Supplementary Materials: Distancing and Moral Dumbfounding Pilot Studies

Cillian McHugh[1], Marek McGann[2], Eric R. Igou[1], and Elaine L. Kinsella [1]

[1]Department of Psychology, University of Limerick

[2]Department of Psychology, Mary Immaculate College

Author Note

Correspondence concerning this article should be addressed to Cillian McHugh, Email: cillian.mchugh@ul.ie

Abstract

Moral dumbfounding occurs when people maintain a moral judgment in the absence of supporting reasons. Drawing on dual-process approaches to moral judgment, one possible explanation for moral dumbfounding proposes that it occurs as a result of a conflict between intuitive and deliberative processes. Consistent with this explanation, previous research has shown that under manipulations designed to lead to more intuitive thinking rather than deliberative thinking (such as increased cognitive load), people are less likely to provide reasons for their judgments, and more likely to provide dumbfounded responses in a moral dumbfounding task. Building on this work the present research examines if dumbfounded responding can be reduced through experimental manipulations designed to facilitate deliberative thinking (over intuitive thinking). Drawing on construal-level theory, and the finding that distancing facilitates deliberative thinking, we predict that including a distancing manipulation in a moral dumbfounding task will increase reason-giving, and reduce dumbfounded responding. We propose a pre-registered study to test this prediction.

Keywords: moral dumbfounding, distancing, construal-level theory, dual-processes, reasons, intuitions

Supplementary Materials: Distancing and Moral Dumbfounding Pilot Studies

Below we present eight pilot studies to test the hypothesized relationship between distancing and dumbfounded responding. Recent work (McHugh et al., 2023) has proposed a conflict in dual-processes (e.g., Bonner & Newell, 2010) explanation of moral dumbfounding. According to this view, dumbfounding occurs when a habitual (moral judgment) response is in conflict with a deliberative response (providing reasons). This explanation is consistent with dual-process approaches to moral judgment (e.g., Bago & De Neys, 2019; Cushman, 2013; Greene, 2008), as well as with a unimodel (Kruglanski & Gigerenzer, 2011) and categorization (McHugh et al., 2022) approaches.

A key prediction of this explanation is that rates of reason-giving should be influenced by experimental manipulations that impact intuitive vs deliberative thinking. Previous work has demonstrated that inhibiting deliberative thinking through a cognitive load manipulation can reduce reason-giving, leading to higher rates of dumbfounding (McHugh et al., 2023). A corollary of this finding is that reason-giving should be increased under manipulations that encourage deliberative thinking. Drawing on construal-level theory (Förster et al., 2004; Liberman et al., 2002), we predict that increased distance will facilitate the identification of reasons, leading to lower levels of dumbfounded responding.

The Current Pilot Studies

We tested the feasibility of different manipulations of both temporal distance and psychological distance. A-priori power analysis indicated that in order to detect a large effect size ($V = .35$) with 80% power, a sample of $N = 79$ participants is required, in order to detect a medium effect size ($V = .21$) with 80% power a sample of $N = 218$ participants is required, and in order to detect, a small effect size ($V = . 07$) with 80% power a sample of $N = 1966$ is required. The pilot studies below are sufficiently powered to detect large effects only. Based on previous research investigating influences on moral dumbfounding we anticipate a small to medium effect size (e.g., McHugh et al., 2023, report $V = 0.12$). Despite being under-powered, these studies are

descriptively informative, and provide a preliminary estimate of the direction of any possible effect that would be expected from a higher powered study.

Pilot Studies 1 and 2 employ a manipulation of psychological distance compared against no manipulation. Pilot Studies 3, 4, 6, and 7 employ a temporal distance manipulation, comparing increased temporal distance against decreased temporal distance. For consistency across conditions, and in order to enhance the plausibility of the materials, participants in these studies were encouraged to think about the scenarios from the perspective of a third person (thus these studies additionally included a psychological distance manipulation, however this was kept constant across conditions). Pilot Studies 5 and 8 were also employed a temporal distance manipulation, and participants were encouraged to think about the scenarios from a first person perspective. Pilot Studies 1-5, and 8 used the *Julie and Mark* scenario while Pilot Studies 6 and 7 used the *Jennifer* scenario. Pilot Studies 1, 2, 3, 4, 5, 7, and 8, recorded dumbfounded responding using the 'critical slide' (McHugh et al., 2017, 2023), while Pilot Study 6 trialed a more continuous measure of dumbfounded responding. The designs and results of Pilot Studies 1-8 are summarized in Figure 1.

Figure 1

Overview of Pilot Studies 1-8.

Study	Distancing	Manipulation	Conditions	Perspective	Scenario	Measure	N	Result	*p*	direction
Pilot 1	Psychological	Anne vignette	Present vs Absent	First person	Julie & Mark	Critical Slide	120	No sig. Effect	.292	(predicted)
Pilot 2	Psychological	Anne vignette	Present vs Absent	First person vs third person	Julie & Mark	Critical Slide	104 / 76	No sig. Effect	.098 / .101	(predicted)
Pilot 3	Temporal	Anne vignette	Increased vs decreased	Third person	Julie & Mark	Critical Slide	117 / 71	Sig. Effect	.010 / .017	predicted
Pilot 4	Temporal	Anne vignette	Increased vs decreased	Third person	Julie & Mark	Critical Slide	165 / 77	No sig. Effect	.741 / .879	(predicted)
Pilot 5	Temporal	Imagined assignment	Increased vs decreased	First Person	Julie & Mark	Critical Slide	97 / 71	No sig. Effect	.432 / .464	(predicted)
Pilot 6	Temporal	Anne vignette	Increased vs decreased	Third person	Jennifer	Scale Measure	158 / 76	No sig. Effect	.306 / .278	(predicted)
Pilot 7	Temporal	Anne vignette	Increased vs decreased	Third person	Jennifer	Critical Slide	159 / 87	No sig. Effect	.868 / .472	(predicted)
Pilot 8	Temporal	Modified Scenario	Increased vs absent	First Person	Julie & Mark	Critical Slide	286 / 124	No sig. Effect	.482 / .681	(opposite)

Note. Sample sizes listed are the total samples that passed the attention checks, without exclusion based on responses to Need for Closure, followed by the sample size when excluding participants who failed the Need for Closure attention check.

Pilot Study 1 - Psychological Distance, "Julie and Mark"

The aim of Pilot Study 1 was to investigate if a psychological distance manipulation influenced participants' ability to justify their moral judgment. We also measured social desirability (Ballard, 1992; Crowne & Marlowe, 1960; Strahan & Gerbasi, 1972) as a potential correlate/moderator variable.

Pilot Study 1: Method

Pilot Study 1: Participants and Design

Pilot Study 1 was a between subjects design. The dependent variable was rates of providing reasons/dumbfounding (measured using to the critical slide with 3 response options: 1: reason-giving; 2: nothing-wrong; 3: dumbfounded response - an admission of not having reasons). The independent variable was psychological distancing with two levels: present and absent. To manipulate distancing participants were told that a philosophy student (Anne) had been asked to consider the moral scenario, and participants were asked to consider the reasons they might use to justify Anne's judgement. Social desirability (Ballard, 1992; Crowne & Marlowe, 1960; Strahan & Gerbasi, 1972) was included as an additional potential predictor variable.

A total sample of 120 participants (62 female, 58 male; $M_{age} = 38.02$, min = 22, max = 75, $SD = 11.9$) took part. Participants were recruited through MTurk. Participation was voluntary and participants were paid 0.50 US dollars for their participation. Participants were recruited from English speaking countries or from countries where residents generally have a high level of English (e.g., The Netherlands, Denmark, Sweden).

Pilot Study 1: Procedure and Materials

Data were collected using an online survey generated using Questback (Unipark, 2013). The survey opened with questions relating to basic demographics. Following this, participants were presented with two statements relating to the norm principle

(taken from McHugh et al., 2020; Royzman et al., 2015).[1] At this point, the distancing group were presented with an additional set of instructions prior to being presented with the the *Julie and Mark* scenario (taken from Haidt et al., 2000). Participants in the control group did not receive the additional instructions, and proceeded directly to the *Julie and Mark* scenario. The distancing instructions, and the *Julie and Mark* scenario read as follows:

Distancing Instructions.

