# Abstract

Occasionally, people judge an action as wrong, but when asked why, they may simply assert “*it’s just wrong!*” or even “*it’s wrong, but I can’t explain why*”. Maintaining a moral judgment without supporting reasons is known as moral dumbfounding. But why are people dumbfounded? One explanation proposes that it occurs as a result of a conflict between intuitive and deliberative processes. Previous research demonstrated that a manipulation designed to lead to more intuitive thinking rather than deliberative thinking (increased cognitive load), reduced the likelihood of providing reasons, and increased dumbfounded responding. We examine if dumbfounded responding can be reduced by facilitating 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 experiment (*N* = 2,400) to test this prediction.

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

# 

# Space to Think: Testing the Effect of Distancing on Moral Dumbfounding – A Registered Report

*“This line of thinking is morally wrong”* (US Representative discussing role of parents/schools in raising children in the House of Representatives in 2023)

*“The Bill is unnecessary and it is poorly drafted, but above all, it is deeply wrong…”* ((UK Politician discussing free speech in Universities in UK Parliament in 2021)

*“I appeal to the Minister to put this Bill on the shelves of his Department and to leave it to become surrounded by cobwebs … Artificial contraception is morally wrong.”* (Irish Politician discussing contraception in the Dáil in 1979)

People often appeal to morality to support their positions in arguments or discussions. Such moral appeals are typically presented as self-evident (Skitka, 2010) and justifications or reasons not provided. In politically charged, contentious debates proponents may argue for something *“because it is the right thing to do”* while opponents may reject something *“because it is just plain wrong”*. It is not always apparent if the speakers could provide a justification if pressed. In fact, there is good reason to expect that, at least in some cases, speakers would fail to provide reasons in support of their moral position (e.g., McHugh, Zhang, et al., 2023; see also Sim, 2016). Interestingly, it is not even clear if a position would be strengthened by the inclusion of reasons; research has shown that including consequentialist justifications can reduce the perceived moral relevance of an argument, potentially reducing its moral appeal (Kreps & Monin, 2014).

This failure to provide reasons for a moral judgment is known as moral dumbfounding (Haidt et al., 2000; McHugh et al., 2017). While perhaps not as prevalent as initially thought (e.g., Royzman et al., 2015; see also McHugh et al., 2020), there is a growing body of evidence demonstrating that it is a robust phenomenon (e.g., in a cross-cultural study McHugh, Zhang, et al., 2023 reported rates of dumbfounded responding ranging from 9% to 36% depending on the scenario and population). Despite a growing body of empirical work demonstrating the phenomenon, it remains poorly understood, with only limited work examining the specific mechanisms that underpin dumbfounded responding.

# The Importance of Studying Dumbfounding

Since its earliest empirical demonstrations (Haidt et al., 2000), moral dumbfounding has influenced theorizing about moral judgment. It is typically cited as evidence that moral judgments can be grounded in intuitive, automatic, or emotional processes (Cushman et al., 2010; Haidt, 2001; Prinz, 2005), informing the development of both intuitionist (Haidt, 2001) and dual-process theories of moral judgment (Byrd & Conway, 2019; Cushman, 2013).

While its influence on the development of theory is apparent, the cognitive processes that lead to dumbfounding are not well accounted for within current theorizing. There have been some attempts to explain the phenomenon (e.g., Green, 2025; Haidt, 2001; McHugh et al., 2022), however, there is only limited empirical work directly testing these explanations (McHugh, McGann, et al., 2023). Furthering our understanding of moral dumbfounding is therefore critical for the advancement of theory in moral psychology. That is, a successful theory of moral psychology should be able to provide a clear account of the phenomenon. This is the first aim of the current work.

In addition to the theoretical importance of understanding moral dumbfounding, There is also a practical benefit to studying it. Understanding moral dumbfounding may lead to the development of techniques by which it can be reduced. This would be of value both because the experience of it is often morally uncomfortable, but also because dumbfounding limits effective communication and discussion. Interventions that reduce its likelihood may improve debates addressing morally charged issues in settings from schools to parliaments. Testing one such strategy is the second aim of the current research. We propose a pre-registered experiment that will (a) examine one possible explanation for moral dumbfounding and (b) test the efficacy of one strategy that might help reduce dumbfounding: psychological distancing.

# Moral Dumbfounding as a Conflict in Dual-Processes

McHugh, McGann, et al. (2023) proposed a dual-process explanation of moral dumbfounding. Dual-process approaches contrast intuitive/habitual responding, e.g., *killing is wrong!*, with deliberative responding, e.g., *sacrificing one to save five may be justified* (e.g., Bago & De Neys, 2019; Cushman, 2013; Greene, 2008). McHugh, McGann et al. (2023) argue that moral dumbfounding results from a conflict in dual-processes, that is, when a habitual/intuitive response - making a moral judgment - differs from a response that results from deliberation - providing reasons for a moral judgment (e.g., Bonner & Newell, 2010; De Neys & Glumicic, 2008).

