Pilot Study 1

Blinded[1], Blinded[2], Blinded[1], and & Blinded[1]

[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

<p style="text-align:center">Pilot Study 1</p>

The aim of this pilot study was to develop and test materials that could be used to study the dilution effect for moral characters. We developed diagnostic and non-diagnostic character descriptions. We hypothesized that moral evaluations of the diagnostic descriptions would be more severe (more immoral) than for the non-diagnostic descriptions.

Pilot Study 1: Method

Pilot 1: Participants and design

The pilot study was a within-subjects design. The independent variable was description type with two levels, *diagnostic* and *non-diagnostic*. We used two dependent variables. The first dependent variable was the four item moral perception scale (MPS-4), participants rated the characters on four dimensions using 7-point bipolar scales. The dimensions and scale endpoints were: Bad-Good, Immoral-Moral, Violent-Peaceful, Merciless-Empathetic, this showed excellent reliability, $\alpha = 0.93$. The second dependent variable was a single item moral perception measure (MM-1) which consisted of a 100-point slider ranging from $0 = $ *Very Immoral* to $100 = $ *Very Moral.* Both dependent variables were taken from Walker et al. (2021).

A total sample of 235 (89 female, 142 male, 1 non-binary, 1 prefer not to say; M_{age} $= 36.45$, min $= 20$, max $= 72$, $SD = 10.23$) started the survey. Participants were recruited from MTurk.

We removed participants who failed both manipulation checks ($n = 23$), leaving a total sample of 212 participants (80 female, 128 male, 1 non-binary, 1 prefer not to say; $M_{\text{age}} = 36.63$, min $= 20$, max $= 72$, $SD = 10.34$).

Pilot 1: Procedure and materials

Data were collected using an online questionnaire presented with Qualtrics (www.qualtrics.com). Participants were presented with descriptions of six characters.

Moral character descriptions were developed by combining descriptions relating to three different moral foundations. These descriptions were adapted from the items of the extended character morality questionnaire (Grizzard et al., 2020). A sample description reads: *Imagine a person named Sam. Throughout their life they have been known to be cruel, act unfairly, and to betray their own group.* Full text of these descriptions can be found in the supplementary materials.

We developed neutral descriptions that included information relating to physical appearance/attributes, hobbies/activities, and family information, e.g., *Imagine a person named Jackie. They have red hair, play tennis four times a month, and have one older sibling and one younger sibling.*

Character descriptions did not specify the gender of the charcters, and all characters had names that could be either male or female (Sam, Robin, Francis, Alex, Jackie, Charlie). All participants read six descriptions, four moral descriptions and two neutral. Pilot Study 1 was pre-registered at https://aspredicted.org/3VK_8FD.

Pilot 1: Results

The means and standard deviations for MPS-4 for each scenario are as follows: *Sam* (diagnostic), $M_{\text{MPS-4}} = 4.35$, $SD_{\text{MPS-4}} = 1.90$, *Francis* (diagnostic), $M_{\text{MPS-4}} = 4.46$, $SD_{\text{MPS-4}} = 1.73$, *Alex* (diagnostic), $M_{\text{MPS-4}} = 4.44$, $SD_{\text{MPS-4}} = 1.79$, *Robin* (diagnostic), $M_{\text{MPS-4}} = 4.35$, $SD_{\text{MPS-4}} = 1.96$, *Jackie* (non-diagnostic), $M_{\text{MPS-4}} = 5.40$, $SD_{\text{MPS-4}} = 1.01$, *Charlie* (non-diagnostic), $M_{\text{MPS-4}} = 5.38$, $SD_{\text{MPS-4}} = 1.01$. For the diagnostic descriptions, there was no significant variation depending on the description, $F(3,600) = 1.58$, $p = .194$, partial $\eta^2 = 0.00$. For the non-diagnostic descriptions there was no significant difference in ratings depending on description, $t(211) = -0.67$, $p = .506$, $d = 0.05$.

The means and standard deviations for MM-1 for each scenario are as follows: *Sam* (diagnostic), $M_{\text{MM-1}} = 55.67$, $SD_{\text{MM-1}} = 30.47$; *Francis* (diagnostic), $M_{\text{MM-1}} = 58.22$, $SD_{\text{MM-1}} = 28.61$; *Alex* (diagnostic), $M_{\text{MM-1}} = 56.80$, $SD_{\text{MM-1}} = 29.45$; *Robin* (diagnostic), $M_{\text{MM-1}} = 55.49$, $SD_{\text{MM-1}} = 31.38$; *Jackie* (non-diagnostic), $M_{\text{MM-1}} = 73.00$, $SD_{\text{MM-1}} = 14.72$; *Charlie* (non-diagnostic), $M_{\text{MM-1}} = 72.94$, $SD_{\text{MM-1}} = 14.79$. For the diagnostic descriptions, we observed significant variation depending on the description, $F(3,608) = 3.01$, $p = .032$, partial $\eta^2 = 0.001$. When correcting for multiple comparisons, pairwise comparisons did not reveal significant differences between descriptions. We note that without correction, *Francis* appeared to be rated as more moral than both *Robin* ($p = .012$), and *Sam* ($p = .009$). For the non-diagnostic descriptions there was no significant difference in ratings depending on description, $t(211) = -0.09$, $p = .929$, $d = 0.01$.

