

Reply to Jern et al.: On Asking the Right Questions

R.I.M. Dunbar

Department of Experimental Psychology

University of Oxford

Old Road Campus

Oxford OX3 7LG

Text length: 480 words

We are delighted that Jern and colleagues (1) share our interest in the psychopharmacological underpinnings of social behaviour (2). It is a topic of growing interest and major significance. It is, however, important to correct some misunderstandings. First, we fully appreciate the concern about type I errors in multiple comparisons; it has been widely discussed in the ‘genome-wide’ genetics literature (3-6), and we addressed it at length in our *SI* (2). However, it is also widely recognised most correction procedures unhelpfully inflate type II errors (7-10). We followed professional advice and current best practice in genome-wide research (3-6). Ironically, had we reported a study on just one neurochemical¹, no one would have quibbled. The results do not suddenly become non-significant simply because we

¹ In our paper, we used the term *neuroendocrine* to cover the six neuro-peptides/amines/steroids we considered. Biochemically, of course, they have different properties, although from the social point of view they are all equally functional. To avoid confusion, I use the term *neurochemical* here.

considered several neurochemicals and their interactions simultaneously. Second, although Jern et al. obviously wouldn't be aware of this, we have two smaller independent follow-up samples (currently under review) that show the same pattern. Our results *are* statistically robust. Third, more importantly, the substantive issue that we address in our paper is the fact that all studies on this topic to date invariably focus on just one social neurochemical (usually oxytocin) without making any attempt to check whether the reported effects might be confounded by the others. Only three of the six we consider actually play substantive roles, and that in itself weakens concerns about multiple comparisons: they are not *all* significant. Moreover, the strong signature for dopamine was completely unexpected and alone merits attention. Ours is a more nuanced story. Fourth, Jern et al. (1) question our contingency table analysis. Jern et al. test one hypothesis (essentially that dealt with by our main analysis), but we test a rather different one: given that some of the associations between SNPs and social domains are stronger than might be expected, how likely is it that these 'significant' effects are randomly distributed across the three social domains? That is a perfectly legitimate use of the χ^2 test. We used conventional significance values simply as a convenient criterion for distinguishing potentially interesting associations from the rest and, in this context, need make no particular claim about their individual statistical significances as such. This is an issue about sampling grain that has been a longstanding problem in quantitative ecology (11). We could have placed our criterion anywhere on the p-value continuum. The tradeoff is between samples that are too coarse-grained or too fine-grained to show anything at all and something in between; the proper question is whether is there any level of grain in between where something interesting emerges. We show that there is. We rest our case.

References

1. Jern, P., Verweij, K.J.H., Barlow, F.K. & Zietsch, B.P. (2017). Reported associations between receptor genes and human sociality are explained by methodological errors, and do not replicate. *Proceedings of the National Academy of Sciences, USA*.
2. Pearce, E., Wlodarski, R., Machin, A. & Dunbar, R.I.M. (2017). Variation in the β -endorphin, oxytocin, and dopamine receptor genes is associated with different dimensions of human sociality. *PNAS* doi: 10.1073/pnas.1700712114
3. Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. *Proc Natl Acad Sci USA* 100:9440-9445.
4. Johnson, R.C. et al., 2010. Accounting for multiple comparisons in a genome-wide association study (GWAS). *BMC Genomics* 11: 724.
5. Rice TK, Schork NJ, Rao DC (2008) Methods for Handling Multiple Testing. *Genetic Dissection of Complex Traits*, ed Genetics BT-A in (Academic Press), pp 293–308.
6. Singer JB (2009) Candidate gene association analysis. *Cardiovascular Genomics: Methods and Protocols*, ed DiPetrillo K (Humana Press, Totowa, NJ), pp 223–230.
7. Perneger, T.V. (1998). What's wrong with Bonferroni adjustments. *British Medical Journal* 316: 1236.
8. García, L.V. (2004). Escaping the Bonferroni iron claw in ecological studies. *Oikos* 105: 657-663.
9. Garamszegi, L.Z. (2006). Comparing effect sizes across variables: generalization without the need for Bonferroni correction. *Behavioral Ecology* 17: 682-687.
10. Narum, S.R. (2006). Beyond Bonferroni: less conservative analyses for conservation genetics. *Conservation Genetics* 7: 783-787.
11. Greig-Smith, P. (1983). *Quantitative Plant Ecology*. Oxford: Wiley-Blackwell.