

Study 3

Blinded¹, Blinded², Blinded¹, and & Blinded¹

¹ Blinded

² Blinded

Author Note

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

Abstract

Six studies etc.

Keywords: keywords

Word count: TBC

Study 3

Study 3

In Study 1 we found evidence for the moral dilution effect for judgments of *bad* moral characters. In Study 2 we failed replicate this effect for judgments of *good* moral characters. The aim of Study 3 was to test if valence (good vs. bad) moderates the moral dilution effect. We hypothesized that valence (good vs bad) would interact with condition in producing a dilution effect, such that the dilution effect would be observed for bad characters but not for good characters. Study 3 was pre-registered at https://aspredicted.org/QDF_XT1.

Study 3: Method

Study 3: Participants and design

Study 3 was a 2×2 within-subjects factorial design. The first independent variable was condition with two levels, diagnostic and non-diagnostic. The second independent variable was valence of character description, with two levels morally good and morally bad. We used the same two dependent variables as in previous studies, the four item moral perception scale (MPS-4, $\alpha = 0.94$), and the single item moral perception measure MM-1.

A total sample of 1095 (700 female, 386 male, 2 non-binary, 0 other; 2 prefer not to say, $M_{\text{age}} = 36.42$, $\text{min} = 19$, $\text{max} = 77$, $SD = 10.65$) started the survey. Participants were recruited from MTurk and paid \$0.40 for their participation.

Participants who failed both manipulation checks were removed ($n = 221$), leaving a total sample of 874 participants (550 female, 320 male, 0 other, 0 prefer not to say; $M_{\text{age}} = 36.37$, $\text{min} = 19$, $\text{max} = 77$, $SD = 10.72$).

Study 3: Procedure and materials

Again, data were collected using an online questionnaire presented with Qualtrics (www.qualtrics.com). Participants were presented with four descriptions of characters taken from Studies 1 and 2. To ensure consistency across character judgments, we selected descriptions that related to the same moral foundations (care, fairness, and loyalty). We used the same four character names as in previous studies. The *good* characters were *Sam* and *Robin*, and the *bad* characters were *Francis* and *Alex*, e.g., *Imagine a person named Robin. Throughout their life they have been known to show compassion and empathy for others, act with a sense of fairness and justice, and, never to break their word.* or, *Imagine a person named Alex. Throughout their life they have been known to be cruel, act unfairly, and to betray their own group.* Full descriptions for each character are in the supplementary materials. One description for each the *good* and *bad* characters was randomly assigned to include non-diagnostic information for each participant thus all participants were exposed to all conditions (see https://osf.io/mdnpv/?view_only=77883e3fbc3d45f1a35fe92d5318cb67 for details of the randomization blocks). Study 3 was pre-registered at https://aspredicted.org/QDF_XT1

Study 3: Results

The means and standard deviations for MPS-4 for each scenario are as follows: *Sam*, $M_{\text{MPS-4}} = 5.90$, $SD_{\text{MPS-4}} = 1.03$, *Francis*, $M_{\text{MPS-4}} = 4.07$, $SD_{\text{MPS-4}} = 2.07$, *Alex*, $M_{\text{MPS-4}} = 4.03$, $SD_{\text{MPS-4}} = 2.03$, *Robin*, $M_{\text{MPS-4}} = 5.85$, $SD_{\text{MPS-4}} = 1.05$. There was significant variation depending on the description, $F(1,1080) = 442.71$, $p < .001$, partial $\eta^2 = 0.24$. Both the *good* characters (*Robin* and *Sam*) were rated significantly more favorably than both the *bad* characters (*Alex* and *Francis*; all $ps < .001$). There were no differences between *Robin* and *Sam* (*good*: $p = .366$) or between *Alex* and *Francis* (*bad*; ($p = .648$)).

The means and standard deviations for MM-1 for each scenario are as follows: *Sam* (diagnostic/moral), $M_{\text{MM-1}} = 81.01$, $SD_{\text{MM-1}} = 15.23$; *Francis* (diagnostic/moral), $M_{\text{MM-1}}$

= 51.49, $SD_{MM-1} = 33.18$; *Alex* (diagnostic/moral), $M_{MM-1} = 50.89$, $SD_{MM-1} = 32.14$; *Robin* (diagnostic/moral), $M_{MM-1} = 80.81$, $SD_{MM-1} = 15.16$. There was significant variation depending on the description, $F(1,1080) = 458.92$, $p < .001$, partial $\eta^2 = 0.254$. Again, the *good* characters (*Robin* and *Sam*) were rated significantly more favorably than the *bad* characters (*Alex* and *Francis*; all $ps < .001$). There were no differences between *Robin* and *Sam* (*good*: $p = .776$) or between *Alex* and *Francis* (*bad*; ($p = .683$)).

