality rating scales to be .7-.8 at best, and a reliability no higher than .7 for the **BOLD** (Blood-Oxygen-Level Dependent) signal. If each of these measures was [impossibly] perfect, then the highest possible correlation would be $\sqrt{.8 * .7}$, or .74.

This observation prompted the authors to conduct a [meta-analysis](#) of the literature. They identified 54 papers that met their criteria for fMRI studies reporting correlations between the BOLD response in a particular brain region and some social/emotional/personality measure. In most cases, the Methods sections did not provide enough detail about the statistical procedures used to obtain these correlations. Therefore, a questionnaire was devised and sent to the corresponding authors of all 54 papers:

APPENDIX 1: fMRI Survey Question Text

Would you please be so kind as to answer a few very quick questions about the analysis that produced, i.e., the correlations on page XX. We expect this will just take you a minute or two at most.

To make this as quick as possible, we have framed these as multiple choice questions and listed the more common analysis procedures as options, but if you did something different, we'd be obliged if you would describe what you actually did.

The data plotted reflect the percent signal change or difference in parameter estimates (according to some contrast) of...

1. ...the average of a number of voxels.
2. ...one peak voxel that was most significant according to some functional measure.

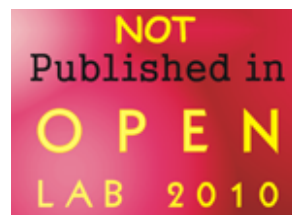


[Subscribe to comments](#)

[Subscribe to The Neurocritic by Email](#)



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 Unported License](#).



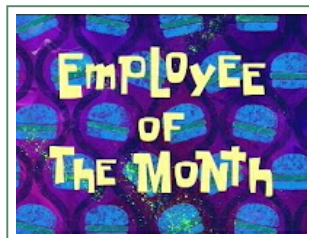
3. ...something else?

etc.....

Thank you very much for giving us this information so that we can describe your study accurately in our review.

They received 51 replies. Did these authors suspect the final product could put some of their publications in such a negative light?

SpongeBob: What if Squidward's right? What if the award is a phony? **Does this mean my whole body of work is meaningless?**



After providing a nice overview of fMRI analysis procedures (beginning on page 6 of the [preprint](#)), Vul et al. present the results of the survey, and then explain the problems associated with the use of non-independent analysis methods.

...23 [papers] reported a correlation between behavior and **one peak voxel**; 29 reported the mean of a number of voxels. ... Of the 45 studies that used functional constraints to choose voxels (either for averaging, or for finding the 'peak' voxel), 10 said they used functional measures defined within a given subject, **28 used the across-subject correlation to find voxels**, and 7 did something else. All of the studies using functional constraints used the same data to select voxels, and then to measure the correlation. **Notably, 54% of the surveyed studies selected voxels based on a correlation with the behavioral individual-differences measure, and then used those same data to compute a correlation within that subset of voxels.**

Therefore, for these 28 papers, voxels were selected *because* they correlated highly with the behavioral measure of interest. Using simulations, Vul et al. demonstrate that this glaring "non-independence error" can produce significant correlations out of noise!

This analysis distorts the results by selecting noise exhibiting the effect being searched for, and any measures obtained from such a non-independent analysis are biased and untrustworthy (for a formal discussion see Vul & Kanwisher, in press, [PDF](#)).

And the problem is magnified in correlations that used activity in *one* peak voxel (out of a grand total of between 40,000 and 500,000 voxels in the entire brain) instead of a cluster of voxels that passed a statistical threshold. Papers that used non-independent analyses were much more likely to report implausibly high correlations, as illustrated in the figure below.