According to this view, rates of reason-giving should be reduced in situations where deliberation is inhibited. McHugh, McGann, et al. (2023) tested this prediction demonstrating that a cognitive load manipulation (requiring participants to attend to a secondary task while responding to the moral dumbfounding protocol) led to reduced rates of reason-giving and higher rates of dumbfounding (McHugh, McGann, et al., 2023).

 The reverse prediction can also be made: under conditions where deliberative responding is more likely, rates of reason-giving should increase, and rates of dumbfounding should decrease. We draw on research on construal level theory and psychological distance (e.g., Trope & Liberman, 2010), and predict that under conditions of increased psychological distance, rates of reason-giving should increase. The proposed studies not only extend and build on McHugh McGann, et al., (2023), but also address a limitation associated with this work. Specifically, it is unclear if manipulating cognitive load inhibited participants’ ability to provide reasons, or if participants under cognitive load were simply less motivated to provide reasons. The manipulations proposed here do not have the same limitation.

# Distancing and Dumbfounding

Psychological distance is the degree to which something is removed from direct experience, or the perception that something is close or distant from the self (this could be temporal, spatial, social, or hypothetical, see Trope & Liberman, 2010). Previous research has shown that varying psychological distance impacts how people think about things (e.g., Förster et al., 2004; though see Maier et al., 2022).

Most relevant for the current research is the suggestion that manipulations intended to increase psychological distance will lead to an increase in deliberative thinking (e.g., Bar-anan et al., 2006; Evans & Stanovich, 2013; Trope & Liberman, 2010). Specifically, greater psychological distance is said to be associated with higher-level construals, which are associated with more abstract (Bar-anan et al., 2006) and analytical or deliberative thinking (Evans & Stanovich, 2013), as well as less emotional thinking (Powers & LaBar, 2019). Conversely, decreased psychological distance is said to be associated with lower level construals, and associated with more emotional thinking (Powers & LaBar, 2019), less abstraction (more concrete thinking), and less deliberative (more intuitive) thinking (Evans & Stanovich, 2013). Together, this suggests that increasing psychological distance should facilitate deliberative thinking (and lead to higher rates of reason-giving in a dumbfounding task).

McHugh, McGann, et al. (2023), mapped the three responses in the dumbfounding paradigm according to their relative amount of deliberation. This approach assumes that making a moral judgment in the dumbfounding paradigm involves an intuitive/habitual response (in line with model-free responses described by Cushman, 2013; or habitualized categorizations described by McHugh et al., 2022). In support of this view, participants in studies of moral dumbfounding are typically asked to judge the (often shocking) behavior of others, rather than being asked to compare possible outcomes and make recommend a decision from a first person perspective (see Monin et al., 2007). In contrast, providing a reason for a moral judgment involves greater deliberation. Applied to the responses in the dumbfounding paradigm, providing a reason for a moral judgment (reason-giving) involves the most deliberation; providing a dumbfounded response (dumbfounding) involves the least amount of deliberation.

Characterizing the third response, nothing-wrong, is somewhat more complex. It is plausible that there is some level of deliberation involved in engaging with the counter-arguments and recognizing that an initial judgment may need either revision, or stronger justification. Having recognized that an initial position needs either revision or stronger justification, we propose (in line with McHugh, McGann, et al., 2023) that generating and articulating additional reasons to defend the initial position involves more deliberation than simply revising a judgment (based on deliberation already completed). In this way, while both nothing-wrong, and reason-giving involve more deliberation than providing a dumbfounded response, reason-giving involves more deliberation than nothing-wrong.

In line with this, and based on the prediction that increasing psychological distance should facilitate deliberative thinking, we propose our first hypothesis: that under conditions of increased psychological distance, rates of reason-giving will increase, leading to lower rates of dumbfounding, and lower rates of nothing-wrong. In Figure 1 we have mapped each response according to the relative amount of deliberation, and relative construal level and psychological distance.

Hypothesis 1 is based on a straightforward mapping of the responses in the dumbfounding paradigm both to relative amount of deliberation and to construal level. However, this does not account for the possible relationship between construal level and moral judgments more generally. This poses a particular challenge for attempting to incorporate nothing-wrong responses, that is, if construal level influences people’s judgments, our results in relation to nothing-wrong responses would be confounded by this. To account for this we will take the following steps. First we will explicitly examine participants’ judgments using a measure that is separate from the reason-giving/dumbfounding/nothing-wrong measure. Second, we will conduct all analyses twice, where our first analyses includes all participants, and the second analyses excludes participants who selected nothing-wrong to directly examine reason-giving against dumbfounding.

