

Study 6

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 6

Study 6 - Pre-registered Replication

In response to the mixed findings in Studies 1-5, the aim of Study 6 was to test if cognitive load influences participants' ability to justify their judgments in high-powered, large sample study. Study 6 was pre-registered (see https://aspredicted.org/XZP_UHW).

Study 6: Method

Study 6: Participants and design

Study 6 was a between subjects design. The dependent variable was rates of providing reasons/dumbfounding (measured using to the critical slide as in previous studies). The primary independent variable was cognitive load with two levels: present and absent. To manipulate cognitive load, a stream of numbers scrolled across the screen above the question text, and participants were required to pay attention to how many times they saw a given number. Scenario served as a secondary independent variable, we used four scenarios: *Julie and Mark (Incest)*, *Jennifer (Cannibal)*, *Trolley*, *Heinz*.

A total sample of 1899 participants (984 female, 876 male, 17 non-binary, 1 other, 5 prefer not to say; $M_{\text{age}} = 43.22$, $\text{min} = 18$, $\text{max} = 84$, $SD = 15.85$) started the survey. Participants in this sample were recruited from Prolific ($n_{UK} = 963$, $n_{US} = 936$).

Participants who failed both manipulation checks ($n = 7$) or who had missing data for the measures of interest were removed, leaving a total sample of 1686 participants (867 female, 799 male, 14 non-binary, 1 other, 5 prefer not to say; $M_{\text{age}} = 43.81$, $\text{min} = 18$, $\text{max} = 83$, $SD = 15.76$), $n_{UK} = 842$, $n_{US} = 844$.

Study 6: Procedure and materials

Data were collected using an online questionnaire developed with jsPsych and distributed with cognition.run. Participants were presented with one of four moral scenarios (*Julie and Mark*, *Jennifer*, *Trolley*, *Heinz*, see supplementary materials for full wording), previously used in studies of moral dumbfounding (McHugh et al., 2017). The sequence of responses was similar to previous studies, but with content of questions and counter arguments updated to match the relevant scenarios (see supplementary materials).

To manipulate cognitive load manipulation in Study 6, drawing on Greene, Morelli, Lowenberg, Nystrom, and Cohen (2008) we included a video stream of numbers scrolling above the question text. The video was wide enough to display 3 numbers at a time, and the numbers scrolled past at a speed of 2 numbers per second. Participants were asked to attend to and report (on a subsequent page) how many times a particular number appeared in the stream, while answering the target question. Following an initial training task, the video was presented while participants made their initial judgments, while they responded to the critical slide, and while they were providing their revised judgments.

Two attention check tasks were included for all participants, these included a brief paragraph of text where instructions for the correct response were embedded within the text. The wording of the text was misleading such that if participants skimmed or only read some of the text they would likely provide an incorrect response. In order to keep the survey short to facilitate the recruiting of a large sample, we did not record Need for Cognition in Study 6.

Participants clicked on the survey link and were randomly assigned to either the experimental condition or the control condition, within which they were randomly presented with one of the four scenarios. The study was complete within 5 minutes.

Study 6: Results

One thousand three hundred sixty-five participants (80.96%) rated the behavior described as wrong initially, and one thousand three hundred forty three participants (79.66%) rated the behavior as wrong at the end of the task. Initial ratings ($M = 2.26$, $SD = 1.63$) were significantly more severe than revised ratings ($M = 2.34$, $SD = 1.66$), $t(1685) = -2.69$, $p = .007$; $d = 0.07$. Inspection of the binned judgments revealed that two hundred (11.86%) participants changed the valence of their judgments, breakdown of the changes in judgments is in Table XX (full sample) and Table XX (by scenario) in the supplementary materials.

A 2×2 factorial ANOVA revealed significant differences in initial judgments depending on both condition $F(1, 1678) = 26.65$, $p < .001$, partial $\eta^2 = .016$, and scenario $F(3, 1678) = 69.30$, $p < .001$, partial $\eta^2 = .110$. Participants under cognitive load were significantly ($p < .001$) less harsh in their judgments ($M = 2.46$, $SD = 1.75$) than those in the control condition ($M = 2.07$, $SD = 1.49$). Participants rated *Jennifer* as the most wrong ($M = 1.53$, $SD = 1.13$), followed by *Julie and Mark* ($M = 2.05$, $SD = 1.65$, $p < .001$), then *Heinz* ($M = 2.49$, $SD = 1.65$, $p < .001$), with *Trolley* receiving the least severe judgment ($M = 2.98$, $SD = 1.69$, $p < .001$). There was no significant condition \times scenario interaction $F(3, 1678) = 0.46$, $p = .708$, partial $\eta^2 < .001$.