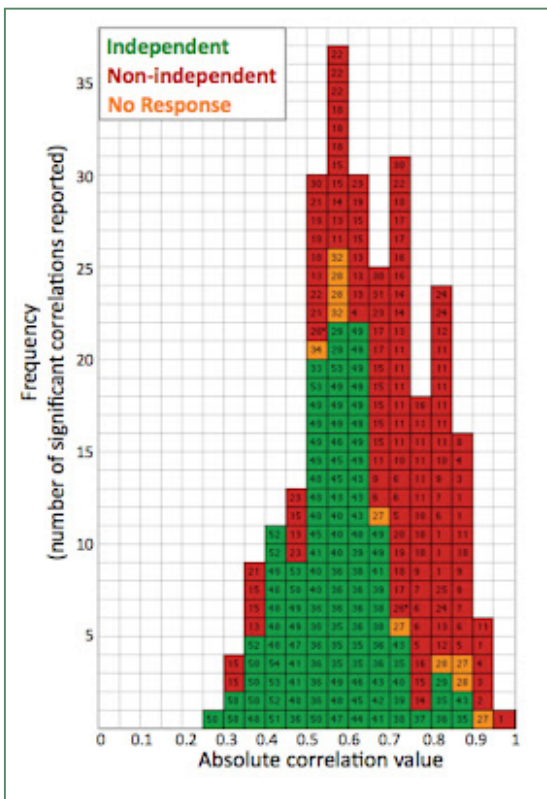


Figure 5 (Vul et al., 2009). The histogram of the correlations values from the studies we surveyed, color-coded by whether or not the article used non-independent analyses. Correlations coded in green correspond to those that were achieved with independent analyses, avoiding the bias described in this paper. However, those in red correspond to the 54% of articles surveyed that reported conducting non-independent analyses – these correlation values are certain to be inflated. Entries in orange arise from papers whose authors chose not to respond to our survey.

Not so coincidentally, some of these same papers have been flagged (or flogged) in this very blog. The Neurocritic's very first post 2.94 yrs ago, [Men are Torturers, Women are Nurturers...](#), complained about the overblown conclusions and misleading press coverage of a particular paper (Singer et al., 2006), as well as its methodology:

And don't get me started on their methodology -- *a priori* regions of interest (ROIs) for pain-related empathy in fronto-insular cortex and anterior cingulate cortex (like the relationship between those brain regions and "pain-related empathy" are well-established!) -- and on their [pink-and-blue color-coded tables](#)!

Not necessarily the most sophisticated deconstruction of analytic techniques, but it was the first...and it did question how the regions of interest were selected. And of course how the data were interpreted and presented in the press.

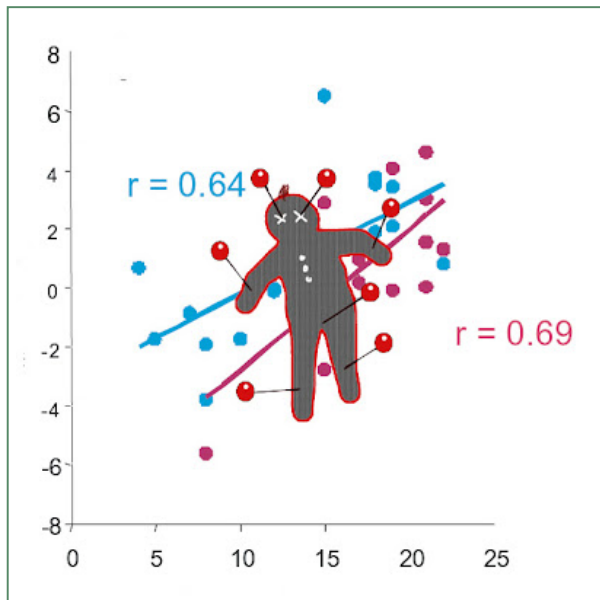
SUMMARY from The Neurocritic : Ummm, it's nice they can generalize from 16 male undergrads to the evolution of sex differences that are universally valid in all societies.

As you can tell, this one *really* bothers me...

And what are the conclusions of Vul et al.?

To sum up, then, we are led to conclude that a disturbingly large, and quite prominent, segment of social neuroscience research is using seriously defective research methods and producing a profusion of numbers that should not be believed.

Finally, they call upon the authors to re-analyze their data and correct the scientific record.



Footnotes

¹ Kanwisher was elected to the prestigious [National Academy of Sciences](#) in 2005.

² The authors note that the problems are probably not unique to neuroimaging papers in this particular subfield, however.

References

Eisenberger NI, Lieberman MD, Williams KD. (2003). [Does rejection hurt? An fMRI study of social exclusion.](#) *Science* 302:290-2.

Singer T, Seymour B, O'Doherty J, Kaube H, Dolan RJ, Frith CD. (2004). [Empathy for pain involves the affective but not sensory components of pain.](#) *Science* 303:1157-62.