Anne is a student of philosophy. She generally shows a good understanding of the subject matter, and this is reflected in her grades. Sometimes, however, she may adopt a position on an issue and struggle (or even fail) to defend it.

She is currently taking a course in ethics and has been asked to study the following scenario.

How should Anne judge the actions of the two people involved?

What reasons would you use to explain why she should make that judgement?

Julie and Mark Scenario.

Julie and Mark, who are brother and sister, are travelling together in France. They are both on summer vacation from college. One night they are staying alone in a cabin near the beach. They decide that it would be interesting and fun if they tried making love. At very least it would be a new experience for each of them. Julie was already taking birth control pills, but Mark uses a condom too, just to be safe. They both enjoy it, but they decide not to do it again. They keep that night as a special secret between them, which makes them feel even closer to each other.

––––––––

[1] Responses to this are not relevant for the research question of interest here and as such we do not discuss this in our analyses.

Participants rated how right or wrong the behavior of Julie and Mark was on a 7-point Likert scale (where, 1 = *Morally wrong*; 4 = *neutral*; 7 = *Morally right*), and were asked to provide reasons for their judgement (or what reasons they would use to explain to Anne why she should make that judgment). Participants then read a series of counter-arguments (developed by McHugh et al., 2017), which refuted commonly used justifications for rating the behavior as "wrong".

To measure dumbfounding we used the *critical slide* developed by McHugh et al. (2017). This included a statement defending the behavior and a question asking how the behavior could be wrong ("Julie and Mark's behaviour did not harm anyone, how can there be anything wrong with what they did?"). There were three possible answer options: (a) "It's wrong and I can provide a valid reason" (reason-giving), (b) "It's wrong but I can't think of a reason" (dumbfounding), and (c) "There is nothing wrong" (nothing-wrong). The order of these response options was randomized. The selecting of option (b), the admission of not having reasons, was taken to be a dumbfounded response, and we note that this measure provides a conservative measure of dumbfounded responding (McHugh et al., 2017). Participants who selected (a) were prompted to type a reason once they progressed to the next page.

Following the critical slide rated the behavior again on a 7-point Likert scale (where, 1 = *Morally wrong*; 4 = *neutral*; 7 = *Morally right*). They then responded to the credulity check questions devised by Royzman et al. (2015), and answered the three questions relating to the application of the harm principle (McHugh et al., 2020).[2] Finally participants completed the short version of the Marlowe-Crowne (Crowne & Marlowe, 1960) social desirability scale (devised by Strahan & Gerbasi, 1972; see also Ballard, 1992). This consisted of ten questions (e.g., "There have been occasions when I took advantage of someone.", "I never resent being asked to return a favor.") to which participants selected "True" or "False".

_____

[2] As with the norm principles questions above responses to these questions are not relevant for the research question of interest here and as such we do not discuss this in our analyses.

Pilot Study 1: Results

Figure 2

Study 1: Responses to critical slide for the experimental group (N = 60) and the control group (N = 60); (error bars represent standard error of the proportion)
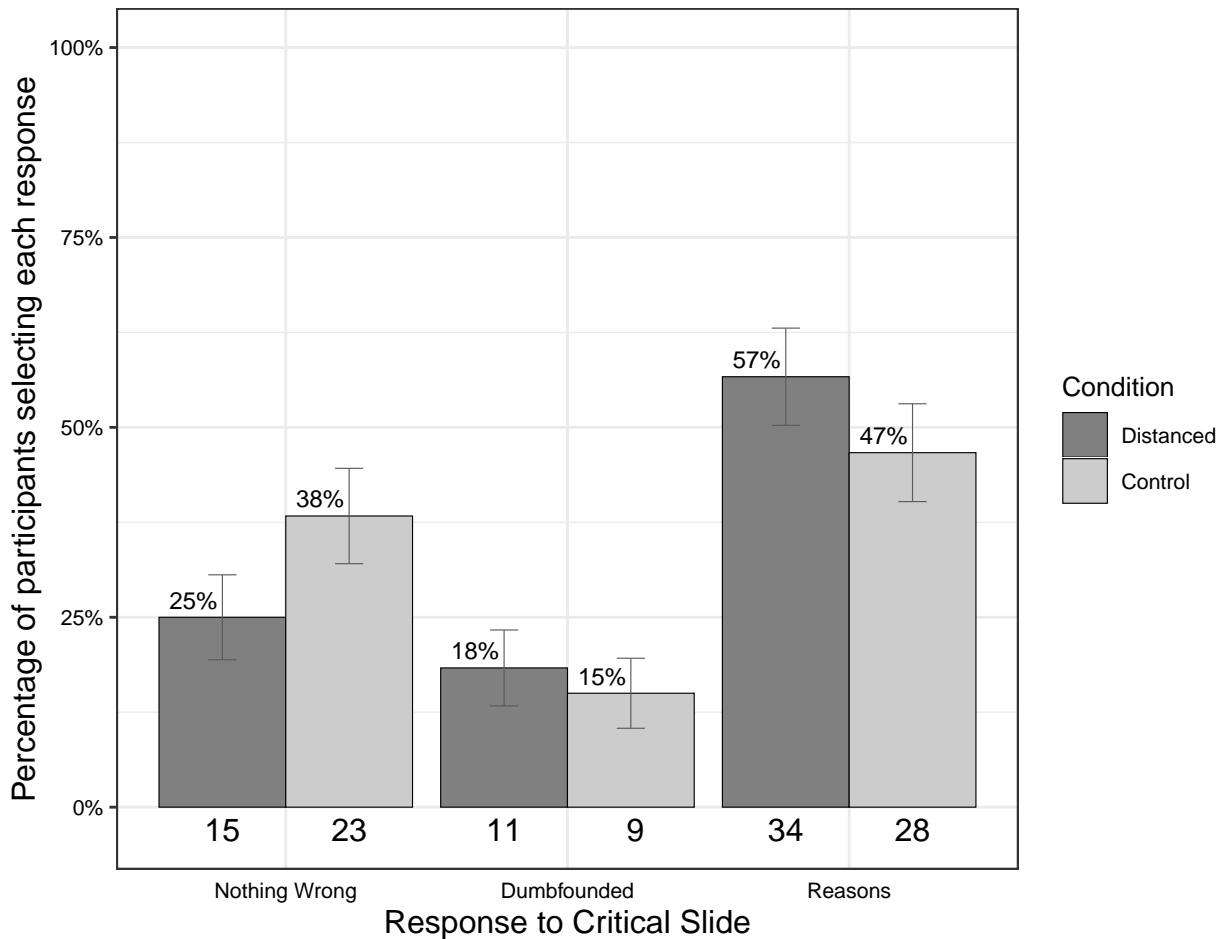
Judgments

