

Humanizing Pilot

Blinded¹ & Blinded¹

¹ Blinded

Author Note

Department of Psychology, University of BLINDED. All procedures performed in studies involving human participants were approved by institutional research ethics committee and conducted in accordance with the Code of Professional Ethics of the Psychological Society of Ireland, and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study. The authors declare that there are no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. All authors consented to the submission of this manuscript.

Correspondence concerning this article should be addressed to Blinded, blinded.
E-mail: blinded

Abstract

15

16 Supplementary analysis to accompany the manuscript The Moral Dilution Effect:
17 Irrelevant Information Influences Judgments of Moral Character.

18 *Keywords:* keywords

19 Word count: X

Humanizing Pilot

Study S4: Humanizing vs Dilution Pilot Study

To examine the possibility that the observed dilution effect may have been confounded by a humanizing effect of the non-diagnostic information rather than as a result of dilution, we conducted a follow-up pilot study with additional measures of both humanness and relatableness. The assumption is that if the non-diagnostic information is having a humanizing effect, the humanness and relatable ratings for the non-diagnostic descriptions than for diagnostic descriptions (for both *bad* and *good* actors).

Study S4: Method

Study S4: Participants and Design. Study S4 was a within-subjects design. The independent variable was information type with three levels *good*, *bad*, and *neutral* (non-diagnostic). There were three dependent variables: morality ratings, humanness ratings, and relatableness ratings.

For morality ratings, we predicted that morality ratings would be highest for the *good* descriptions, and lowest for the *bad* descriptions, with the *neutral* descriptions being rated lower than *good* and higher than *bad*.

If the dilution effect is being driven by a humanizing effect of non-diagnostic information, this would predict a different pattern for humanness and relatableness ratings. Specifically, humanness and relatableness ratings would be expected to be highest for the *neutral* (non-diagnostic) conditions compared to the *bad* and *good* conditions For the humanness and relatableness ratings. If this pattern is not observed, it suggests that our findings may be evidence of a dilution effect rather than a humanizing effect.

A total sample of 100 (41 female, 59 male, 0 non-binary, 0 other; 0 prefer not to say, $M_{\text{age}} = 40.87$, $\text{min} = 19$, $\text{max} = 70$, $SD = 12.65$) started the survey. Participants were

recruited from Prolific and paid £0.40 for their participation.

Study S4: Procedure and Materials. The materials were the similar to those used in Study 3. Participants read six descriptions, two *good*, two *bad*, and two *neutral* (non-diagnostic). The content of these descriptions was taken directly from Study 3. For each description participants rated the character’s morality, the humanness, and the relatableness. The morality rating was the same as the MM-1 rating used in previous studies. The humanness and relatableness ratings were adapted from Martin and Mason (2022). Participants read the following instructions “The following scale represents humanness levels. Please rate the humanness of X”, 0 = *Not at all Human*, 100 = *Very Human*; “The following scale represents relatableness levels. Please rate the relatableness of X”, 0 = *Not at all Relatable*, 100 = *Very Relatable*.

Sam

Imagine a person named Sam. Throughout their life they have been known to always help and care for others, treat everyone fairly and equally, and show a strong sense of loyalty to others.

Robin

Imagine a person named Robin. Throughout their life they have been known to protect and provide shelter to the weak and vulnerable, uphold the rights of others, and show respect for authority.

Francis

Imagine a person named Francis. Throughout their life they have been known to cause others to suffer emotionally, to deny others their rights, and to cause chaos or disorder.

Alex

Imagine a person named Alex. Throughout their life they have been known to be

69 cruel, act unfairly, and to betray their own group.

70 *Jordan*

71 Imagine a person named Jordan. They have red hair, play tennis four times a month,
72 and have one older sibling and one younger sibling.

73 *Charlie*

74 Imagine a person named Charlie. They are left-handed, drink tea in the morning,
75 and have two older siblings and one younger sibling.

76 **Study S4: Results**

77 **Morality.** We conducted a linear-mixed-effects model to test if information type
78 influenced morality ratings. Our outcome measure was morality rating, our predictor
79 variable was valence/information-type; we allowed intercepts to vary across participants.
80 Overall, the model significantly predicted participants responses, and provided a better fit
81 for the data than the baseline model, $\chi^2(2) = 945.33$, $p < .001$. Information type
82 significantly influenced morality ratings, $F(2, 498.00) = 1,216.56$, $p < .001$. Tukey's
83 post-hoc pairwise comparisons indicated that the highest rated descriptions were the good
84 descriptions, ($M = 89.10$, ($SD = 13.47$), and these were significantly higher ($p < .001$)
85 than the bad descriptions ($M = 13.71$, and significantly higher ($p < .001$) than the
86 neutral/non-diagnostic descriptions ($SD = 18.34$). The neutral/non-diagnostic descriptions
87 were also significantly higher ($p < .001$) than the bad descriptions.

88 **Humanness.** We conducted a linear-mixed-effects model to test if information type
89 influenced humanness ratings. Our outcome measure was humanness rating, our predictor
90 variable was valence/information-type; we allowed intercepts across participants. Overall,
91 the model significantly predicted participants responses, and provided a better fit for the
92 data than the baseline model, $\chi^2(2) = 616.31$, $p < .001$. Information type significantly
93 influenced morality ratings, $F(2, 498.00) = 602.40$, $p < .001$. Tukey's post-hoc pairwise

comparisons indicated that the highest rated descriptions were the good descriptions, ($M = 88.84$, ($SD = 13.15$), and these were significantly higher ($p < .001$) than the bad descriptions ($M = 29.62$, and significantly higher ($p < .001$) than the neutral/non-diagnostic descriptions ($SD = 28.59$). The neutral/non-diagnostic descriptions were also significantly higher ($p < .001$) than the bad descriptions.

Relatableness. We conducted a linear-mixed-effects model to test if information type influenced relatableness ratings. Our outcome measure was relatableness rating, our predictor variable was valence/information-type; we allowed intercepts across participants. Overall, the model significantly predicted participants responses, and provided a better fit for the data than the baseline model, $\chi^2(2) = 655.42$, $p < .001$. Information type significantly influenced morality ratings, $F(2, 498.00) = 665.91$, $p < .001$. Tukey's post-hoc pairwise comparisons indicated that the highest rated descriptions were the good descriptions, ($M = 79.64$, ($SD = 17.94$), and these were significantly higher ($p < .001$) than the bad descriptions ($M = 19.23$, and significantly higher ($p < .001$) than the neutral/non-diagnostic descriptions ($SD = 22.61$). The neutral/non-diagnostic descriptions were also significantly higher ($p < .001$) than the bad descriptions.

References

- Martin, A. E., & Mason, M. F. (2022). What does it mean to be (seen as) human? The importance of gender in humanization. *Journal of Personality and Social Psychology*, 123(2), 292–315. <https://doi.org/10.1037/pspa0000293>