Singer T, Seymour B, O'doherty JP, Stephan KE, Dolan RJ, Frith CD. (2006) [Empathic neural responses are modulated by the perceived fairness of others.](#) *Nature* 439:466-9.

Edward Vul, Christine Harris, Piotr Winkielman, & Harold Pashler (2009). Voodoo Correlations in Social Neuroscience. *Perspectives on Psychological Science*, in press. [PDF](#)

Vul E, Kanwisher N. (in press). Begging the question: The non-independence error in fMRI data analysis. *To appear in Hanson, S. & Bunzl, M (Eds.), Foundations and Philosophy for Neuroimaging.* [PDF](#)

posted by The Neurocritic @ 5:30 AM

[28 comments](#) 

28 Comments:

At [January 06, 2009 1:32 AM](#),  [Ahmed Aldebrn Fasih](#) said...

Fantastic review article! My own signal processing involves physical randomness and it *is* hard work to make sure one's statistics is honest, but the examples provided here (choosing the peak voxel based on its correlation to the behavior desired!) are for me an all-time low in mechanistic, unreflective application of statistics.

But I anticipate news of even worse misuses of statistics will surely come to light. I'm waiting for a paper like this to consider the many new cancer-carcinogen studies which guide policy. I also heard of data presented in the 2006 AAAS meeting that showed that 80% of epidemiological studies failed to replicate (these also guide policy).


At [January 06, 2009 9:12 PM](#),  [DRCONTRARIAN](#) said...

Completely off-topic, but since I don't 'tweet', I have taken the 'old-fashioned' option of commenting on a few of my favourite blogs (and you are on the list :-)) to engage in discussion about the question: what will change the world? The question was asked and answered by over 100 of the brightest on Edge.org (it is not a blog) so I have summarised their views on my blog. One answer is 'social media literacy' - and I wonder what you (and your readers) think?

I think you might enjoy 92, 93 and 94 amongst others

At [January 07, 2009 12:11 AM](#),  [The Neurocritic](#) said...

Aldebrn,

As mentioned by another [commenter](#), there's the 2005 *PLoS Medicine* article by John P.A. Ioannidis, [Why Most Published Research Findings Are False](#) .

But then the debunker gets debunked here, [ASSESSING THE UNRELIABILITY OF THE MEDICAL LITERATURE: A RESPONSE TO "WHY MOST PUBLISHED RESEARCH FINDINGS ARE FALSE"](#).

At [January 08, 2009 2:19 AM](#),  Anonymous said...

How often do the statistics actually back the conclusions in the research I read?

Very seldom.

Stat. analysis is not being taught well, I would guess.

I agree with Aldebrn here- this is one of many coming.

(I see a problem in science in general- the idea is to make a hypothesis and then find evidence to support it. I see this in biology and astrophysics and neuroscience especially. Good science is an attempt to disprove the current theory. My two bits...)

Sonic

At [January 08, 2009 10:28 AM](#),  Anonymous said...

There are two major problems with your (and Vul et al's) discussion of reliability and the impact on correlations.

First, reliability does not refer to measures but to scores on measures. Thus, you cannot say that a measure has a reliability of X. Rather scores on the measure (i.e., in a particular setting) have a reliability of X. It is thus not very meaningful to apply reliabilities observed in other contexts to a new situation since the reliability of scores on the measure could be significantly higher in a new situation. Second (and more importantly), observed unreliability does not place an absolute constraint on observed correlations - which is why correlations corrected for unreliability are occasionally greater than 1. This (as has been pointed out by the likes of Murphy, 2003) is typically due to underestimates of reliability. Cronbach's alpha is, for example, typically a lower bound on the reliability of a set of items. It is thus quite possible to observe a correlation that is larger than the square root of the product of two reliability estimates. This is contrary to your statement that the correlation between (scores on) two measures with reliabilities of .8 and .7 cannot be greater than .74. I am surprised that a journal as good as *Perspectives* has allowed these rather elementary errors to slip through.

At [January 08, 2009 11:28 AM](#),  [The Neurocritic](#) said...

Thanks for commenting.

First, I'm aware that the scores or hemodynamic measures in question are taken on the same day, and that test-retest reliability applies to scores or measures taken at different times. I'm sure Vul et al. know this as well. However, the personality/emotionality scores and the hemodynamic measures are typically not obtained at the same time.