Seventy six participants (63.87%) rated the behavior of Julie and Mark as wrong initially, and Seventy two participants (60.5%) rated the behavior as wrong at the end of the task. There was no significant difference between initial ratings ($M = 2.9$, $SD = 2.1$) and revised ratings ($M = 2.9$, $SD = 2$), $t(118) = -0.72$, $p = .476$, $d = 0.07$.
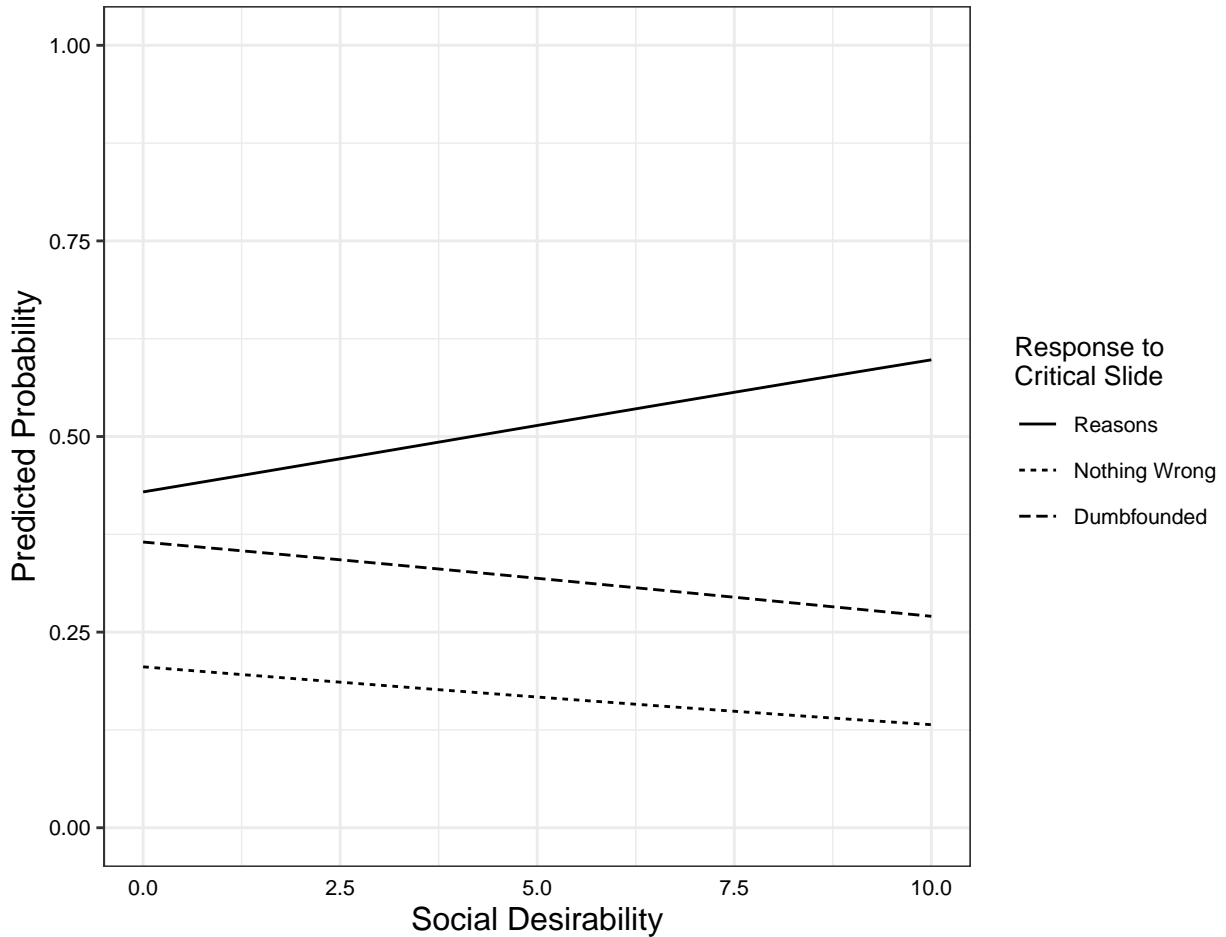
Distancing and Judgments

There was no difference in initial judgement depending on distance manipulation: $t(113.8) = -0.93$, $p = .356$, $d = 0.17$, ($M_{distanced} = 2.7$, $SD_{distanced} = 2.2$, $M_{control} = 3.1$, $SD_{control} = 1.9$). There was no difference in revised judgement depending

Figure 3

Study 1: Probability of selecting each response to the critical slide depending on Social Desirability

on distance manipulation: $t(113.99) = $ -1.23, $p = $ .223, $d = 0.22$, ($M_{\text{distanced}} = 2.7$, $SD_{\text{distanced}} = 2.2$, $M_{\text{control}} = 3.2$, $SD_{\text{control}} = 1.9$).

Distancing and Dumbfounding

There was no significant association between experimental condition and response to the critical slide, $\chi^2(2, N = 119) = 2.286$, $p = .319$, $V = 0.14$, the observed power was 0.25. The responses to the critical slide for the experimental group ($N = 59$) and the control group ($N = 60$) are displayed in Figure 2.

Social Desirability and Dumbfounding

There was no significant association between Social Desirability and response to the critical slide $\chi^2(2, N = 119) = 0.84$, $p = .656$, The observed power was 0.12. The

predicted probabilies of each response depending on social desirability are displayed in Figure 3.

Pilot Study 2 - Psychological Distance, "Julie and Mark"

As with Pilot Study 1, the aim of Pilot Study 2 was to investigate if a psychological distance manipulation influenced participants' ability to justify their moral judgment. There were two changes from Pilot Study 1. First, we modified our distancing manipulation to include an explicit instruction to think about the scenario from the perspective of a third party. In addition to recording social desirability (Ballard, 1992; Crowne & Marlowe, 1960; Strahan & Gerbasi, 1972), Pilot Study 2 additionally included need for closure (Kruglanski et al., 2013; Kruglanski, 2013; Kruglanski & Webster, 1996) and the short form of the cognitive reflection test (Frederick, 2005; Thomson & Oppenheimer, 2016; Toplak et al., 2011) as potential correlate/moderator variables.

Pilot Study 2: Method

Pilot Study 2: Participants and Design

Pilot Study 2 was a between subjects design. As in Pilot Study 1, the dependent variable was rates of providing reasons/dumbfounding - measured using to the critical slide (with the same 3 response options: 1: reason-giving; 2: nothing-wrong; 3: dumbfounded response - an admission of not having reasons). The independent variable was psychological distancing with two levels: present and absent. The distancing manipulation was the same as in Pilot Study 1 (the *Anne* vignette) with the inclusion of an explicit instruction to consider the scenario from Anne's perspective. Social desirability (Ballard, 1992; Crowne & Marlowe, 1960; Strahan & Gerbasi, 1972), need for closure (Kruglanski et al., 2013; Kruglanski, 2013; Kruglanski & Webster, 1996) cognitive reflection (Frederick, 2005; Thomson & Oppenheimer, 2016; Toplak et al., 2011) were recorded as potential correlate/moderator variables.

A total sample of 104 participants (49 female, 55 male; $M_{age} = 37.46$, min = 19, max = 83, $SD = 12.2$) took part. The measure of need for closure includes a "lie score", whereby if participants score above a threshold on a combination of specific items they

are deemed to be lying (example lie score items include claiming to never have been late for an appointment, or never having met someone the didn't like). Following the removal of participants who scored above the lie score threshold, we were left with a sample of 76 participants (37 female, 39 male; M_{age} = 38.67, min = 19, max = 83, SD = 12.56) who were eligible for analysis. Participants were recruited through MTurk in the same way as in Pilot Study 1 (same payment amount, same country selection).

Pilot Study 2: Procedure and Materials

The procedure and materials were similar to Pilot Study 1 with a change to the distance manipulation and the inclusion of additional measures. The distance manipulation for Pilot Study 2 included an explicit instruction to think about the scenario from the perspective of a third party. The revised manipulation read as follows:

Distancing Instructions.

Anne is a student of philosophy. She generally shows a good understanding of the subject matter, and this is reflected in her grades. Sometimes, however, she adopts a position on an issue in class and struggles (or fails) to defend it when challenged by others.

She is currently taking a course in ethics and has been asked to study the following scenario.

While reading the story on the next page, try to imagine how the philosophy student Anne will judge the actions of the two people.

In particular try to think about reasons she may use to defend her judgement.

Try to think about the story from Anne's perspective rather than your own.

In addition to social desirability, we additionally measured need for closure (Kruglanski et al., 2013; Kruglanski, 2013; Kruglanski & Webster, 1996) cognitive reflection (Frederick, 2005; Thomson & Oppenheimer, 2016; Toplak et al., 2011). The

Need for Closure Scale contains 47 questions (e.g., "I'd rather know bad news than stay in a state of uncertainty.") to which participants respond on a 6 point Likert scale, where 1 = *strongly disagree*, and 6 = *strongly agree.* The CRT is a brief test of analytical thinking. It contains three questions, each of which has an answer that seems intuitively correct, but is actually wrong (e.g., If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?)

Pilot Study 2: Results

Below we present two sets of results. First we present the results for the full sample, second we present the results for the sample with exclusions based on participants' "lie score".

Judgments (no exclusions)

Seventy two participants (69.23%) rated the behavior of Julie and Mark as wrong initially, and Sixty six participants (63.46%) rated the behavior as wrong at the end of the task. There was no significant difference between initial ratings ($M = 2.5$, $SD = 2$) and revised ratings ($M = 2.4$, $SD = 2$), $t(89) = 0$, $p = 1.000$, $d = 0$.

Distancing and Judgments (no exclusions)

There was no difference in initial judgement depending on distance manipulation: $t(96.67) = 0.35$, $p = .729$, $d = 0.07$, ($M_{\text{distanced}} = 2.5$, $SD_{\text{distanced}} = 2$, $M_{\text{control}} = 2.4$, $SD_{\text{control}} = 2.1$). There was no difference in revised judgement depending on distance manipulation: $t(86.71) = -0.27$, $p = .789$, $d = 0.06$, ($M_{\text{distanced}} = 2.3$, $SD_{\text{distanced}} = 1.9$, $M_{\text{control}} = 2.5$, $SD_{\text{control}} = 2.2$).