We conducted a linear-mixed-effects model to test if our predictors influenced MPS-4 responses. Our outcome measure was MPS-4, our predictor variables were condition and valence; we allowed intercepts and the effects of condition and valence to vary across participants. Overall, the model significantly predicted participants responses, and provided a better fit for the data than the baseline model, $\chi^2(5) = 3,420.34$, $p < .001$. As expected, on its own, condition did not influence responses to the MPS-4, $F(1, 1746) = 0.01$, $p = .937$; valence significantly predicted responses, $F(1, 1746) = 587.37$, $p < .001$; and there was a significant condition \times valence interaction, $F(1, 1746) = 9.13$, $p = .003$. and was not a significant predictor in the model when controlling for scenario, $b = 0.00$, $t(1,746.00) = -0.08$, $p = .937$.

We conducted a linear-mixed-effects model to test if our predictors influenced MM-1 responses. The model was the same as the previous model, with a change to the outcome measure, our outcome measure for this model was MM-1. As above, our predictor variables were condition and valence; we allowed intercepts and the effects of condition and valence to vary across participants. Overall, the model significantly predicted participants responses, and provided a better fit for the data than the baseline model, $\chi^2(5) = 3,441.43$, $p < .001$. As expected, on its own, condition did not influence responses to the MPS-4, $F(1, 1746) = 0.03$, $p = .852$; valence significantly predicted responses, $F(1, 1746) = 638.14$, $p < .001$; and there was a significant condition \times valence interaction, $F(1, 1746) = 17.23$, $p < .001$. and was not a significant predictor in the model when controlling for scenario, $b = -0.03$, $t(1,746.00) = -0.19$, $p = .852$.

Differences in the *Bad* Descriptions

To interpret the interaction effect, we conducted separate analyses for the Good and Bad descriptions.

We conducted a linear-mixed-effects model to test if condition influenced MPS-4 responses. Our outcome measure was MPS-4, our predictor variable was condition; we allowed intercepts and the effect of condition to vary across participants. Overall, the model did not significantly predict participants responses, or provide a better fit for the data than the baseline model, $\chi^2(3) = 5.40$, $p = .145$. Condition did not significantly influence MPS-4 responses $F(1, 872.00) = 3.54$, $p = .060$, and was not a significant predictor in the model $b = -0.03$, $t(872.00) = -1.88$, $p = .060$, see Figure 1.

We conducted a linear-mixed-effects model to test if condition influenced MM-1 responses. Our outcome measure was MM-1, our predictor variable was condition; we allowed intercepts and the effect of condition to vary across participants. Overall, the model significantly predicted participants responses, and provided a better fit for the data than the baseline model, $\chi^2(3) = 8.67$, $p = .034$. Condition significantly influenced MM-1 responses $F(1, 872.00) = 7.01$, $p = .008$, and was a significant predictor in the model $b = -0.69$, $t(872.00) = -2.65$, $p = .008$, see Figure 1.

Differences in the *Good* Descriptions

To interpret the interaction effect, we conducted separate analyses for the Good and Bad descriptions.

We conducted a linear-mixed-effects model to test if condition influenced MPS-4 responses. Our outcome measure was MPS-4, our predictor variable was condition; we allowed intercepts and the effect of condition to vary across participants. Overall, the model significantly predicted participants responses, and provided a better fit for the data than the baseline model, $\chi^2(3) = 13.66$, $p = .003$. Condition significantly influenced MPS-4

responses $F(1, 872.00) = 6.82, p = .009$, and was a significant predictor in the model $b = 0.03, t(872.00) = 2.61, p = .009$, see Figure 1.