A 2×2 factorial ANOVA revealed significant differences in revised judgments depending on both condition $F(1, 1678) = 12.82$, $p < .001$, partial $\eta^2 = .008$, and scenario $F(3, 1678) = 80.69$, $p < .001$, partial $\eta^2 = .126$. Participants under cognitive load were significantly ($p < .001$) less harsh in their judgments ($M = 2.47$, $SD = 1.71$) than those in the control condition ($M = 2.20$, $SD = 1.59$). Participants rated *Jennifer* as the most wrong ($M = 1.54$, $SD = 1.12$), followed by *Julie and Mark* ($M = 2.15$, $SD = 1.73$, $p < .001$), then *Heinz* ($M = 2.52$, $SD = 1.58$, $p = .003$), with *Trolley* receiving the least severe judgment ($M = 3.14$, $SD = 1.72$, $p < .001$). There was no significant condition \times scenario

interaction $F(3, 1678) = 1.34$, $p = .260$, partial $\eta^2 = .002$.

Dumbfounding was recorded using the critical slides, participants who selected the admission of not having reasons on the critical slide were identified as dumbfounded. Four hundred and seventeen participants (24.73%) selected “It’s wrong but I can’t think of a reason.” One thousand and thirty-two participants (61.21%) selected “It’s wrong and I can provide a valid reason”; and two hundred and thirty-seven participants (14.06%) selected “There is nothing wrong.”

A chi-squared test for independence revealed a significant association between experimental condition and response to the critical slide, $\chi^2(2, N = 1686) = 25.48$, $p < .001$, $V = 0.12$, the observed power was 0.997. As predicted, under cognitive load fewer participants (458; 55.45%) provided reasons than in the control condition (574; 66.74%), and more participants (245; 29.66%) selected “It’s wrong but I can’t think of a reason.” than in the control group (172; 20%). The responses to the critical slide for the experimental group ($N = 826$) and the control group ($N = 860$) are displayed in Figure 1. The observed counts, expected counts and standardised residuals are displayed in Table 1.

This pattern was observed for all scenarios individually with the exception of *Julie and Mark*, which showed no association between experimental condition and cognitive load, $\chi^2(2, N = 418) = 0.49, p = .783, V = 0.25, \text{power} = 0.601$. The association was significant for *Jennifer* $\chi^2(2, N = 418) = 17.33, p < .001, V = 0.24, \text{power} = 0.623$, *Trolley* $\chi^2(2, N = 418) = 10.95, p = .004, V = 0.25, \text{power} = 0.614$, and *Heinz*, $\chi^2(2, N = 418) = 7.16, p = .028, V = 0.25, \text{power} = 0.608$, see Figure 2. Supplementary Tables XX-XX show the direction of the effect for each scenario. Under cognitive load, fewer participants provided reasons and more participants provided a dumbfounded response for *Jennifer*, *Trolley*, and *Heinz*.

A chi-squared test for independence revealed a significant association between scenario and response to the critical slide, $\chi^2(6, N = 1686) = 61.34, p < .001, V = 0.19$, the observed power was 1. Participants were significantly more likely to select “There is nothing wrong” for *Julie and Mark* ($p = .002$), more likely to provide reasons ($p = .002$) and less likely to select “There is nothing wrong” ($p < .001$) for *Jennifer*, and more likely to be dumbfounded by *Trolley* ($p = .031$).