Furthermore, even if one accepts that the observed correlation can be greater than .74, that doesn't invalidate the fact that 54% of the surveyed papers used non-independent analysis methods, which inflated the correlation values.

Although your comment will have a reasonably sized audience here, you might want address the authors directly in a letter to the editor of *Perspectives on Psychological Science*...

At [January 09, 2009 9:32 PM](#),  Anonymous said...

Anonymous who raises concerns about the definition of reliability etc.: it seems that Edward Vul has actually posted responses to your queries here:

<http://www.edvul.com/voodoo CORR.php>

At [January 09, 2009 10:47 PM](#),  [The Neurocritic](#) said...

Thanks to the latest Anonymous Commenter for pointing out Ed Vul's more proficient answers to the questions from the January 08, 2009 10:28 AM [Anonymous](#).

[Supplementary Q and A](#)

Since some interesting questions have been raised about our paper (in these blogs and otherwise), we'll do our best to address them here.

Q: Since reliabilities apply to scores on measures rather than the measures themselves, how can you use reliabilities from other samples to make inferences about the scores used in the particular

studies you describe?

A: Like nearly all social scientists, we assume that the reliability of a measure estimated from scores obtained on one sample will generalize to other samples. It is true that these measures of reliability will vary from sample to sample, but this is true of any measure ever obtained from a sample population. We hope (and we assume the authors of the articles do too) that the participants sampled into the reported studies were representative (with respect to the inferences in question). If they are, then we have no reason to suspect that the reliability of scores on any of the measures in these samples will differ substantially from those of other samples that have been previously used to evaluate these measures.

Q: Does reliability put an absolute constraint on the correlation that may be obtained?

A: No, reliability puts a constraint on the expected value of the correlation. Noise may make the correlation higher sometimes, and lower at other times. We argue that these articles have selected favorable noise that increases the apparent correlation, thus causing these estimates to systematically exceed the maximum possible expected value of a correlation between these measures. We should reiterate again that we think that the theoretical upper bound on the expected correlation to be much higher than what should reasonable be expected. The upper bound assumes a perfect underlying correlation.

At [January 10, 2009 11:44 AM](#),  Anonymous said...

It's unfortunate that the blog world is getting all excited by this. This is hardly a new issue, nor one endemic to "social neuroscience". Non-independence is certainly an issue if you take the correlation results as being more than they are (and I'm sure more than a few authors are happy to have you think that).

However, the underlying analysis is sound. In a nutshell what most of the studies criticized in this paper have done is this. They've run a whole-brain regression of brain activity against a psychological variable. They then take the resulting map and apply a statistical correction for multiple comparisons (using whatever method they prefer. FDR, RFT, or a value based on monte-carlo simulations etc.) they usually apply a cluster extent threshold such that for clusters must be composed of 10 contiguous voxels to be considered significant (this helps get rid of spurious findings from single-voxels which have a stronger likelihood of being due to chance). *So far this is identical to the contrasting of conditions that is the standard analysis in nearly all of published fmri studies.*

What is making everyone all flustered is that the researchers then use regions of interest based on this whole-brain regression map to create scatterplots of signal change VS a psychological variable. Yes, these scatterplots will necessarily be significant since the previous analysis said they were. But if you wanted to be a convinced of a correlation, what would you rather see? A cluster and a t-value, or a scatterplot so that you can see yourself whether the correlation is sound or being driver by outliers?

Perhaps if authors wrote "for illustrative purposes we also show a

scatterplot of the observed correlation" people would relax a little. But is that really necessary? Most imagers understand that this is the value of the scatterplots.

This issue also comes up with traditional contrast methods. For example, conditions A and B are contrasted. Regions show up as more active in $A > B$. Then signal change is extracted from ROIs based on the peak voxels showing maximum difference between A and B and those signal change values are plotted. The resulting bar graphs are also suffer from non-independence, however they are informative as they tell you the direction of the effect relative to baseline. Now not everyone buys the idea of a resting baseline consisting of a fixation cross, but that's a separate argument.

Finally, let's not get too carried away with functional localizers. They're only as good as the localizer you use. And they can lead one to ignore potentially interesting effects that occurred outside of what these localizers activated.