Distancing and Dumbfounding (no exclusions)

There was no significant association between experimental condition and response to the critical slide, $\chi^2(2, N = 0) = 4.636$, $p = .098$, $V = 0.21$, the observed power was 0.47. The responses to the critical slide for the experimental group ($N = 52$) and the control group ($N = 52$) are displayed in Figure 4.

Figure 4

Pilot Study 2: Responses to critical slide for the experimental group (N = 52) and the control group (N = 52); (No exclusions; error bars represent standard error of the proportion)
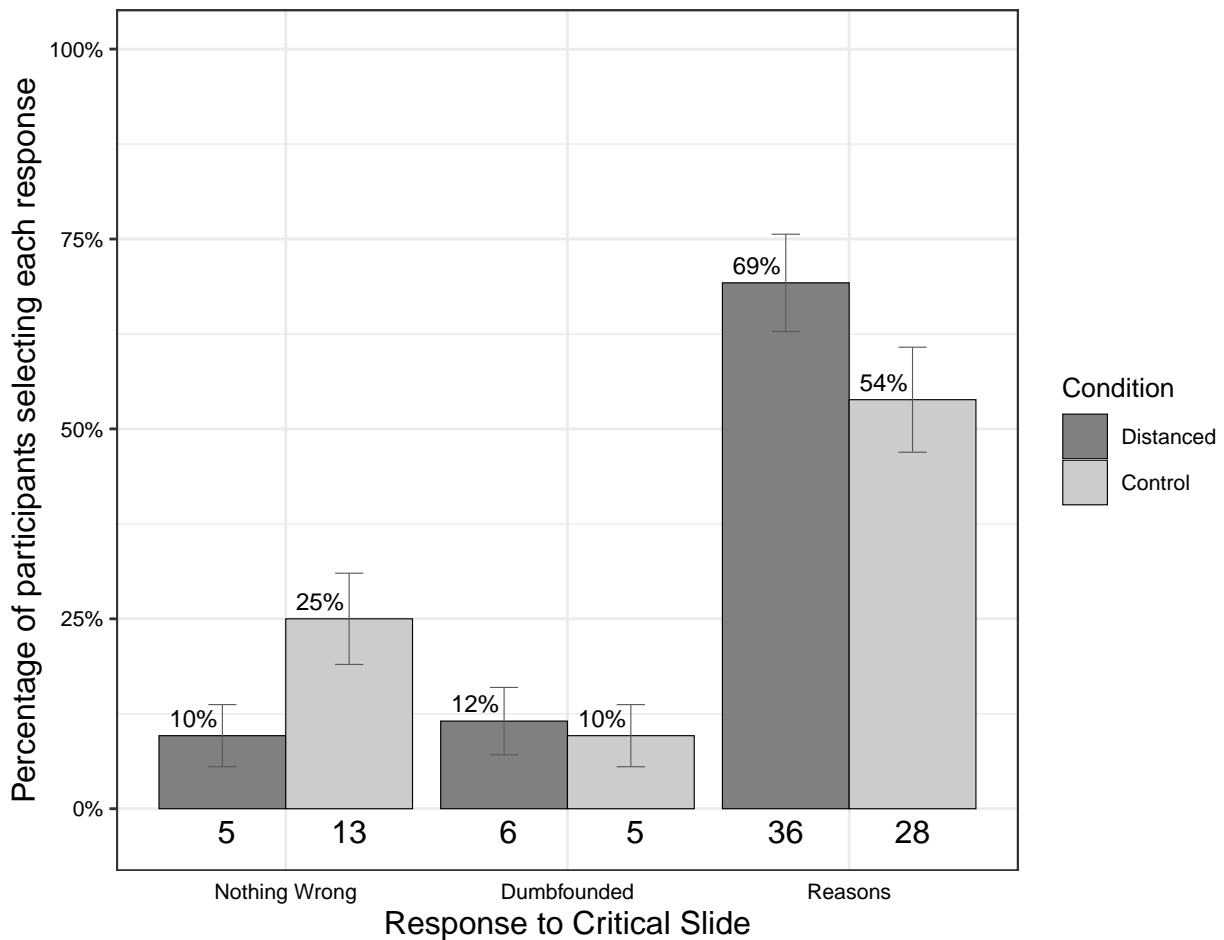
### Individual Difference Predictors (no exclusions)

A multinomial logistic regression was conducted to test if the individual difference measures predicted dumbfounding. Overall the model did not significantly predict responses to the critical slide $\chi^2(6, N = 104) = 2.8$, $p = .834$, The observed power was 0.3.

Judgments (with exclusions)

Fifty one participants (67.11%) rated the behavior of Julie and Mark as wrong initially, and Forty seven participants (61.84%) rated the behavior as wrong at the end of the task. There was no significant difference between initial ratings ($M = 2.5$, $SD =$

Figure 5

Pilot Study 2: Responses to critical slide for the experimental group (N = 39) and the control group (N = 37); (With exclusions; error bars represent standard error of the proportion)
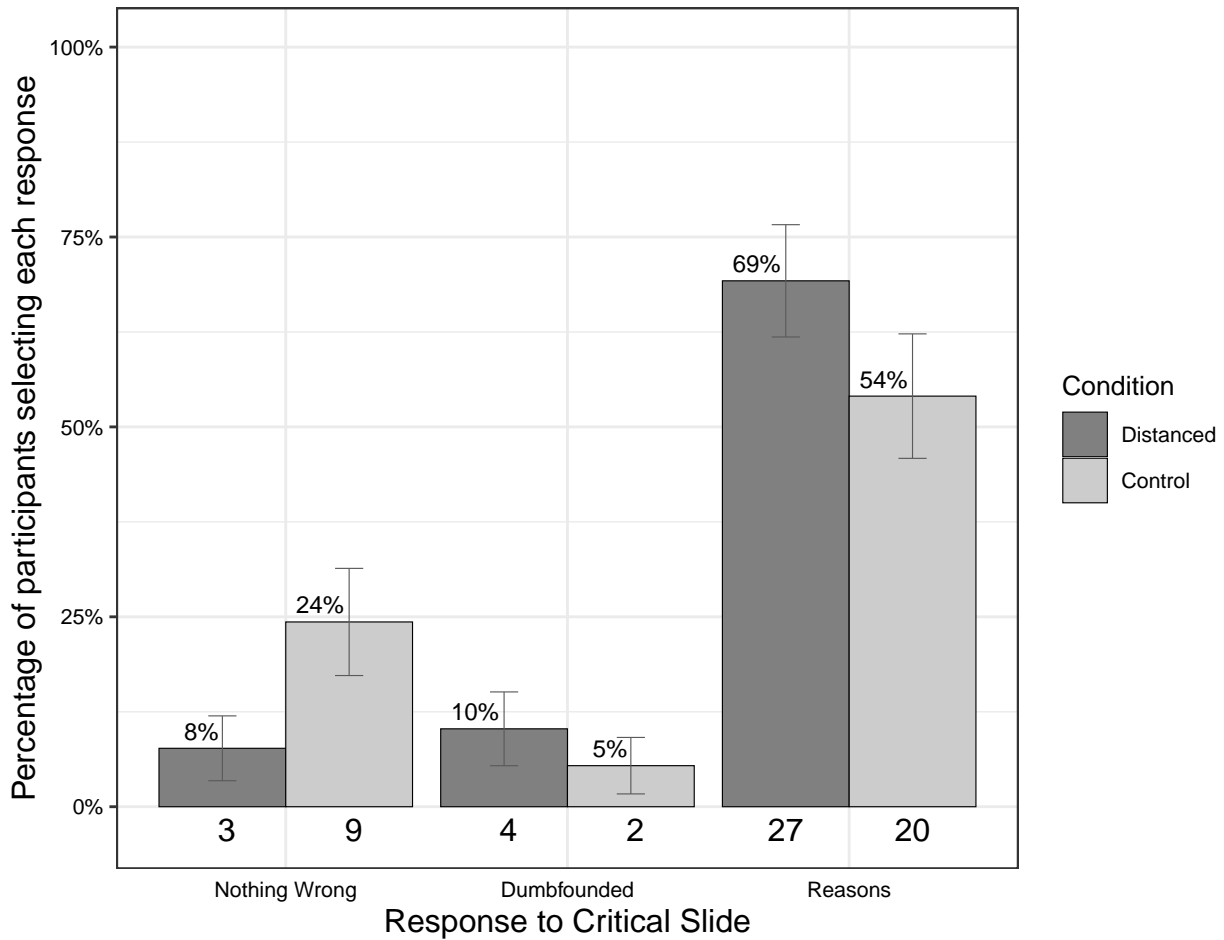
2.1) and revised ratings ($M = 2.4$, $SD = 2.1$), $t(61) = 0.65$, $p = .517$, $d = 0.08$.