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, $\chi^2(2) = 860.16$, $p < .001$. Condition was a significant predictor in the model $b = -0.49$, $t(211.05) = -8.54$, $p < .001$, with the non-diagnostic descriptions being rated as more moral than the diagnostic descriptions of immoral characters 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, $\chi^2(2) = 924.82$, $p < .001$. Condition was a significant predictor in the model $b = -8.22$, $t(210.98) = -8.60$, $p < .001$, with the non-diagnostic descriptions being rated as more moral than the diagnostic descriptions, see Figure 1.

Grizzard, M., Fitzgerald, K., Francemone, C. J., Ahn, C., Huang, J., Walton, J., ...

Eden, A. (2020). Validating the extended character morality questionnaire. *Media Psychology, 23*(1), 107–130. https://doi.org/10.1080/15213269.2019.1572523

Walker, A. C., Turpin, M. H., Fugelsang, J. A., & Białek, M. (2021). Better the two devils you know, than the one you don't: Predictability influences moral judgments of immoral actors. *Journal of Experimental Social Psychology, 97*, 104220. https://doi.org/10.1016/j.jesp.2021.104220
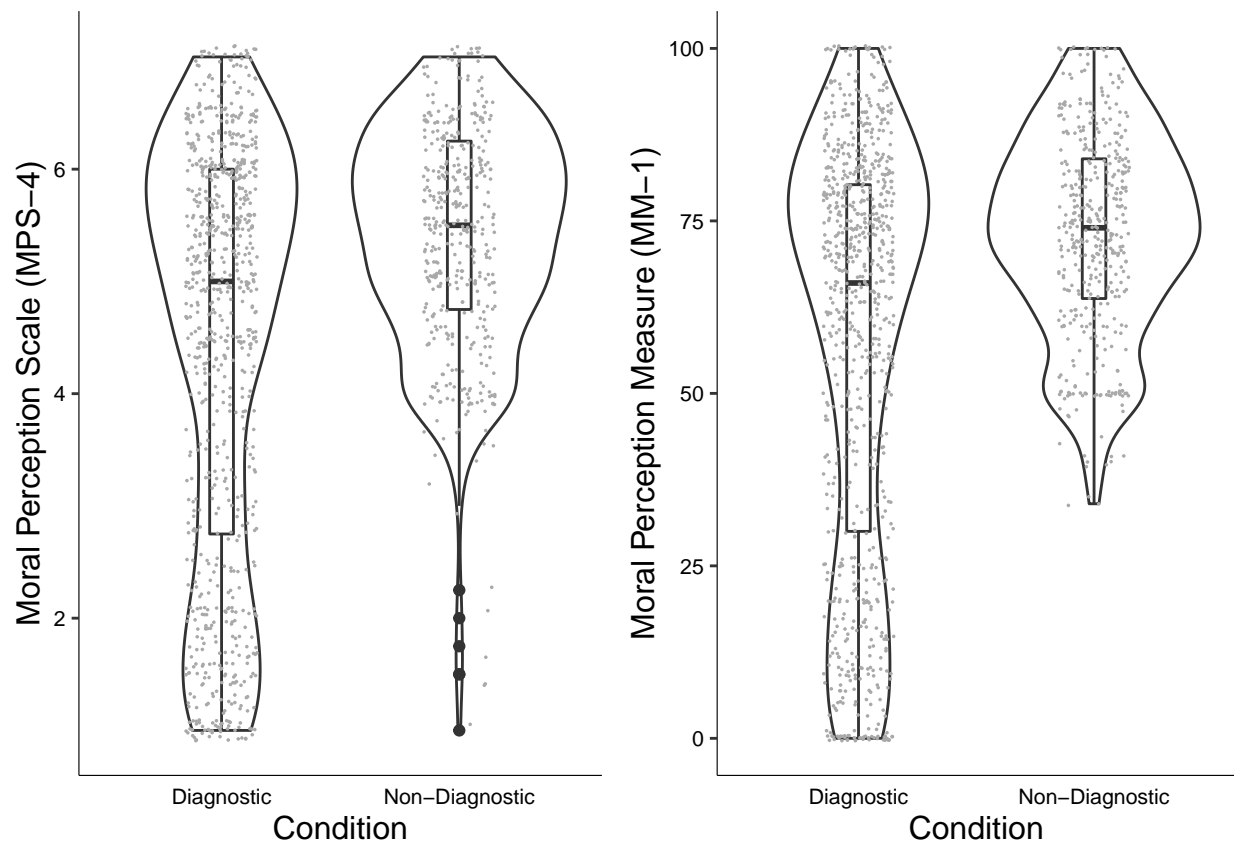
Figure 1

Pilot Study 1: Differences in moral perception depending on condition