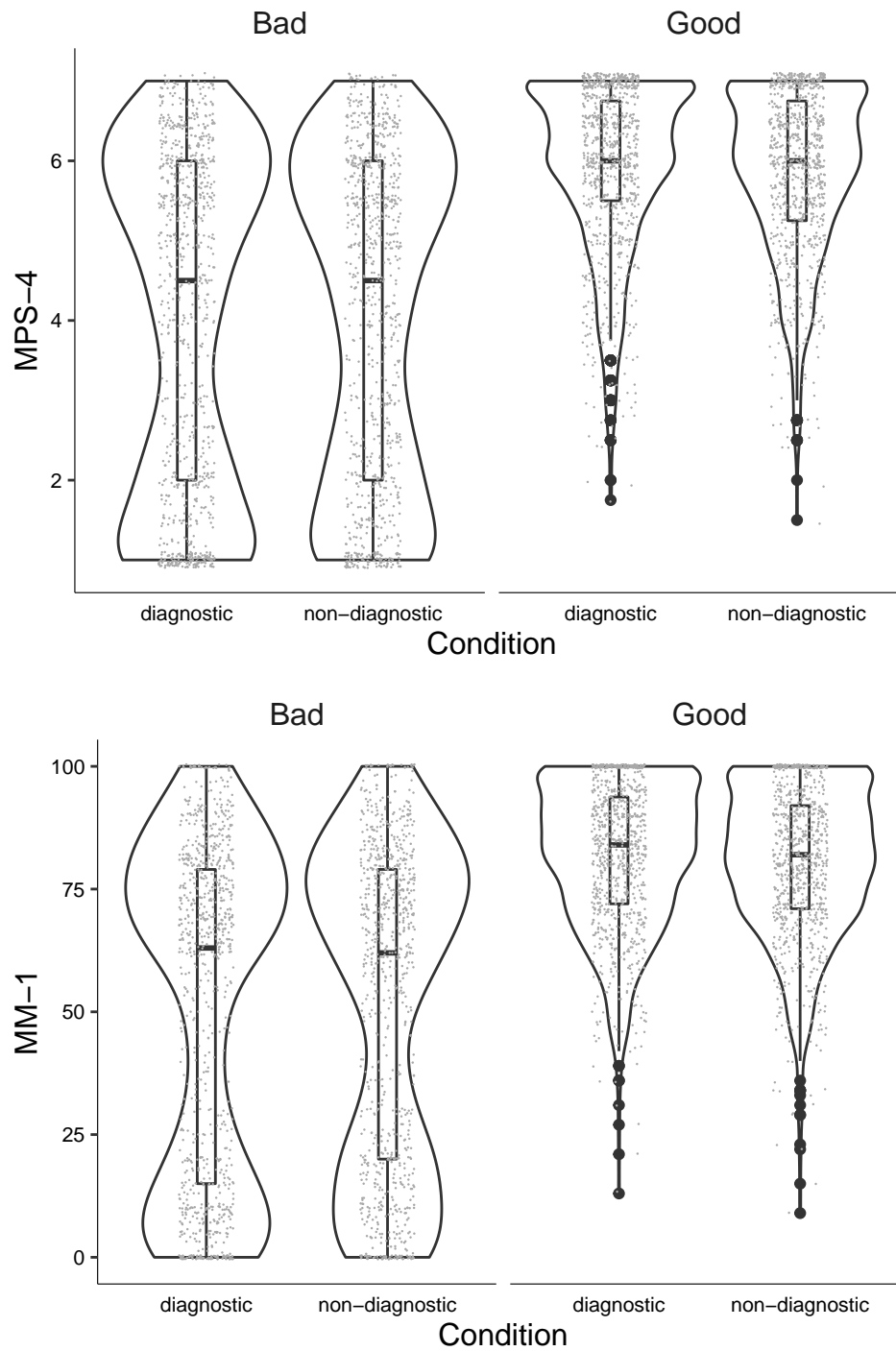


Figure 1

Study 3: Differences in moral perception depending on condition

We conducted a linear-mixed-effects model to test if condition influenced MM-1 responses. Our outcome measure was MM-1, our predictor variable was condition; we allowed intercepts and the effect of condition to vary across participants. Overall, the model significantly predicted participants responses, and provided a better fit for the data than the baseline model, $\chi^2(1) = 11.97$, $p < .001$. Condition significantly influenced MM-1 responses $F(1, 873) = 12.04$, $p < .001$, and was a significant predictor in the model $b = 0.63$, $t(873) = 3.47$, $p < .001$, see Figure 1.

In the supplementary analyses we report the effect of condition on moral perception for each description individually.

The aim of Study 3 was to test if the moral dilution effect was moderated by valence of description. Based on the results of Studies 1 and 2 we hypothesized that a dilution effect would be observed for judgments of *bad* characters, but not for judgments of *good* characters. Interestingly, in Study 2 we found a dilution effect for *good* characters across both measures, however we only found a dilution effect for *bad* characters when using the single item measure of moral perception.