Distancing and Judgments (with exclusions)

There was no difference in initial judgement depending on distance manipulation: $t(66.79) = 0.22$, $p = .827$, $d = 0.05$, ($M_{\text{distanced}} = 2.6$, $SD_{\text{distanced}} = 2$, $M_{\text{control}} = 2.5$, $SD_{\text{control}} = 2.2$). There was no difference in revised judgement depending on distance manipulation: $t(56.44) = -0.61$, $p = .543$, $d = 0.16$, ($M_{\text{distanced}} = 2.2$, $SD_{\text{distanced}} = 1.8$, $M_{\text{control}} = 2.5$, $SD_{\text{control}} = 2.3$).

Distancing and Dumbfounding (with exclusions)

There was no significant association between experimental condition and response to the critical slide, $\chi^2(2, N = 0) = 4.581$, $p = .101$, $V = 0.25$, the observed power was 0.47. The responses to the critical slide for the experimental group ($N = 39$) and the control group ($N = 37$) are displayed in Figure 5.

Individual Difference Predictors (with exclusions)

A multinomial logistic regression was conducted to test if the individual difference measures predicted dumbfounding. Overall the model did not significantly predict responses to the critical slide $\chi^2(6, N = 76) = 5.36$, $p = .499$, The observed power was 0.53.

Pilot Study 3 - Temporal Distance, "Julie and Mark"

For both Pilot Studies 1 and 2 there was no significant association between response to the critical slide and the psychological distance manipulation. Pilot Study 3 was designed to test an alternative distance manipulation, temporal distance. The aim of Pilot Study 3 was to investigate if manipulating temporal distance influenced participants' ability to justify their moral judgment. There were two changes from Pilot Study 1. First, we modified our distancing manipulation to include an explicit instruction to think about the scenario from the perspective of a third party. In addition to recording social desirability (Ballard, 1992; Crowne & Marlowe, 1960; Strahan & Gerbasi, 1972), Pilot Study 2 additionally included need for closure (Kruglanski et al., 2013; Kruglanski, 2013; Kruglanski & Webster, 1996) and the short form of the cognitive reflection test (Frederick, 2005; Thomson & Oppenheimer, 2016; Toplak et al., 2011) as potential correlate/moderator variables.

Pilot Study 3: Method

Pilot Study 3: Participants and Design

As with previous studies, Pilot Study 3 was a between subjects design. The dependent variable was rates of providing reasons/dumbfounding - measured using to the critical slide. The independent variable was temporal distance with two levels:

increased and decreased. All participants read a vignette similar to the distancing manipulation used in Pilot Studies 1 and 2 and temporal distance was manipulated by varying the deadline by which the assignment would need to be completed by - either five weeks in the future (increased) or by 5pm today (decreased). Need for closure (Kruglanski et al., 2013; Kruglanski, 2013; Kruglanski & Webster, 1996) and cognitive reflection (Frederick, 2005; Thomson & Oppenheimer, 2016; Toplak et al., 2011) were recorded as potential correlate/moderator variables. Social desirability was not recorded.

A total sample of 105 participants (47 female, 58 male; $M_{\text{age}} = 37.51$, min $= 21$, max $= 73$, $SD = 12.5$) took part. As in Pilot Study 2, participants who scored above the threshold on the need for closure lie score were removed, and this left a sample of 71 participants (35 female, 36 male; $M_{\text{age}} = 40.7$, min $= 21$, max $= 73$, $SD = 13.4$). Participants were recruited through MTurk in the same way as in Pilot Study 1 (same payment amount, same country selection).

Pilot Study 3: Procedure and Materials

The procedure and materials were similar to Pilot Study 3 with a change to the distance manipulation and the removal of the social desirability measure. The same moral scenario ("Julie and Mark") was used. The revised manipulation read as follows:

Increased Temporal Distance.

Anne is a student of philosophy. She generally shows a good understanding of the subject matter, and this is reflected in her grades. Sometimes, however, she may adopt a position on an issue and struggle (or even fail) to defend it.

She is currently taking a course in ethics. For this ethics course, Anne and has been asked to study the following scenario. She must identify if the behaviour described is right or wrong, and provide reasons for her judgement. **She must complete and submit this assignment within the next five weeks**.

While reading the story on the next page, try to think about the story from Anne's perspective rather than your own.

Decreased Temporal Distance:.

Anne is a student of philosophy. She generally shows a good understanding of the subject matter, and this is reflected in her grades. Sometimes, however, she may adopt a position on an issue and struggle (or even fail) to defend it.

She is currently taking a course in ethics. For this ethics course, Anne and has been asked to study the following scenario. She must identify if the behaviour described is right or wrong, and provide reasons for her judgement. **She must complete and submit this assignment today**.

While reading the story on the next page, try to think about the story from Anne's perspective rather than your own.

Pilot Study 3: Results

As for Pilot Study 2, we present two sets of results. First we present the results for the full sample, second we present the results for the sample with exclusions based on participants' "lie score".

Judgments (no exclusions)

Sixty three participants (60%) rated the behavior of Julie and Mark as wrong initially, and Sixty four participants (60.95%) rated the behavior as wrong at the end of the task. There was no significant difference between initial ratings ($M = 2.8$, $SD = 2$) and revised ratings ($M = 2.9$, $SD = 2$), $t(104) = -0.58$, $p = .566$, $d = 0.06$.

Distancing and Judgments (no exclusions)

There was no difference in initial judgement depending on distance manipulation: $t(102.72) = -1$, $p = .318$, $d = 0.2$, ($M_{increased} = 2.6$, $SD_{increased} = 2$, $M_{decreased} = 3$, $SD_{decreased} = $ NA). There was no difference in revised judgement depending on distance manipulation: $t(102.97) = -1.32$, $p = .190$, $d = 0.26$, ($M_{increased} = 2.6$, $SD_{increased} = 2$, $M_{decreased} = 3.2$, $SD_{decreased} = 2$).

