A Quartet of Interactions

Eric-Jan Wagenmakers
University of Amsterdam
Correspondence concerning this article should be addressed to:
Eric-Jan Wagenmakers
University of Amsterdam, Department of Psychology
Weesperplein 4
1018 XA Amsterdam, The Netherlands
E-mail may be sent to EJ.Wagenmakers@gmail.com.

When people practice specific tasks –such as playing video games or speaking a second language– does this boost their general cognitive abilities? To what extent are our basic cognitive processes malleable? Are we to accept the hand that nature has dealt us –a short attention span, a fallible memory, a limited perceptual system– or is there hope that, with the right training and dogged persistence, we can grease the cogwheels of our cognitive machinery, improving performance beyond its apparent boundaries? These questions lie at the heart of a nature versus nurture dilemma that has recently manifested itself in at least three different paradigms. In the "brain training" paradigm, researchers study whether training on specific memory tasks results in transfer and helps increase the overall effectiveness of human working memory (e.g., Jaeggi, Buschkuehl, Jonides, & Shah, 2011; but see Shipstead, Redick, & Engle, 2012); in the "video gaming" paradigm, researchers study whether playing violent video games leads to improvements in visual perception (e.g., Bejjanki et al., 2014; but see van Ravenzwaaij, Boekel, Forstmann, Ratcliff, & Wagenmakers, 2014); finally, in the "bilingual" paradigm, researchers study whether people who speak a second language display better executive functioning – for instance, bilinguals are presumed to be relatively good in suppressing irrelevant information (e.g., Bialystok, Craik, & Luk, 2012; but see Hilchey, Saint-Aubin, & Klein, in press).

All three paradigms face similar challenges: they require an enormous investment of resources –training people in memory tasks or video games, finding a matched set of bilingual and monolingual participants–, they present a myriad of ways in which a positive result can arise from confounds or contaminations, and they are beset by packs of skeptics and detractors who argue that the positive results are entirely spurious.

The article under discussion represents the most recent volley on the part of the skeptics (Paap, Johnson, & Sawi, in press). The arguments by Paap et al. (in press) are

---

This work was supported by an ERC grant from the European Research Council. Correspondence concerning this article may be addressed to Eric-Jan Wagenmakers, University of Amsterdam, Department of Psychology, Weesperplein 4, 1018 XA Amsterdam, the Netherlands. Email address: EJ.Wagenmakers@gmail.com.

compelling, and even the most dispassionate reader will feel depressed at the suggestion that the collective research effort on bilingual advantages in executive functioning has been a waste of time, effort, and resources. The presence of such research waste has profound negative ramifications that extend well beyond the research topic at hand. The integrity and intellectual health of our discipline requires that sham research paradigms are identified and eliminated as quickly as possible. Luckily there are many ways in which this can be accomplished, for instance through preregistration, data sharing, increasing statistical power, use of Bayesian inference, and engaging in adversarial collaborations (see Chambers, 2013; Matzke et al., 2015; Nosek et al., 2015 and references therein).

Here I emphasize another antidote against the self-deception inherent to sham research paradigms: always plot the data. Indeed, "Graphs are essential to good statistical analysis" (Anscombe, 1973, p. 17). In the days of Anscombe, plots could be obtained only with "trouble, cunning and a fighting spirit" (Anscombe, 1973, p. 21), but this situation has changed and nowadays plots are easy to produce using any statistical software package. The importance of visual data inspection is often demonstrated by Anscombe's quartet, a collection of four X-Y data sets that each yield a Pearson correlation coefficient of $r = .816$, even though they allow very different conclusions about the degree of dependence between X and Y.

Here I constructed a similar quartet, but this time for interactions instead of correlations. Figure 1 shows hypothetical high-N results from a typical 2×2 factorial design with conditions C_1 and C_2 and two different groups. For concreteness, C_1 may index the congruent condition, and C_2 the incongruent condition; mean response time (RT) is on the y-axis; the dots designate the bilinguals whereas the Xs refer to the monolinguals. In all panels, the monolinguals show a 200 ms congruency effect, an effect twice as large as that displayed by the bilinguals. It is tempting to conclude that such data provide support for the bilingual advantage theory of executive functioning.

Indeed, this conclusion is warranted for the data from the top-left panel, where performance for the congruent condition is the same for both groups. The interaction shown in the top-right panel, however, can be undone by a monotonic transformation (Loftus, 1978; Wagenmakers, Krypotos, Criss, & Iverson, 2012), meaning that its existence is tied to the measurement scale – the interaction may be present for RT but absent for speed (i.e., 1/RT). The interaction in the bottom-left panel is driven in part by an unexplained performance difference for the congruent condition. Nevertheless, if the bilingual advantage for the incongruent condition is statistically compelling, then this pattern does provide support for the theory. The results from the bottom-right panel, in contrast, are inconsistent with the bilingual advantage theory because they originate solely from an unexplained performance difference for the congruent condition. Interactions more exotic than the ones displayed in Figure 1 provide even less support for the bilingual advantage theory, and raise more questions than they answer (e.g., Hilchey et al., in press; Paap et al., in press; Redick & Webster, 2014; Redick, 2015).

