



Purity matters more than harm in moral judgments of suicide: Response to Gray (2014)



Joshua Rottman^{a,*}, Deborah Kelemen^a, Liane Young^b

^a Department of Psychological and Brain Sciences, Boston University, 64 Cummington Mall, Boston, MA 02215, United States

^b Department of Psychology, Boston College, 140 Commonwealth Ave., Chestnut Hill, MA 02467, United States

ARTICLE INFO

Article history:

Received 3 June 2014

Accepted 10 June 2014

Available online 10 July 2014

Keywords:

Morality

Suicide

Purity

Harm

Exclusion

Effect sizes

ABSTRACT

Many people judge suicide to be immoral. We have found evidence that these moral judgments are primarily predicted by people's belief that suicide taints the soul and by independent concerns about purity. This finding is inconsistent with accounts that define morality as fundamentally based upon harm considerations. In this commentary, we respond to a critique of our finding, and we provide further support for our original conclusions. Even when applying new exclusion criteria to our data, an examination of effect sizes demonstrates that concerns about purity robustly and meaningfully explain variance in moral judgments of suicide. While harm concerns sometimes predict moral judgments of suicide alongside purity concerns, they reliably explain a much smaller proportion of the variance than do purity concerns. Therefore, data from six studies continue to suggest that the relevance of harm concerns for moral judgments of suicide is substantially overshadowed by the contribution of purity concerns.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Suicide is undeniably harmful: it precludes the flourishing of one's future self, leads to tremendous devastation among family and community members, and increases the likelihood of other suicides. While some scholars point to such harm-based arguments to condemn suicide (e.g., Hecht, 2013), our studies demonstrate that moral judgments of suicide are instead primarily predicted by concerns about purity (Rottman, Kelemen, & Young, 2014). However, Gray (2014) takes issue with this conclusion, arguing that harm concerns predict moral judgments of suicide more meaningfully than we previously claimed. Here, we outline the reasons why Gray's critiques fail to undermine our prior conclusion. In particular, we demonstrate that even if concerns about harm do sometimes

reach statistical significance in predicting moral judgments of suicide, these harm-based concerns have relatively minor explanatory power compared to purity-based concerns.

2. Exclusion criteria and the bivariate outlier

Gray (2014) argues that our exclusion criteria were restrictive and that our conclusions were distorted by insufficient power. Specifically, he takes issue with our decision to exclude the fastest (but not slowest) participants. However, a minority of participants recruited via Amazon Mechanical Turk are motivated by payment rather than by the desire to produce high-quality work. They therefore hastily click through studies without carefully reading materials, making it prudent for researchers to filter out unreliable data by excluding people who progress through a study too quickly (Mason & Suri, 2012). By contrast, we do not think that going unusually slowly indicates

* Corresponding author. Tel.: +1 970 250 7048; fax: +1 617 353 6933.
E-mail address: rottman@bu.edu (J. Rottman).

any lack of comprehension or diligence; thus, our selection criteria were warranted.

Gray (2014) suggests that our removal of very rapid responders means that we have excluded participants who are answering most intuitively and are therefore most likely to perceive harm in immoral actions. This is unlikely, as reaction time was uncorrelated with assessments of harm (for Study 1, $r = .01$). Also, a median split between “fast” and “slow” participants shows that faster responders were not more likely to perceive a correlation between immorality and harm (for Study 1, $r_{fast} = .29$; $r_{slow} = .45$). Moreover, extremely fast reaction times are generally a poor proxy for intuitive responding (Evans, Dillon, & Rand, 2014).

Gray (2014) additionally finds a bivariate outlier that is an overly influential point in one of the six studies reported in our paper. Because the participant in question passed attention checks, and because her responses were sensible and internally consistent, we believe this data point represents a rare but valid observation in the Study 1 population rather than a contaminant. More importantly, *post hoc* selection can be a worrisome practice, and therefore in all of our studies we excluded participants solely on the basis of a *priori* decisions. By determining appropriate sample sizes and exclusion criteria before running analyses, our findings were insulated from motivated biases.

3. Comparison of effect sizes

As Gray (2014) notes, significance testing uses a somewhat arbitrary threshold as a heuristic for identifying a true positive result. He cites Cumming (2014), who asserts that effect sizes are more useful and interpretable. Indeed, a predictor variable is not necessarily meaningful if it is statistically significant but explains a relatively minor proportion of the unique variance in an outcome variable. For this reason, effect size testing is becoming an increasingly popular method for measuring a variable’s impact, and therefore we reported effect sizes for all variables in Studies 1 and 2 of our paper (Rottman et al., 2014). Here, we present squared semi-partial correlations (indicating the unique contribution of each predictor variable), using Gray’s suggested exclusion criteria (see Figs. 1 and 2). Consistent with our original exclusion criteria, these new analyses omitted non-American participants and participants

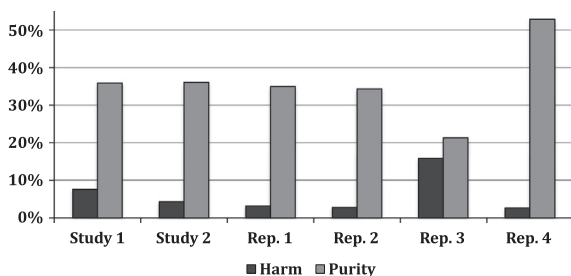


Fig. 1. Effect sizes (squared semi-partial correlations) indicating the unique percentage of variance in moral judgments of suicide that is explained by ratings of harm and ratings of impurity. For Study 2, “harm to others” was selected as the harm variable.