Figure 6

Pilot Study 3: Responses to critical slide for the increased temporal distance group (N = 53) and the reduced temporal distance group (N = 52); (No exclusions; error bars represent standard error of the proportion)
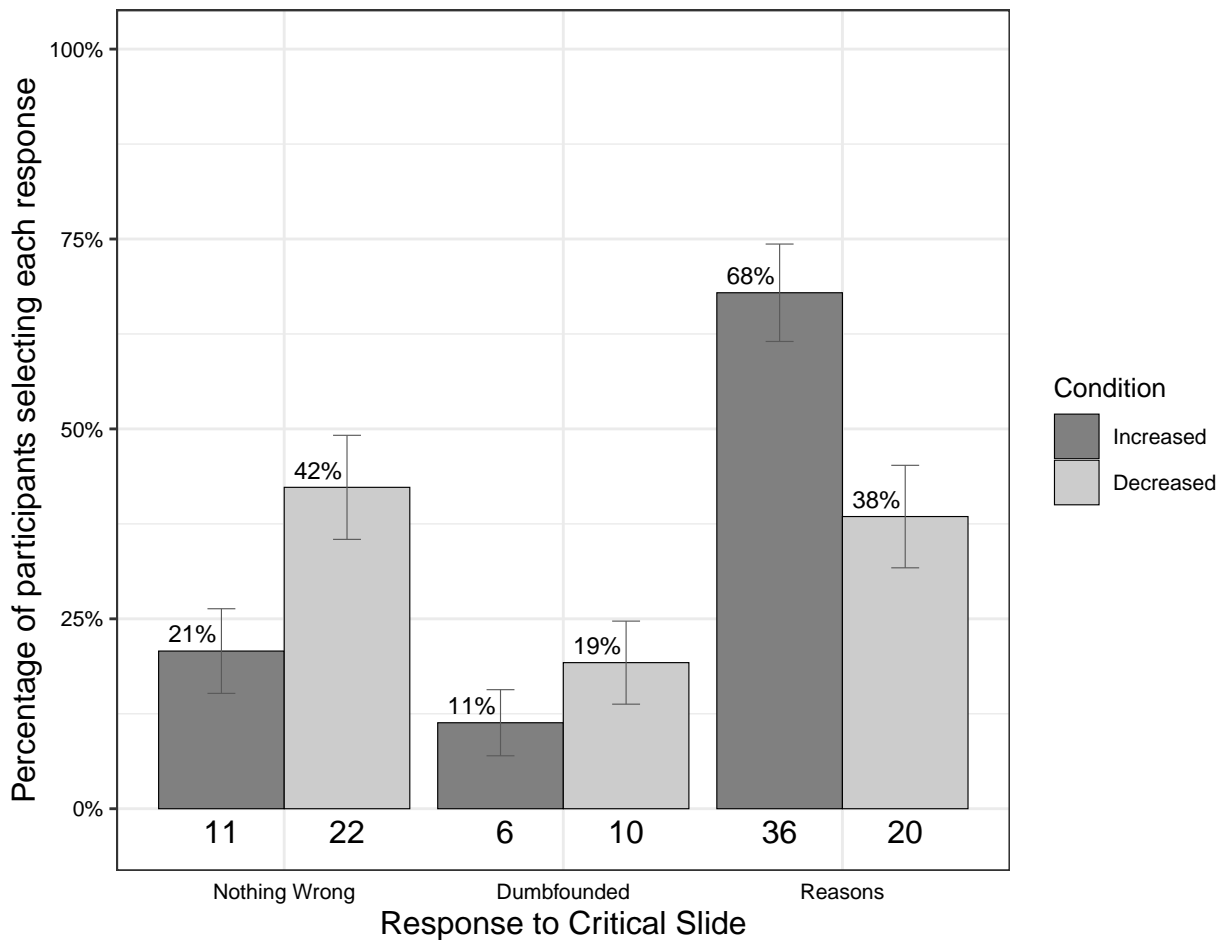
Distancing and Dumbfounding (no exclusions)

There was no significant association between experimental condition and response to the critical slide, $\chi^2(2, N = 0) = 9.229$, $p = .010$, $V = 0.3$, the observed power was 0.78. The responses to the critical slide for the experimental group ($N = 53$) and the control group ($N = 52$) are displayed in Figure 4.

Individual Difference Predictors (no exclusions)

A multinomial logistic regression was conducted to test if the individual difference measures predicted dumbfounding. Overall the model significantly predicted responses to the critical slide $\chi^2(4, N = 105) = 13.22$, $p = .010$, The observed power

was 0.91. The model explained between 4.09% (Cox and Snell R square) and 5.75% (Nadelkerke R squared) of the variance in responses to the critical slide. Need for closure was a significant predictor in the model. As Need for closure increased, participants were significantly more likely to provide reasons than to present as dumbfounded, Wald = 7.25, $p = .007$, odds ratio = 0.94, 95% CI [0.9, 0.98], or to select nothing-wrong , Wald = 5.66, $p = .017$, odds ratio = 0.97, 95% CI [0.94, 0.99].

Figure 7

Study 1: Probability of selecting each response to the critical slide depending on Social Desirability
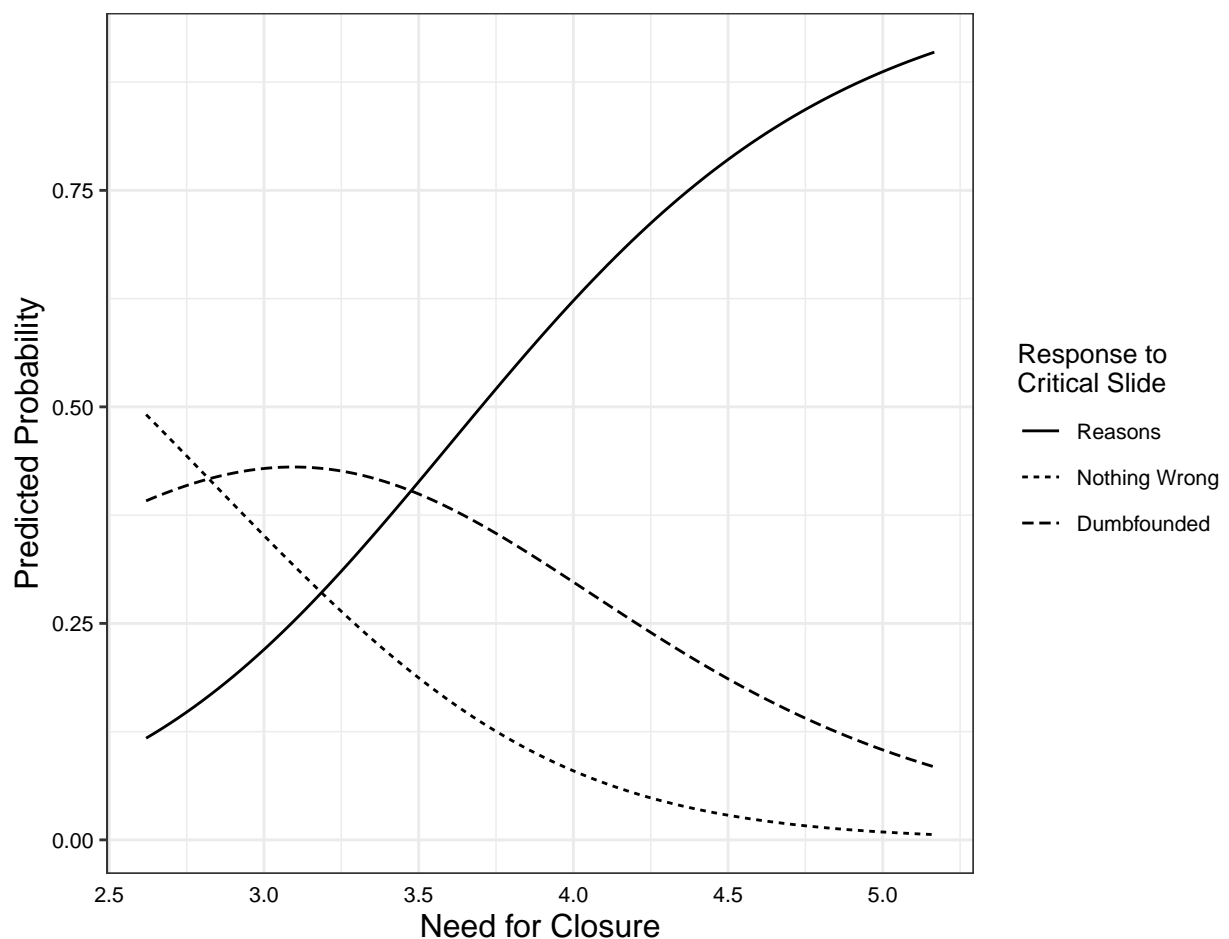
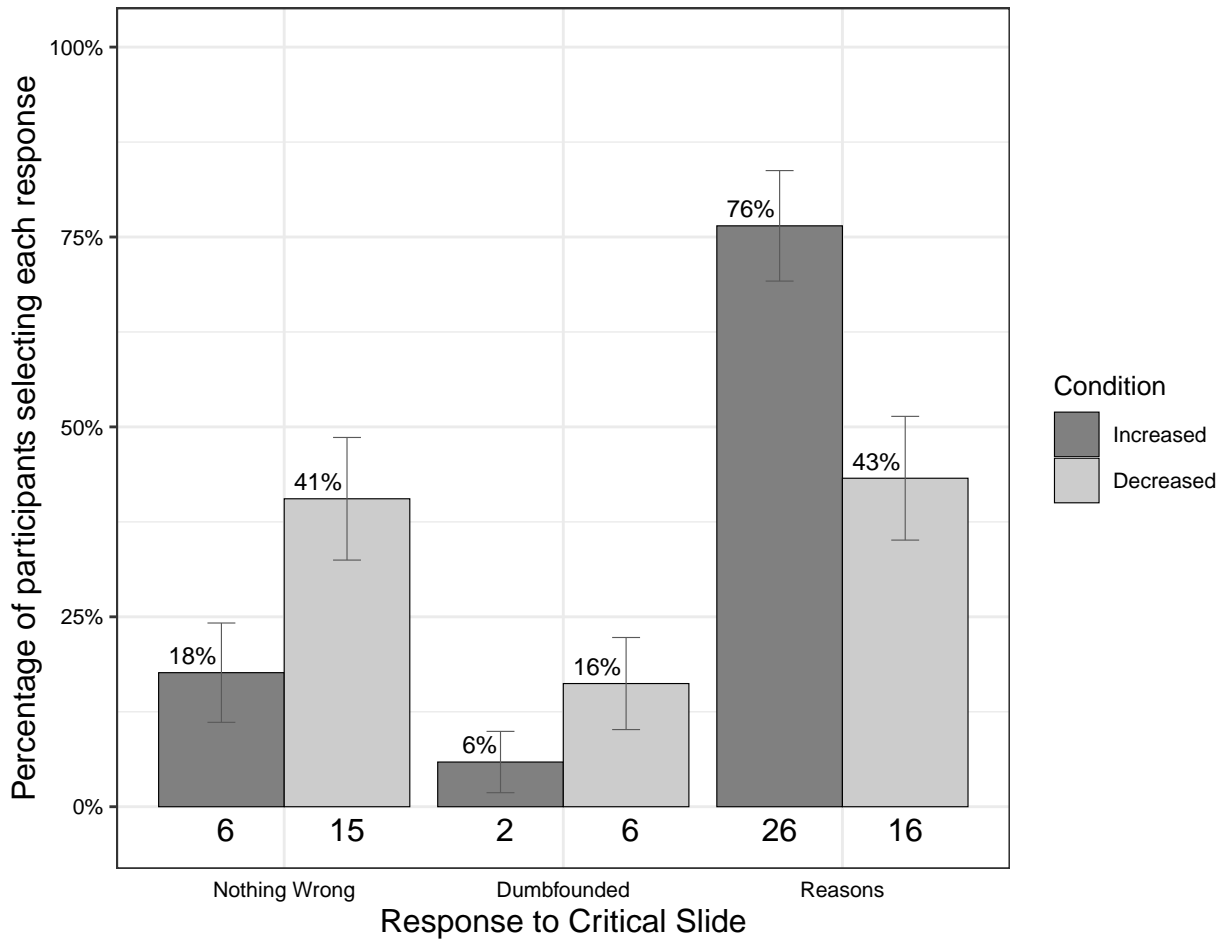### Judgments (with exclusions)

