Commentary

Psychological Science
2017, Vol. 28(4) 547–550

Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797617697445
www.psychologicalscience.org/PS
SAGE

In a World of Big Data, Small Effects Can Still Matter: A Reply to Boyce, Daly, Hounkpatin, and Wood (2017)

Sandra C. Matz[1], Joe J. Gladstone[2], and David Stillwell[3]
[1]Department of Psychology, University of Cambridge; [2]UCL School of Management, University College London; and [3]Cambridge Judge Business School, University of Cambridge

Received 10/10/16; Revision accepted 2/12/17

We make three points in response to Boyce, Daly, Hounkpatin, and Wood (2017). First, we clarify a misunderstanding of the goal of our analyses, which was to investigate the links between life satisfaction and spending patterns, rather than spending volume. Second, we report a simulation study we ran to demonstrate that our results were not driven by the proposed statistical artifact. Finally, we discuss the broader issue of why, in a world of big data, small but reliable effect sizes can be valuable.

Our Goal Was to Study Spending Patterns, Not Spending Volume

Boyce et al. argue that our findings do not provide sufficient evidence for the contention that spending more money buys greater happiness when that money is spent on products and services that fit a person's personality. We agree, because this is not the argument we made in our article. Our argument was that spending the money one already has on products and services that match one's personality (and thus psychological needs) results in greater happiness. That is, we studied spending patterns (what people buy) rather than spending volume (how much of it they buy). In fact, we emphasized the relative unimportance of spending more money in our measures, analyses, and study designs. First, we considered the fact that many small purchases can result in greater happiness than a few large ones (e.g., Dunn, Gilbert, & Wilson, 2011) in the calculation of the basket-participant-match variable itself. Rather than weighting all spending categories by the amount spent, we assigned an equal weight to each of them. Second, we highlighted that the coefficients for income and total spending on life satisfaction were nonsignificant (Matz, Gladstone, & Stillwell, 2016, Table 3), which suggests that earning or spending more money does not affect life satisfaction and happiness. Finally, we did not vary the amount spent in our experimental study, but instead focused exclusively on manipulating the spending pattern (i.e., whether or not a voucher fitted the personality of the recipient).

Our Results Were Not Exclusively Driven by the Proposed Statistical Artifact

Boyce et al. suggest that the effect of the match between participants' shopping baskets and their personality might be driven "by participant personality, which is known to relate to well-being" (p. 545). We also had this concern. That is why we controlled for participant personality as well as the overall extremity of profiles in the original analysis. We found that the effects of basket-participant match remained stable when controlling for these variables (Matz et al., 2016, Table 3).

To provide further evidence for our effect's robustness, we ran a simulation on our data set that was similar to the one reported by Boyce et al. In 1,000 iterations, we randomly allocated the basket personality calculated for each participant in Study 1 to another participant (i.e., we "swapped" participants' shopping baskets at random). We then calculated the basket-participant match and ran regression analyses including the control variables used by Boyce et al. (our original Model 1 but without total spending) and the original control variables of Model 2 (see the Supplemental Material available online for more details). Out of 1,000 iterations, basket-participant match was significant at an alpha level of .05 in only 3.5% of

Corresponding Author:
Sandra C. Matz, University of Cambridge, Department of Psychology, Downing Street, Cambridge, CB2 3EB, United Kingdom
E-mail: sm917@cam.ac.uk