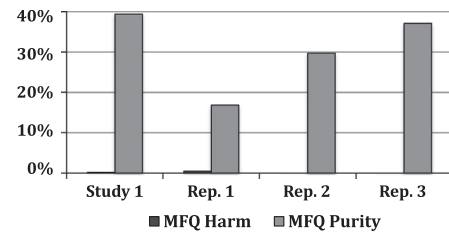


Fig. 2. Effect sizes (squared semi-partial correlations) indicating the unique percentage of variance in moral judgments of suicide that is explained by MFQ Harm and MFQ Purity. This questionnaire was not administered in Study 2 or in Replication 4.

who failed attention checks. However, participants were no longer excluded based upon reaction times, while data points with high Mahalanobis distances were omitted. We employed a conservative cutoff value of $p < .001$ to identify these outliers, and we considered all variables in each regression analysis rather than searching for bivariate outliers (Tabachnick & Fidell, 2013).

Even applying these new exclusion criteria, an examination of effect sizes continues to indicate that purity-based concerns explain a more substantial proportion of variance in moral judgments of suicide than harm-based concerns. A meta-analysis combining data from all six studies reveals that purity judgments explain 34.5% of the total variance ($f^2 = .61$), compared to 5.6% of total variance explained by harm judgments ($f^2 = .10$). This difference is even more pronounced when comparing effect sizes of scores on the Purity and Harm subscales of the Moral Foundations Questionnaire; a meta-analysis with data from all four studies employing this questionnaire finds that MFQ Purity explains 30.3% of variance in moral judgments of suicide ($f^2 = .44$), while MFQ Harm explains 0.0% of this variance ($f^2 = .00$). Given Cohen’s (1992) guidelines that “small,” “medium,” and “large” effect size indexes for f^2 are .02, .15, and .35, respectively, the results of these meta-analyses demonstrate that harm has a small-to-medium effect in explaining moral judgments of suicide, while purity has a consistently large effect. Across both measurements of these variables, harm is therefore a considerably less important construct than purity in accounting for variation in moral judgments of suicide.

4. Conclusion

The hypothesis that all moral judgments can be explained by reference to harm (Gray, Schein, & Ward, in press; Gray, Young, & Waytz, 2012) is interesting and provocative, and we agree that harm concerns can illuminate a great deal of moral psychology. Yet moral judgments are multifaceted (e.g., Greene, 2013; Sinnott-Armstrong & Wheatley, 2014; Young & Dungan, 2012), and our study is one of many that have demonstrated the rich plurality of the moral mind (see Graham et al., 2013; Haidt, 2012). There is room for conciliation between these views, however. For example, it is possible that certain moral judgments are driven by assessments of purity but consistently lead to later perceptions of harm (Gray et al.,

in press). In sum, it would be damaging for moral psychologists to ignore harm, but it would also be sully to ignore purity.

Acknowledgments

Preparation of this commentary was supported by NSF GRF DGE-1247312 to J.R., a John Templeton Foundation Grant to L.Y., and NSF 1007984 to D.K. We thank Tim Brown, Ben Emmert-Aronson, Jesse Graham, Kurt Gray, and Hao Wu for helpful comments and discussions.

References

- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159.
- Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, 25(1), 7–29. <http://dx.doi.org/10.1177/0956797613504966>.
- Evans, A. M., Dillon, K. D., & Rand, D. G. (2014). Reaction times and reflection in social dilemmas: Extreme responses are fast, but not intuitive. Available at SSRN: <<http://ssrn.com/abstract=2436750>>.
- Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S., et al. (2013). Moral Foundations Theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology*, 47, 55–130.
- Gray, K. (2014). Harm concerns predict moral judgments of suicide: Comment on Rottman, Kelemen, and Young (2014). *Cognition*.
- Gray, K., Schein, C., & Ward, A. F. (in press). The myth of harmless wrongs in moral cognition: Automatic dyadic completion from sin to suffering. *Journal of Experimental Psychology: General*. doi: <http://dx.doi.org/10.1037/a0036149>.
- Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, 23(2), 101–124. <http://dx.doi.org/10.1080/1047840X.2012.651387>.
- Greene, J. D. (2013). *Moral tribes: Emotion, reason, and the gap between us and them*. New York: Penguin.
- Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. New York: Pantheon.
- Hecht, J. M. (2013). *Stay: A history of suicide and the philosophies against it*. New Haven, CT: Yale University Press.
- Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods*, 44(1), 1–23.
- Rottman, J., Kelemen, D., & Young, L. (2014). Tainting the soul: Purity concerns predict moral judgments of suicide. *Cognition*, 130(2), 217–226. <http://dx.doi.org/10.1016/j.cognition.2013.11.007>.
- Sinnott-Armstrong, W., & Wheatley, T. (2014). Are moral judgments unified? *Philosophical Psychology*, 27(4), 451–474. <http://dx.doi.org/10.1080/09515089.2012.736075>.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Allyn and Bacon.
- Young, L., & Dungan, J. (2012). Where in the brain is morality? Everywhere and maybe nowhere. *Social Neuroscience*, 7(1), 1–10. <http://dx.doi.org/10.1080/17470919.2011.569146>.