At the moment this paper has generated a witch hunt among bloggers. If this reaches reviewers it's going to be a pain for years to come. Perhaps conventions need to be changed such that people are more explicit in noting that the scatterplots are not-independent. But non-independence doesn't invalidate the original regression analysis.

At [January 10, 2009 11:52 AM](#),  Anonymous said...

Thanks for your reply to the concerns I raised. I love the paper but the psychometrician in me remains (perhaps pedantically) hung up on your statement in your paper (p. 2-3) that

"Thus, the reliabilities of two measures provide an upper bound on the possible correlation that can be observed between the two measures (Nunnally, 1970)."

As I noted earlier - reliabilities refer to scores (not measures), correlations are observed between scores on measures (not measures), and score reliability does not provide an upper bound on possible correlations.

As I said - perhaps pedantic but I could not help hearing the disapproving voice of my old psychometrics professor as I read that sentence. Congratulations on a wonderful paper nonetheless!

At [January 11, 2009 12:26 AM](#),  The Neurocritic said...

vtw - thanks for your comments, but did you actually read the paper? Did you look at **Figure 4: A simulation of a non-independent analysis on pure noise?**

In addition, Vul et al. noted that:

38% of our respondents who reported the correlation of the peak voxel (the voxel with the highest observed correlation) rather than the average of all voxels in a cluster passing some threshold.

And why do you say that bloggers are conducting a witch hunt? The manuscript has been accepted for publication in [Perspectives on Psychological Science](#), so of course reviewers will see it.

At [January 11, 2009 12:36 AM](#),  The Neurocritic said...

To the Anonymous of January 10, 2009 11:52 AM,

I'm sure Vul et al. very much appreciate your interest in their paper. And at least you admit you were being pedantic... ;-)

At [January 11, 2009 9:08 AM](#),  Anonymous said...

Hey Neurocritic.

Yes I read the paper closely. It's quite good and timely, there only two issues I disagree with. Many of the simulations that the authors use to build their argument do not directly compare to the way imagers actually run the analysis (appendix 2 is about signal inflation not the possibility of false positives). More on this in a minute. Secondly, the authors make the very strong claim that results of many of the studies might be false. Once again, the underlying regression analysis is no less sound than doing whole-brain GLMs for contrasting conditions. The same possibility of finding spurious results by chance still applies. Which is true of statistics in general, not just neuroimaging. Hence the convention of $p < 0.05$.

Back to the simulations. These are very illuminating, but in my mind the simulation the truly matches that way that imagers do neuroimaging is not in the paper. As I'm certain you and many readers here know, imagers use a cluster extent threshold to ensure that the analysis reveals clusters rather than single voxels. So to put the simulations back in terms of, for instance, the author's stock exchange example. The problem then is not that the weather station readings will correlate with some measures of the stock exchange by pure chance. Rather they should posit it as what are the odds that the weather station readings correlate with 10 stocks which are ordered consecutively on some list. The authors do address this in point F (page 18) using alphasim (which I also use). And they are correct, most papers have a tendency to chose a threshold that is more liberal than FWE $p < 0.05$. One of the conventions in many studies is $p < 0.001$ and $K=10$. This isn't true of all, but it's true of many and is generally accepted. It's one way of dealing with the very real fact that in imaging doing true FWE corrections leads to far too conservative of a threshold. The alternative is to bump up the cluster size. But bump up too high and you lose the possibility of detecting many smaller structures (amygdala comes to mind).

As for 38% of the papers using the peak voxel. It doesn't really matter what they use, peak or ROI. It is a cartoon. And in this I agree with the paper. The r-values are inflated. The paper makes this point excellently.

The problem is that readers in the blogosphere are taking this to mean that a large portion of the papers using correlations in imaging are reporting nonsense. Which simply isn't the case.

That many of these papers got by based on large r-values. I have no doubt. And so yes this paper is important for demonstrating this problem. But once again, the underlying whole-brain regressions are sound. You have people on some blogs (I followed the links from the authors website) saying things like "all fmri is bogus". And for the next while there's going to be a stigma attached to doing whole-brain regressions with every reviewer and their mothers asking about non-independence, whether or not it applies. I suppose this can be

construed as a good thing. But we all know reviewing is a bit of a voodoo science itself. All it takes is one reviewer who doesn't understand the finer points of this argument to kill a paper.