A multinomial logistic regression was conducted to test the effects of cognitive load and scenario on dumbfounded responding. Overall the model was significant, $\chi^2(8, N = 1686) = 95.9, p < .001$, and explained between 6.07% (Cox and Snell R square) and 7.22% (Nadelkerke R squared) of the variance in responses to the critical slide, the observed power was 1. Participants in the control condition were significantly less likely to provide a dumbfounded response than to provide reasons, Wald = 25.04, $p < .001, OR = 0.55, 95\% \text{ CI } [0.44, 0.70]$, in addition, participants in the control condition were also significantly less likely to select “There is nothing wrong,” than to provide reasons, Wald = 5.23, $p = .022, OR = 0.71, 95\% \text{ CI } [0.54, 0.95]$. For *Jennifer*, participants were significantly less likely to select “There is nothing wrong” than to provide a reason, Wald = 30.87, $p < .001, OR = 0.23, 95\% \text{ CI } [0.13, 0.38]$; while for *Trolley* participants were significantly more likely to present as dumbfounded than to provide a reason, Wald = 6.89, $p = .009, OR = 1.55, 95\%$

CI [1.12, 2.14].

Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008).

Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, 107(3), 1144–1154. <https://doi.org/10.1016/j.cognition.2007.11.004>

McHugh, C., McGann, M., Igou, E. R., & Kinsella, E. L. (2017). Searching for Moral

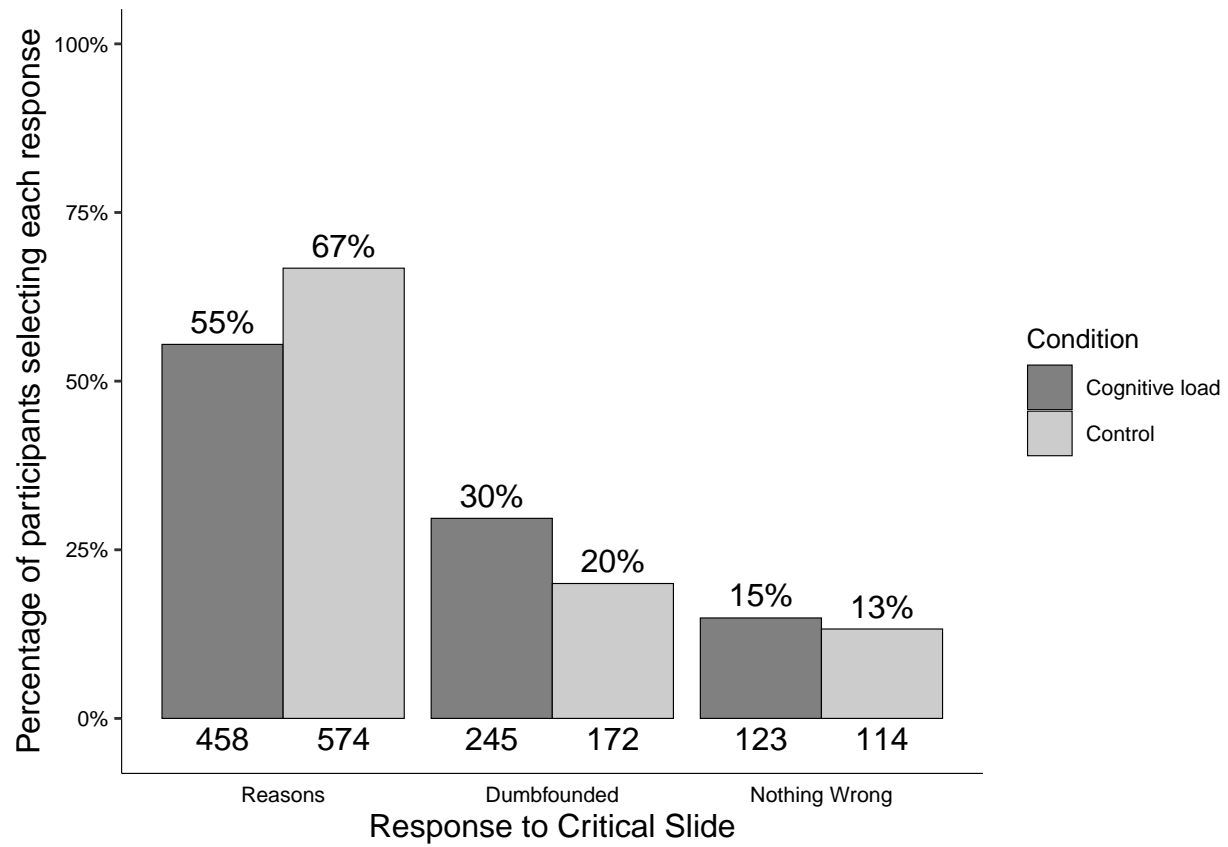
Dumbfounding: Identifying Measurable Indicators of Moral Dumbfounding. *Collabra: Psychology*, 3(1), 1–24. <https://doi.org/10.1525/collabra.79>