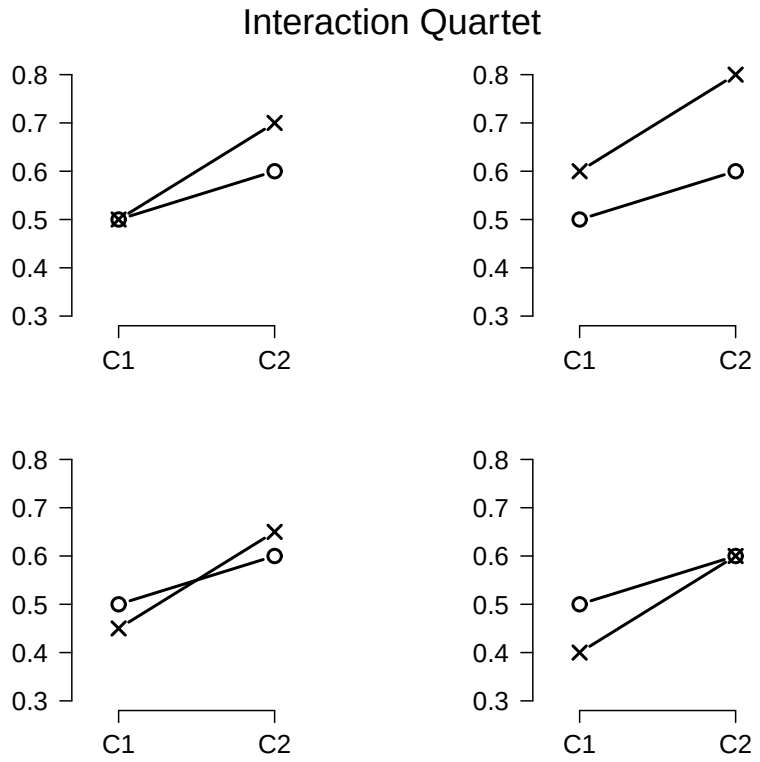

Figure 1. A quartet of interactions, consisting of four hypothetical 2×2 factorial data sets. In all four cases, the effect of an experimental manipulation (i.e., C1 vs. C2; for instance, congruent stimuli vs. incongruent stimuli) increases the value of the dependent variable (e.g., mean response time). In all four cases, the effect is .100 for the group indicated by the dots, and .200 for the group indicated by the Xs: both groups increase, but the increase in one group is twice as large as in the other. Across the four panels, the data for the group indicated by the Xs undergo a constant shift along the y-axis. Despite their surface similarities, the interpretation of the interaction differs dramatically across the four panels. See text for details. Figure available at `http://tinyurl.com/p9kl2aa`, under CC license `https://creativecommons.org/licenses/by/2.0/`.

References

Anscombe, F. J. (1973). Graphs in statistical analysis. *The American Statistician*, *27*, 17–21.

Bejjanki, V. R., Zhang, R., Li, R., Pouget, A., Green, C. S., Lu, Z.-L., & Bavelier, D. (2014). Action video game play facilitates the development of better perceptual templates. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 16961–16966.

Bialystok, E., Craik, F. I. M., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Science*, *16*, 240–250.

Chambers, C. D. (2013). Registered Reports: A new publishing initiative at Cortex. *Cortex*, *49*, 609–610.

Hilchey, M. D., Saint-Aubin, J., & Klein, R. M. (in press). Does bilingual exercise enhance cognitive fitness in traditional non-linguistic executive processing tasks? In J. W. Schwieter (Ed.), *Cambridge handbook of bilingual processing.* Cambridge, UK: Cambridge University Press.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short- and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 10081–10086.

Loftus, G. R. (1978). On interpretation of interactions. *Memory & Cognition*, *6*, 312–319.

Matzke, D., Nieuwenhuis, S., van Rijn, H., Slagter, H. A., van der Molen, M. W., & Wagenmakers, E.-J. (2015). The effect of horizontal eye movements on free recall: A preregistered adversarial collaboration. *Journal of Experimental Psychology: General*, *144*, e1–e15.

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., Ishiyama, J., Karlan, D., Kraut, A., Lupia, A., Mabry, P., Madon, T. A., Malhotra, N., Mayo-Wilson, E., McNutt, M., Miguel, E., Levy Paluck, E., Simonsohn, U., Soderberg, C., Spellman, B. A., Turitto, J., VandenBos, G., Vazire, S., Wagenmakers, E.-J., Wilson, R., & Yarkoni, T. (2015). Promoting an open research culture. *Science*, *348*, 1422–1425.

Paap, K. R., Johnson, H. A., & Sawi, O. (in press). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*.

Redick, T. S. (2015). Working memory training and interpreting interactions in intelligence interventions. *Intelligence*, *50*, 14–20.

Redick, T. S., & Webster, S. B. (2014). Videogame interventions and spatial ability interactions. *Frontiers in Human Neuroscience*, *8*.

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin*, *138*, 628–654.

van Ravenzwaaij, D., Boekel, W., Forstmann, B., Ratcliff, R., & Wagenmakers, E.-J. (2014). Action video games do not improve the speed of information processing in simple perceptual tasks. *Journal of Experimental Psychology: General*, *143*, 1794–1805.

Wagenmakers, E.-J., Krypotos, A.-M., Criss, A. H., & Iverson, G. (2012). On the interpretation of removable interactions: A survey of the field 33 years after Loftus. *Memory & Cognition*, *40*, 145–160.