Anyhow, in the end I agree with the authors, some form of split half or leave-one-out cross-validation is probably the best way to go if you want to report r-values. But strongly disagree that there is something intrinsically wrong with running a whole-brain regression. It's only when you try to express this is an r-value that non-independence creeps in.

Whew, all that being said. Great blog post as usual!

At [January 12, 2009 7:03 AM](#),  Anonymous said...

For a reply by some of the criticized authors read <http://www.bcn-nic.nl/replyVul.pdf>

At [January 13, 2009 10:31 AM](#),  [The Neurocritic](#) said...

Interested parties can read [Voodoo Counterpoint](#) for an excerpt of that rebuttal by Jabbi, Keysers, Singer, and Stephan ([entire PDF](#)).

At [January 27, 2009 11:48 AM](#),  Anonymous said...

[http://www.scn.ucla.edu/pdf/LiebermanBerkmanWager\(invitedreply\).pdf](http://www.scn.ucla.edu/pdf/LiebermanBerkmanWager(invitedreply).pdf)

At [January 27, 2009 1:44 PM](#),  Anonymous said...

I have been following the debate with interest. I just read the reply by Lieberman and colleagues at [http://www.scn.ucla.edu/pdf/LiebermanBerkmanWager\(invitedreply\).pdf](http://www.scn.ucla.edu/pdf/LiebermanBerkmanWager(invitedreply).pdf). They point out that NONE of the "red list" authors they were able to contact (which appeared to be most of them) said they conducted their analysis in the way it was portrayed by the Vul paper. This raises concern that Vul and colleagues may have inadvertently misrepresented the full picture and created a "straw man" to attack (probably unintentionally). Since much of the Vul argument is based on self-reported responses to a handful of survey items, it might be premature to "skewer" the entire field of social neuroscience. Has anyone considered the possibility that the wording or presentation of the Vul survey may itself have introduced some bias or error in their data? The survey items about analysis strategy do seem at least a little cryptic and open to interpretation. It would be a shame to prematurely tarnish years of work, professional reputations, and countless millions of dollars of public research funding based on a couple of answers to a few potentially ambiguous and unvalidated survey items. Just a thought.

At [January 28, 2009 1:35 AM](#),  [The Neurocritic](#) said...

Anonymous LiebermanBerkmanWager - The pointer to your rebuttal is much appreciated. I did link to it in a [new post](#).

Tom - You made a number of important points, thanks for commenting.

At [February 03, 2009 12:28 PM](#),  Anonymous said...

It is interesting to me that Vul et al. imply incompetence and/or fraudulence in fMRI data analyses when their own review of the literature was so clearly biased - leaving out most of the findings that DON'T support their findings.

Kanwisher (Vul's MIT advisor) is well known as a media attention seeking person who appears more interested in making a newsworthy splash than in checking facts. The procedures they are critiquing are almost never used by researchers, and certainly not by most of the researchers they cited (selectively). Incompetence and fraudulence run rampant apparently.

Vul et al. also attack a single area of research (i.e., social neuroscience), when in fact all of neuroscience is susceptible to their criticism...if they were accurate. It suggests some a priori reason to dislike and attack the field, not the method.

I believe that this is an unfortunate attempts to gain notoriety by appealing to those who already hold negative attitude toward social neuroscience, or neuroscience more broadly, or (sadly) science in general.

I implore posters agreeing with Vul et al. to ask themselves if they truly know enough to be agreeing. Have you performed extensive fMRI analyses? Have you reviewed the studies Vul et al. cited? Please do not dismiss some reasearch as "unscientific" while simultaneously failing to engage in critical thinking yourselves.

Cheers!

At [February 03, 2009 5:29 PM](#),  [The Neurocritic](#) said...

To the Anonymous of February 03, 2009 12:28 PM,

Nancy Kanwisher is not an author on the Voodoo paper, so your personal attack on her is irrelevant here.

The rebuttals to Vul et al. have emphasized that the latter's analytic objections are by no means unique to social neuroscience, so you are not alone in your defensiveness. Vul et al. acknowledged the generality of their critique, albeit not in a prominent way. Furthermore, note that two of the authors are social psychologists, one of whom has published both fMRI and EMG papers on the topic.