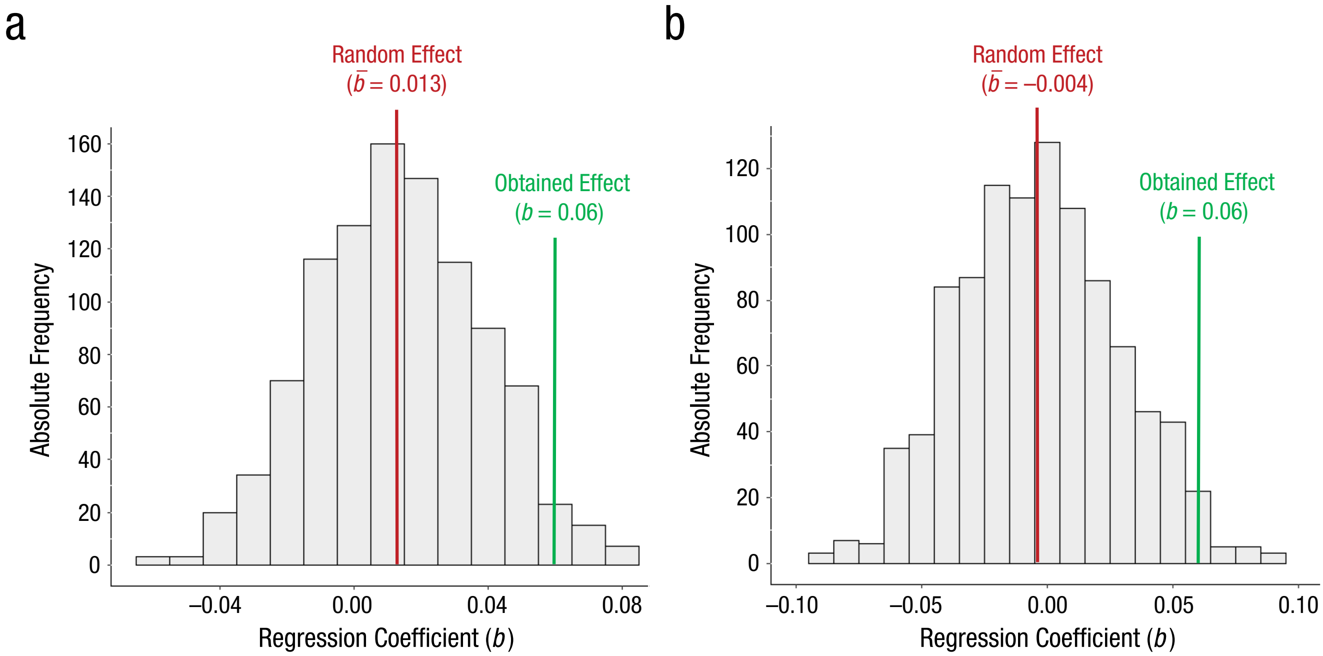

Fig. 1. Distribution of regression coefficients (*b*s) for randomly distributed basket-participant scores as predictors of life satisfaction across 1,000 iterations. Results are shown separately for models that included (a) the control variables of Boyce, Daly, Hounkpatin, and Wood (2017; i.e., our original Model 1 as reported in Matz, Gladstone, & Stillwell, 2016, but without total spending) and (b) the original control variables of our Model 2. In each graph, the average value of the coefficient is shown, along with the coefficient reported in our original article.

cases when including the controls used by Boyce et al., and in only 2.3% of cases when including all control variables of Model 2. The average coefficient (unstandardized) of the randomly generated basket-participant match was 0.013 in Model 1 (SD = 0.03; average β = 0.018; Fig. 1a) and −0.004 in Model 2 (SD = 0.03; average β = −0.005; Fig. 1b). These values contrast with coefficients (b = 0.06 and β = 0.10) reported for real basket-participant match in our original analysis (using the same controls). We take this as strong evidence that the relationship between basket-participant match and life satisfaction is not (or at least not exclusively) driven by the confounding artifacts suggested by Boyce et al.

The Importance of Small Effects

Boyce et al. argue that the relationship we reported between consumers' purchases and their well-being is of no practical relevance. This raises an important debate about the magnitude of meaningful effect sizes.

Traditional social-science research uses small samples (Bertamini & Munafò, 2012; Button et al., 2013), which require large effect sizes to reach statistical significance and be published. Large effect sizes are more likely to be replicable (Open Science Collaboration, 2015), but focusing exclusively on them hinders a nuanced exploration of complex psychological phenomena such as life satisfaction, which are unlikely to be explained by a few strong predictors. Instead of dismissing small effect sizes altogether, researchers should explore new methodologies that enable the discipline of psychology to build a body of small but robust predictors of behavior. New computational social-science approaches (see, e.g., Kosinski, Matz, Gosling, Popov, & Stillwell, 2015) make it possible to collect and analyze behavioral data at an unprecedented scale, which provides psychologists with opportunities to identify weak but reliable signals in a complicated world. However, the danger of using these new methods is that even trivial effects will become statistically significant with large enough samples. Therefore, how can one distinguish between small effects that are meaningful from those that are not?

This requires a scientist's judgment rather than the simple following of arbitrary cutoffs. A full review of situations in which small effect sizes are important would require more space than was available in this response, but other authors have already discussed some situations. Cortina and Landis (2009), for example, note that small effects can suggest strong support for a given phenomenon if they (a) occur in the context of intentionally inauspicious designs, (b) challenge fundamental theoretical assumptions, and (c) have enormous cumulative consequences. One illustration of the latter is Abelson's paradox (Abelson, 1985). Investigating the batting performance of Major League Baseball players, Abelson showed that historical batting performance was barely predictive of players' performance each time they stepped up to the batting plate (less than 0.5%). Yet the cumulative effect over several hundred at bats is so meaningful in practice that it leads to some batters being paid 100,000% more than others.

One challenge in determining the meaning in effect sizes is that it is difficult to directly compare them across studies using different methodologies, such as self-reported surveys and objective behavioral data. In our article, Study 1 was a big-data field study, in which more than 76,000 transactions were mined from customers' bank accounts and in which the measurement of well-being was far removed, in both context and time, from when the bank customers' purchases were being made. Such a procedure is likely to result in smaller—yet probably more realistic—effect sizes than panel surveys because they eliminate well-documented response biases (e.g., consistency motive, covariation bias, or common-method variance). For example, Powdthavee (2008; cited by Boyce et al.) reports a large effect of social contact on subjective well-being in the British Household Panel Survey. While social contact is surely important for well-being, the reported effect size might have been overestimated because respondents were subjectively reporting both their social contact and their well-being at the same time. Respondents who felt negatively about their lives, for example, might be less likely to remember and report short and superficial social contact, such as saying "hello" to a neighbor, than respondents who felt more positive (*mood congruence*; Bower, 1981). Indeed, the results of our Study 2 underline the importance of considering the methodological context when discussing a given effect size; Study 2's behavioral field experiment found a medium-to-large effect size for the interaction between product and participant personality ($\beta = 0.38$).