Fifty participants (70.42%) rated the behavior of Julie and Mark as wrong initially, and Forty nine participants (69.01%) rated the behavior as wrong at the end of the task. There was no significant difference between initial ratings ($M = 2.6$, $SD = 2$)

Figure 8

Pilot Study 3: Responses to critical slide for the increased temporal distance group (N = 34) and the reduced temporal distance group (N = 37); (With exclusions; error bars represent standard error of the proportion)

and revised ratings ($M = 2.7$, $SD = 2$), $t(70) = -0.83$, $p = .410$, $d = 0.1$.

Distancing and Judgments (with exclusions)

There was no difference in initial judgement depending on distance manipulation: $t(68.71) = -1.63$, $p = .107$, $d = 0.39$, ($M_{\text{distanced}} = 2.2$, $SD_{\text{distanced}} = 1.8$, $M_{\text{decreased}} = 2.9$, $SD_{\text{decreased}} = 2.1$). Revised judgments were significantly more severe in the increased temporal distance condition ($M_{\text{increased}} = 2.2$, $SD_{\text{increased}} = 1.9$) than in the reduced temporal distance condition ($M_{\text{decreased}} = 3.1$, $SD_{\text{decreased}} = 2.1$), $t(69) = -2.03$, $p = .046$, $d = 0.48$,

Distancing and Dumbfounding (with exclusions)

There was no significant association between experimental condition and response to the critical slide, $\chi^2(2, N = 0) = 8.126$, $p = .017$, $V = 0.34$, the observed power was 0.72. The responses to the critical slide for the experimental group $(N = 0)$ and the control group $(N = 0)$ are displayed in Figure 8.

Individual Difference Predictors (with exclusions)

A multinomial logistic regression was conducted to test if the individual difference measures predicted dumbfounding. Overall the model did not significantly predict responses to the critical slide $\chi^2(4, N = 71) = 6.77$, $p = .149$, The observed power was 0.64.

Template bits

Level 2 Heading: Flush Left, Bold, Title Case

Subsections of the introduction have level 2 headings. A paragraph after a level 2 Heading is on a new line. Regular paragraphs are indented, flush left, and double-spaced.

You do not need to put text after a heading. You can put a higher-level heading directly underneath if you want.

A Level 2 Heading Without Text Below It

Level 3 Heading: Flush Left, Bold Italic, Title Case

Subsections of a level 2 heading are placed under level 3 headings.

Another Level 3 Heading

Level 4 Heading. A level 4 heading should be indented, flush left, bold, title case, and end with a period. A paragraph after a level 4 or 5 heading is on a new line in this markdown document but will appear as if it were in the same paragraph when rendered. You need at least one paragraph after a level 4 or 5 heading. If you forget the period at the end of the level 4 or 5 heading, it will be added automatically. A period will not be added if the heading ends with a question mark or an exclamation point.

Subsequent paragraphs go on their own lines.

Level 5 Heading. A level 5 heading should be indented, flush left, bold italic, title case, and end with a period. Notice that there was no period after this level 5 heading in the markdown document, but it does appear in the rendered document.

Subsequent paragraphs go on their own lines.

How to Cite References

I am going to cite a reference here in square brackets (**CameronTrivedi2013?**). This reference was in my bibliography file. Here are some variations on parenthetical citations:

- Page references (or any other suffixes are placed after the reference. If you want a comma, you'll need to insert it yourself: (**CameronTrivedi2013?**)

- Prefixes (with or without a comma) are placed before the reference: (e.g., **CameronTrivedi2013?**)

- 2 or more citations separated by a semicolon (**CameronTrivedi2013?**; **cohen2003applied?**)

- Any prefixes or suffixes needing a literal semicolon will confuse Quarto (actually Pandoc). To make it clear that you need to print a semicolon, put a backslash before the semicolon: [FOIL; (**CameronTrivedi2013?**)]

  Text references are possible, too.

- (**CameronTrivedi2013?**) said some interesting things.

- (**cohen2003applied?**) said specific things on specific pages.

- Place the reference's year by itself with a minus sign: (**CameronTrivedi2013?**)

Hypotheses, Aims, and Objectives

The last paragraph of the introduction usually states the specific hypotheses of the study, often in a way that links them to the research design.

Method

General remarks on method. This paragraph is optional.

Not all papers require each of these sections. Edit them as needed. Consult the Journal Article Reporting Standards for what is needed for your type of article.

Participants

Who are they? How were they recruited? Report criteria for participant inclusion and exclusion. Perhaps some basic demographic stats are in order. A table is a great way to avoid repetition in statistical reporting.

Measures

This section can also be titled **Materials** or **Apparatus**. Whatever tools, equipment, or measurement devices used in the study should be described.

Measure A

Describe Measure A.

Measure B

Describe Measure B.

Procedure

What did participants do?

How are the data going to be analyzed?

Results

Descriptive Statistics

Here we describe the basic characteristics of our primary variables.

Let's make a figure. A reference label for a figure in APA format must have the prefix `apafg-`. This is different from the usual Quarto prefix `fig-`.

To refer to any figure or table, put the chunk label in curly braces. For example, see Figure 9. In Figure 10, we import an image.

We can make a table the same way as a figure except that the label prefix is `apatb-`. Again, this is different from the usual quarto prefix `tbl-`, which will put the table table caption in the wrong place and with non-APA formatting.

To refer to this table in text, put the table's reference label in curly braces like so: As seen in Table 1, there is not much information.

