

Title: Facts and findings: A reply to Powell & Nettelbeck (2014)

Authors: Sophie von Stumm (Department of Psychology, Goldsmiths University of London, New Cross, SE14 6NW, UK), Benedikt Hell (Hochschule für Angewandte Psychologie, Institut Mensch in komplexen Systemen, Fachhochschule Nordwestschweiz), Riggbachstrasse 16, 4600 Olten), & Tomas Chamorro-Premuzic (University College London (UCL), 26 Bedford Way, London, WC1H 0AP, UK)

Abstract: In their recent article, [Powell, C., & Nettelbeck, T. (2014). Intellectual curiosity may not incrementally predict academic success. *Personality and Individual Differences*, 64, 7-11], Powell and Nettelbeck question the predictive validity of intellectual curiosity for academic performance. We discuss here three issues we found with their report.

Main text: In their recent article, [Powell, C., & Nettelbeck, T. (2014). Intellectual curiosity may not incrementally predict academic success. *Personality and Individual Differences*, 64, 7-11], Powell and Nettelbeck aspired to replicate, extend and question our previous findings [von Stumm, S., Hell, B., & Chamorro-Premuzic, T. (2011). The hungry mind: Intellectual curiosity as third pillar of academic performance. *Perspectives on Psychological Science*, 6, 574-588]. We are delighted that our work has prompted other researchers to investigate the relationship between intellectual curiosity and academic performance, and we welcome theoretical and empirical advances in this field. However, we found Powell and Nettelbeck's study highly problematic for three reasons, which we outline below.

(1) Misinterpreting previous research

Powell and Nettelbeck interchangeably referred to their study's outcome variable as "school grades" (in abstract), "academic success" (in the title and text), and "academic achievement" (in measures), which was actually assessed by self-reported Tertiary Entry Rank (TER) or Australian Tertiary Admission Rank (ATAR) that were obtained from participants 6 months to 14 years after the actual performance had taken place. Although subjective achievement data are often biased (e.g. Dunning, Johnson, Ehrlinger, & Krueger, 2003; Freund & Kasten, 2013; von Stumm, 2013), and although differences in the time delay between actual

performance and its recall are also likely to affect the results, neither issue was acknowledged by the Powell and Nettelbeck's study title and text. However, what truly concerned us was Powell and Nettelbeck's suggestion that *we*, in our previous work, had similar data to assess academic performance. The authors stated that "[a] possible weakness of [our] study is that TER/ ATAR reflected academic achievement¹ from six months to several years ago. However, these circumstances appear to have similar to those applying to the three studies included in the meta-analysis made by von Stumm et al. (2011), where participants were also mostly undergraduates of about the same ages of students participating here." (p. 10). It is true that the studies included in our meta-analysis reported data from high school and university students. It is, however, false that academic performance was retrospectively self-reported in these studies: two studies included university GPA taken from official university records (Chamorro-Premuzic, Furnham, & Ackerman, 2006; Goff & Ackerman, 1992), and one study cited school grades that were also taken from official school records (Wilhelm, Schulze, Schmiedek, & Suess, 2003). Across all three studies, academic achievement data marked either prospective or contemporaneous performances, not ones that had taken place 6 months to 14 years ago.

(2) Mislabelling study variables

Powell and Nettelbeck referred to "Openness to Ideas" (p. 7), which they noted as a potentially relevant variable in the context of this study, and which they described as one of the six facets of the Openness to Experience measure from the NEO-PI-R (Costa & McCrae, 1992). However in the manuscript's measure section, a different construct, namely Goldberg's (1999) Intellect scale from the IPIP - which had not been mentioned before - was described as the measure that was actually administered in and analyzed for this study. While IPIP-Intellect and NEO-PI-R-Ideas are likely to be highly correlated, they are both scales in their own right and passing one off as the other is confusing to the readership. Furthermore, the exchange of scales diminishes the informative value of the reported factor analyses: the overlap of the NEO-PI-R-Ideas facet with other measures of intellectual curiosity has recently sparked much research interest (e.g. Mussel, 2010; Fleischhauer, Enge, Brocke, Ullrich, Strobel, & Strobel, 2010), while this is less true for Intellect (cf. von Stumm & Deary, 2013).

¹ This sentence is inaccurate in itself: TER/ ATAR data in Powell and Nettelbeck's study reflected *perceived* not actual academic achievement from six months to several years ago.

(3) Misreporting statistics

The authors reported two inaccurate observations about their data. For one, they stated that there was “no sex difference for any variable ($p < .10$)” (p. 9). However, the statistics in table 1 reveal a significant difference in TER/ ATAR between men and women. For the other, the authors stated that “means of the four intellectual curiosity measures were similar to those reported for university populations in previous studies (Goff & Ackerman, 1992; Litman, 2008)” (p. 9). We have checked both references, which reported significantly lower means for Typical Intellectual Engagement and Epistemic Curiosity than described in the current paper (see table 1). While both these reporting errors may result from accidental oversights, they challenge the adequacy of the employed statistical methods, the comparability of data, and the validity of findings.

Concluding remarks

In summary, we find Powell and Nettelbeck’s conclusions untenable. That said, their study offers an opportunity for an important lesson to be drawn. Replications and extensions of previous research findings, including our own, are pivotal for the advancement of science and the accumulation of knowledge and thus, they are often generally encouraged. While we principally agree with such encouragements, we wish to caution here that poor replication studies make negligible contributions to science. To avoid these in the future, we suggest improving the effectiveness of the peer-review process², in this journal as well as elsewhere.

References

- Chamorro-Premuzic, T., Furnham, A., & Ackerman, P. L. (2006). Incremental validity of typical intellectual engagement as predictor of different academic performance measures. *Journal of Personality Assessment*, 87, 261-268.
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI): Professional manual*. Odessa, FL: Psychological Assessment Resources.

² None of us had been invited to review Powell and Nettelbeck’s manuscript before publication.

- Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, 12, 83-86.
<http://dx.doi.org/10.1111/1467-8721.01235>
- Freund, P. A., & Kasten, N. (2012). How smart do you think you are? A meta-analysis on the validity of self-estimates of cognitive ability. *Psychological Bulletin*, 138(2), 296-321.
- Fleischhauer, M., Enge, S., Brocke, B., Ullrich, J., Strobel, A. & Strobel, A. (2010). Same or different? Clarifying the relationship of Need for Cognition to personality and intelligence. *Personality and Social Psychology Bulletin*, 36, 82-96.
- Goff, M., & Ackermann, P. L. (1992). Personality-Intelligence relations: assessment of typical intellectual engagement. *Journal of Educational Psychology*, 84, 537-552.
- Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality Psychology in Europe*, Vol. 7 (pp. 7-28). Tilburg, The Netherlands: Tilburg University Press.
- Mussel (2010). Epistemic curiosity and related constructs: Lacking evidence of discriminant validity. *Personality and Individual Differences*, 49, 506-510
- von Stumm, S. (2013). Intelligence, gender, and assessment method affect the accuracy of self-estimated intelligence. *British Journal of Psychology*: doi: 10.1111/bjop.12031
- von Stumm, S. & Deary, I. J. (2013). Intellect and Cognitive Performance in the Lothian Birth Cohort 1936. *Psychology & Aging*, 28, 680-684.
- Wilhelm, O., Schulze, R., Schmiedek, F., & Süß, H.-M. (2003). Interindividuelle Unterschiede im typischen intellektuellen Engagement [Individual differences in typical intellectual engagement]. *Diagnostica*, 49, 49-60.