A final consideration when deciding whether a small effect may have practical relevance is the extent to which the factor can be manipulated. In their Commentary, Boyce et al. point to other factors that are stronger predictors of well-being (in terms of standardized effect sizes), such as relationships (Powdthavee, 2008), personality (Diener & Lucas, 1999), and stable employment (McKee-Ryan, Song, Wanberg, & Kinicki, 2005). Such factors certainly play an important role in predicting life satisfaction. However, personality cannot easily be changed (McCrae & Costa, 1994), and it is often outside the control of an unemployed person to arrange for stable employment. Consumption choices, on the other hand, are usually under the control of the individual and can be changed relatively easily. Even the very poorest groups in the world (i.e., those living on less than $2 per day) spend substantial amounts on discretionary goods, such as entertainment, celebrations, clothing, and tobacco (Banerjee & Duflo, 2007). Given that consumption is such a universal phenomenon, even small changes can have a big impact if they occur on a large scale. A 1% increase in life satisfaction may be negligible for a single consumer, but if a retail giant such as Amazon aimed at making its customers happier by personalizing their product recommendations, a 1% increase in life satisfaction across its 244 million customers (Kline, 2014) could turn a small effect size into a huge social effect.

Psychologists have typically focused on how their findings apply to individuals. However, by providing the opportunity to understand and influence the behaviors of billions of people around the world, the era of big data encourages—and possibly requires—researchers to think bigger. In this new world, small effects can still matter.

Action Editor

D. Stephen Lindsay served as action editor for this article.

Author Contributions

S. C. Matz ran the analyses. S. C. Matz and J. J. Gladstone wrote the manuscript under the supervision of D. Stillwell.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797617697445

References

Abelson, R. P. (1985). A variance explanation paradox: When a little is a lot. *Psychological Bulletin*, *97*, 129–133.

Banerjee, A. V., & Duflo, E. (2007). The economic lives of the poor. *The Journal of Economic Perspectives*, *21*(1), 141–167.

Bertamini, M., & Munafò, M. R. (2012). Bite-size science and its undesired side effects. *Perspectives on Psychological Science*, *7*, 67–71.

Bower, G. H. (1981). Mood and memory. *American Psychologist*, *36*, 129–148.

Boyce, C. J., Daly, M., Hounkpatin, H. O., & Wood, A. M. (2017). Money may buy happiness, but often so little that it doesn't matter. *Psychological Science, 28*, 544–546. doi:10.1177/0956797616672271

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*, 365–376.

Cortina, J. M., & Landis, R. S. (2009). When small effect sizes tell a big story, and when large effect sizes don't. In C. E. Lance & R. J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends: Doctrine, verity and fable in the* *organizational and social sciences* (pp. 287–308). New York, NY: Routledge.

Diener, E., & Lucas, R. E. (1999). Personality and subjective well-being. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 213–229). New York, NY: Russell Sage Foundation.

Dunn, E. W., Gilbert, D. T., & Wilson, T. D. (2011). If money doesn't make you happy, then you probably aren't spending it right. *Journal of Consumer Psychology*, *21*, 115–125. doi:10.1016/j.jcps.2011.02.002

Kline, D. B. (2014, May 14). How many customers does Amazon have? *The Motley Fool*. Retrieved from http://www.fool.com/investing/general/2014/05/24/how-many-customers-does-amazon-have.aspx

Kosinski, M., Matz, S. C., Gosling, S. D., Popov, V., & Stillwell, D. (2015). Facebook as a research tool for the social sciences. *American Psychologist*, *70*, 543–556.

Matz, S. C., Gladstone, J. J., & Stillwell, D. (2016). Money buys happiness when spending fits our personality. *Psychological Science*, *27*, 715–725. doi:10.1177/0956797616635200

McCrae, R. R., & Costa, P. T. (1994). The stability of personality: Observations and evaluations. *Current Directions in Psychological Science*, *3*, 173–175.

McKee-Ryan, F., Song, Z., Wanberg, C. R., & Kinicki, A. J. (2005). Psychological and physical well-being during unemployment: A meta-analytic study. *The Journal of Applied Psychology*, *90*, 53–76. doi:10.1037/0021-9010.90.1.53

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251). doi:10.1126/science.aac4716

Powdthavee, N. (2008). Putting a price tag on friends, relatives, and neighbours: Using surveys of life satisfaction to value social relationships. *The Journal of Socio-Economics*, *37*, 1459–1480. doi:10.1016/j.socec.2007.04.004