Table 1

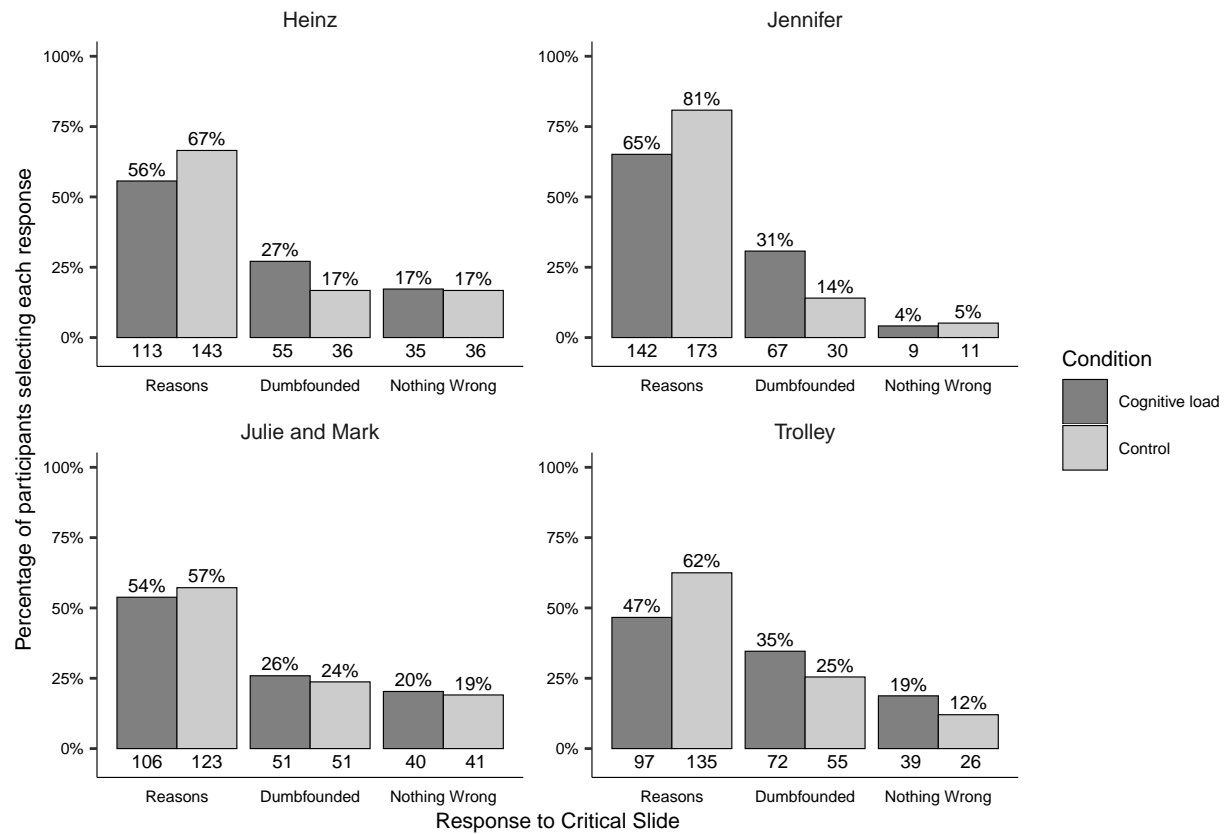
Study 6 – Observed counts, expected counts, and standardised residuals for each response to the critical slide depending on cognitive load (full sample)

		Cognitive Load	Control
Observed count	Reasons	458	574
	Dumbfounded	245	172
	Nothing Wrong	123	114
Expected count	Reasons	505.59	526.41
	Dumbfounded	204.3	212.7
	Nothing Wrong	116.11	120.89
Standardised residuals	Reasons	-4.76**	4.76**
	Dumbfounded	4.6**	-4.6**
	Nothing Wrong	0.97	-0.97

Note. * = sig. at $p < .05$; ** = sig. at $p < .001$

**Figure 1**

Study 6: Responses to critical slide depending on cognitive load

**Figure 2**

Study 6: Responses to critical slide and for the experimental group and the control group for each scenario