What if you want the tables and figures to be at the end of the document? In the .pdf format, you can set the `floatsintext` option to false. For .html and .docx documents, there is not yet an automatic way to put tables and figures at the end. You can, of course, just put them all at the end, in order. The reference labels will work no matter where they are in the text.
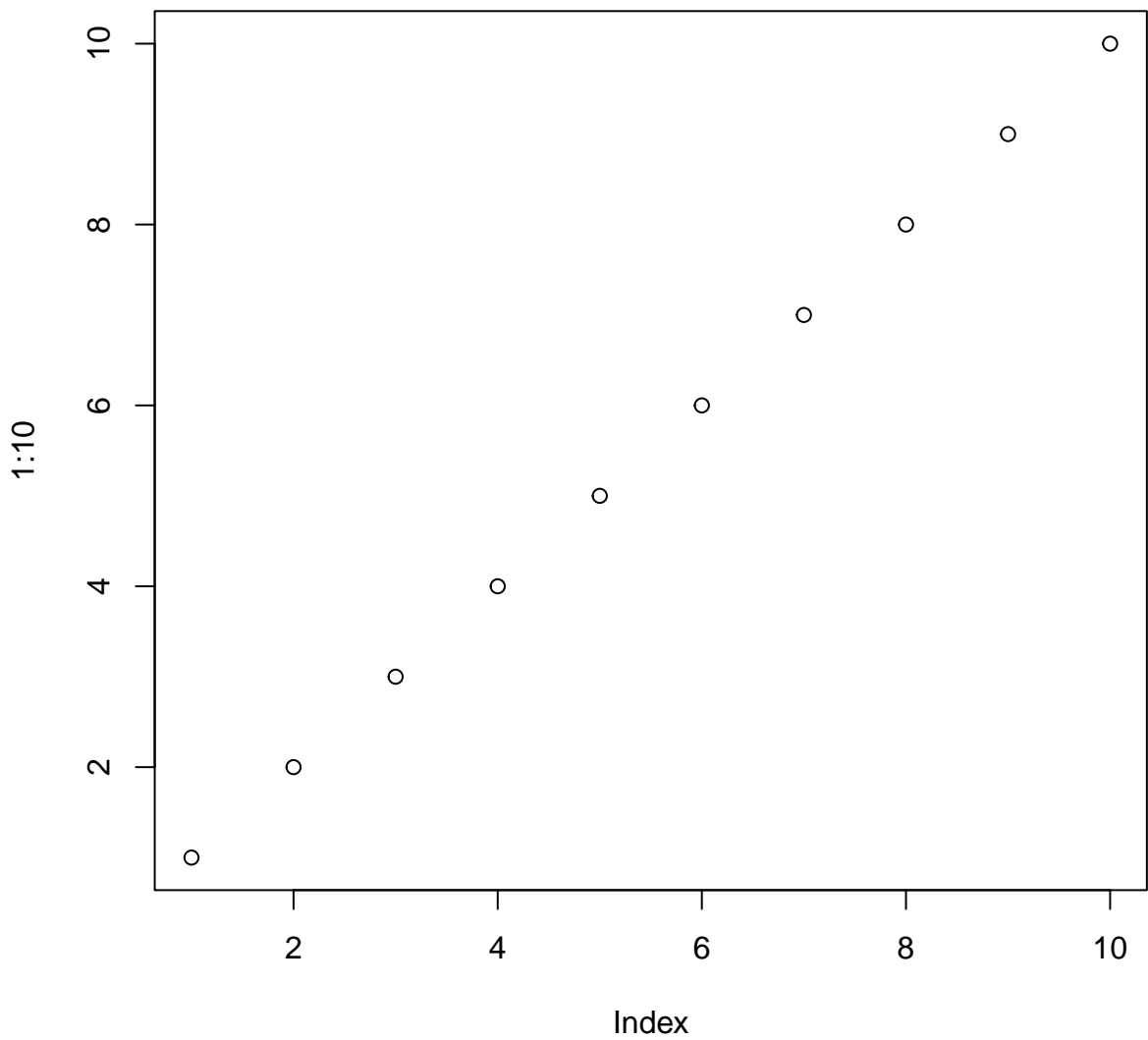
Discussion

Describe results in non-statistical terms.

Figure 9

This is the figure caption.

Note. This is a note below the figure.

Limitations and Future Directions

Every study has limitations. Based on this study, some additional steps might include...

Conclusion

Let's sum this up.

Figure 10

This is an imported graphic.

Note. My note.

References

Bago, B., & De Neys, W. (2019). The intuitive greater good: Testing the corrective dual process model of moral cognition. *Journal of Experimental Psychology: General, 148*(10), 1782–1801. https://doi.org/10.1037/xge0000533

Ballard, R. (1992). Short forms of the marlowe-crowne social desirability scale. *Psychological Reports, 71*(3), 1155–1160. https://doi.org/10.2466/pr0.1992.71.3f.1155

Bonner, C., & Newell, B. R. (2010). In conflict with ourselves? An investigation of heuristic and analytic processes in decision making. *Memory & Cognition, 38*(2), 186–196. https://doi.org/10.3758/MC.38.2.186

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*(4), 349–354. https://doi.org/10.1037/h0047358

Cushman, F. A. (2013). Action, outcome, and value a dual-system framework for morality. *Personality and Social Psychology Review, 17*(3), 273–292. https://doi.org/10.1177/1088868313495594

Förster, J., Friedman, R. S., & Liberman, N. (2004). Temporal construal effects on abstract and concrete thinking: Consequences for insight and creative cognition. *Journal of Personality and Social Psychology*, 177–189.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*(4), 25–42. https://ideas.repec.org/a/aea/jecper/v19y2005i4p25-42.html

Greene, J. D. (2008). The secret joke of kant's soul. In W. Sinnott-Armstrong, *Moral*

Table 1

Here is the table caption.

Numbers	Letters
1	A
2	B
3	C
4	D

Note. Here is the note below the table.

psychology volume 3: The neurosciences of morality: Emotion, brain disorders, and development (pp. 35–79). the MIT press.

Haidt, J., Björklund, F., & Murphy, S. (2000). Moral dumbfounding: When intuition finds no reason. *Unpublished Manuscript, University of Virginia.*

Kruglanski, A. W. (2013). *The psychology of closed mindedness.* Psychology Press.

Kruglanski, A. W., Atash, M. N., De Grada, E., Mannetti, L., & Pierro, A. (2013). *Need for closure scale (NFC). Measurement instrument database for the social science.*

Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and deliberate judgments are based on common principles. *Psychological Review, 118*(1), 97–109. https://doi.org/10.1037/a0020762

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: ”Seizing” and ”freezing.”. *Psychological Review, 103*(2), 263–283. https://doi.org/10.1037/0033-295X.103.2.263

Liberman, N., Sagristano, M. D., & Trope, Y. (2002). The effect of temporal distance on level of mental construal. *Journal of Experimental Social Psychology, 38*(6), 523–534. https://doi.org/10.1016/S0022-1031(02)00535-8

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

McHugh, C., McGann, M., Igou, E. R., & Kinsella, E. L. (2020). Reasons or rationalizations: The role of principles in the moral dumbfounding paradigm. *Journal of Behavioral Decision Making*, *33*(3), 376–392. https://doi.org/10.1002/bdm.2167

McHugh, C., McGann, M., Igou, E. R., & Kinsella, E. L. (2022). Moral judgment as categorization (MJAC). *Perspectives on Psychological Science*, *17*(1), 131–152. https://doi.org/10.1177/1745691621990636

McHugh, C., McGann, M., Igou, E. R., & Kinsella, E. L. (2023). Cognitive load can reduce reason-giving in a moral dumbfounding task. *Collabra: Psychology*, *9*(1), 73818. https://doi.org/10.1525/collabra.73818

Royzman, E. B., Kim, K., & Leeman, R. F. (2015). The curious tale of julie and mark: Unraveling the moral dumbfounding effect. *Judgment and Decision Making*, *10*(4), 296–313.

Strahan, R., & Gerbasi, K. C. (1972). Short, homogeneous versions of the marlow-crowne social desirability scale. *Journal of Clinical Psychology*, *28*(2), 191–193. https://doi.org/10.1002/1097-4679(197204)28:2%3C191::AID-JCLP2270280220%3E3.0.CO;2-G

Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision Making*, *11*(1), 99–113. https://ideas.repec.org/a/jdm/journl/v11y2016i1p99-113.html

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition*, *39*(7), 1275. https://doi.org/10.3758/s13421-011-0104-1

Unipark, Q. (2013). *QuestBack unipark.(2013)*.

Appendix

If there are multiple appendices, label them with level 1 headings as Appendix A, Appendix B, and so forth.