At [February 12, 2009 1:44 PM](#),  Anonymous said...

The Vul paper made an important point about inflated p values that can emerge in neuroimaging research. Unfortunately, the tone of the paper, the method by which they misled researchers to participate, the extent to which they misrepresent findings (i.e., they cherry-picked within the papers they presented), and the attack on a narrow, emerging field when the problem they describe has nothing to do with that field, is shoddy scholarship and will undermine an otherwise important point.

At [March 08, 2009 5:31 PM](#),  Anonymous said...

I'll let interested readers read the array of counterpoints to Vul for themselves (I'd first recommend those of Matt Lieberman et al, see above posts). Vul's article has been thoroughly debunked.

I've been working in fMRI labs for about 3.5 years, and I can say with confidence that any imager with a modest knowledge of standard analysis techniques was able to recognize the faulty reasoning in Vul's criticism. (The extent to which his meta-analysis was sloppily executed

did, however, take me quite by surprise when I read Lieberman's rebuttal.)

What I find most disappointing is the vigor with which non-imagers (and the media at large) blindly accepted Vul's criticism. It is ironic that this group of unquestioning individuals was in fact victim to the exact sort of naive acceptance of published claims that Vul attempted to attack.

At [March 20, 2009 10:50 AM](#),  Anonymous said...

UCLA's ..Dr. Matthew Lieberman and Dr. Naomi Eisenberger's work must be thoroughly investigated! They are guilty of performing unethical and harmful social cognitive research experiments on unconsenting human subjects!

In order for Dr. Matthew Lieberman and Dr. Naomi Eisenberger to get the results for their Social Cognitive Research experiments involving long term emotional and physical pain ...abandonment from loved ones...and being socially rejected,.....they teamed up with a small group of individuals that devised a horrible plan to ruin and destroy innocent ,unconsenting and unknowledgeable patients social environments with systematical harassment!

The Systematical Harassment was created in a way to, deliberately, stir up enough problems with in a person's social life...that they would soon find themselves suffering from the very same social cognitive psychology problems that these two social psychology Doctors just happened to be investigating!!!!

Can anyone tell me , who should be contacted , so...that a deeper investigation into their unethical research methods can begin? I would sure appreciate it!

And, Thank You..So Much! For writing this paper!!!

At [March 31, 2009 11:50 AM](#),  Anonymous said...

The issue of positively biased statistics after voxel selection is quite valid, but also quite well known among neuroimagers. Ed Vul performed a haphazard "meta-analysis" with cherry-picked data, and falsely accused a number of authors of scientific malfeasance. There have been three significant rebuttals written: Jabbi, Lieberman, and Nichols. Vul has replied (ratherly weakly) to the Jabbi rebuttal, but has so far ignored the others which contain a much stronger and better supported condemnation of his paper.

It is indeed disappointing that so many have jumped on the Ed Vul bandwagon without first acquainting themselves with the standards in the field, or with the work reported in the "accused" papers.

At [April 08, 2009 11:50 PM](#),  [The Neurocritic](#) said...

Anonymous of March 31, 2009 11:50 AM,

You might want to read Vul et al.'s rebuttal ([PDF](#)) to Lieberman et al. ([PDF](#)) and Nichols et al. ([PDF](#)). While you're at it, you can take a gander at the other articles in [PERSPECTIVES ON PSYCHOLOGICAL SCIENCE](#), Vol. 4, Issue No. 3 (May 2009), including those by statisticians who support the arguments of Vul et al.

At [March 24, 2010 1:24 PM](#),  [Matthew Lieberman](#) said...

For anyone interested, there was a public debate on Voodoo Correlations last fall at the Society of Experimental Social Psychologists between Piotr Winkielman (one of the authors on the Voodoo paper) and myself (Matt Lieberman). The debate has been posted online.

<http://www.scn.ucla.edu/Voodoo&Typell.html>

At [March 25, 2010 1:06 PM](#),  [The Neurocritic](#) said...

Matt - Thanks for providing the link. I posted the videos here:

[Voodoo and Type II: Debate between Piotr Winkielman and Matt Lieberman](#)

At [March 15, 2016 5:06 PM](#),  Anonymous said...

Just curious, did any of the accused authors re do their analysis?

[Post a Comment](#)

[<< Home](#)