**Figure 1**

*Hypothesized level of deliberation for each response in the dumbfounding paradigm*

*A diagram of a level of construction

AI-generated content may be incorrect.*

*Note.* Adapted from McHugh, McGann, et al. (2023)

# Construal Level and Moral Judgments

Previous research has shown that construal level can influence moral judgments, with mixed findings (e.g., Alper, 2020; Žeželj & Jokić, 2014). On the one hand, some research has shown that under conditions of increased psychological distance/higher level construals people show greater concern for moral issues, as evidenced by more severe condemnation of targets who transgress or a reduced willingness to endorse committing a given act (Agerström et al., 2013; Agerström & Björklund, 2009b, 2009a; Eyal et al., 2008; Lammers, 2012). On the other hand, the opposite relationship has also been found, that is, people have been found to have greater concern for moral issues under conditions of decreased psychological distance/lower level construals (Eyal et al., 2014; Gong & Medin, 2012). Still others have found inconsistent effects of construal level on moral judgments, where the presence and direction of the varies depending on the sample being studied, and on the construal level manipulation employed (Gamliel et al., 2017; Žeželj & Jokić, 2014). Informed by this we propose our second hypothesis: participants’ moral judgments will be influenced by construal level.

Based on the previous research there are two competing predictions as follows: Hypothesis 2a: higher construal will lead to stronger condemnation of the behaviors, and lower construals will lead to weaker condemnations (e.g., Agerström & Björklund, 2009a; Eyal et al., 2008); Hypothesis 2b Higher construal will lead to weaker condemnation, and lower construals will lead to stronger condemnations (e.g., Gong & Medin, 2012; see also Ledgerwood et al., 2010 regarding potential susceptibility to counter arguments).

# The Role of Ambiguity and Construal Level

Above we have identified competing predictions regarding the relationship between construal level and moral judgment. One possible way to reconcile these competing predictions is to examine the role of ambiguity. Amaral and Jiao (2023) showed opposing effects of construal level on participants willingness to behave unethically, demonstrating that the direction of the effects depended on the ambiguity (or relative salience of trade-offs) of the scenarios. They showed that for unambiguous scenarios (without salient trade-offs, e.g., cleaning a toilet with a flag, cheating on a partner) participants reported being more likely to behave unethically in conditions of lower level construals (reduced psychological distance) compared to higher level construals (increased psychological distance), higher level construals predicted greater concern for moral values. In contrast for ambiguous scenarios (where there is a more salient trade-off between achieving a goal and committing an immoral act, e.g., keeping an item that was received by mistake), participants reported being more likely to behave unethically for higher level construals (increased psychological distance) compared to lower level construals (decreased psychological distance). In line with these findings we make a similar prediction here, for our third hypothesis: Ambiguity (trade-off salience) will interact with the effect of construal level on moral judgments. Specifically, for ambiguous (high trade-off salience) scenarios (*Trolley*, *Heinz*), lower construals will lead to stronger condemnation of the behavior, while in unambiguous (low trade-off salience) scenarios (*Jennifer*, *Julie and Mark*) higher construals will lead to stronger condemnation of the behavior.

It is possible that ambiguity also influences rates of reason-giving/dumbfounding. We propose two exploratory hypotheses regarding the possible effects of ambiguity on reason-giving. Our fourth hypothesis is that rates of reason-giving/dumbfounding will vary depending on ambiguity (trade-off salience/ambiguity). Specifically, we expect that reason-giving will be higher for ambiguous scenarios than for unambiguous scenarios. The ambiguous scenarios used in the studies presented here contain trade-offs framed in utilitarian terms, and we anticipate that this will prompt participants to think of utilitarian reasons to justify their judgments. Furthermore, these utilitarian reasons may be robust to the counter-arguments, leading to higher rates of reason-giving. In contrast, unambiguous scenarios will be more likely be considered in deontological terms, which may not be as robust to the counter-arguments leading to lower rates of reason-giving.

In line with the prediction that ambiguity will interact with construal level in predicting moral judgments, it is plausible to expect a similar interaction in relation to reason-giving. Our fifth (exploratory) hypothesis predicts that ambiguity (trade-off salience) will interact with the effect of construal level on reason-giving: We propose three competing predictions. Hypothesis 5a: the predicted effect of construal level on reason-giving/dumbfounding will only be observed for unambiguous (low trade-off salience) scenarios, with no effect of construal level on ambiguous scenarios. Specifically, for unambiguous scenarios, higher construals will lead to increased reason-giving. For ambiguous scenarios, reason-giving will be higher overall, and unaffected by construal level. Hypotheis 5b: The predicted effect of construal level on reason-giving/dumbfounding will only be observed for ambiguous (high trade-off salience) scenarios, with no effect of construal level on unambiguous scenarios. For ambiguous scenarios, higher construals will lead to increased reason-giving. For unambiguous scenarios, reason-giving will be lower overall and will be unaffected by construal level. Hypothesis 5c: The predicted effect of construal level on reason-giving/dumbfounding will be observed for unambiguous scenarios, but will be reversed for ambiguous scenarios. Specifically, for unambiguous scenarios higher construals will lead to increased reason-giving. For ambiguous scenarios the lower construals will lead to greater attention to contextual factors (Alper, 2020), which may facilitate reason-giving compared to higher level construals.

Above we have made predictions about the expected patterns of responses we expect to see depending on construal level and ambiguity. We make one final prediction regarding the content of the reasons participants provide in support of their judgments. Our sixth (exploratory) hypothesis is that the content of the reasons provided by participants will vary depending on construal level. Specifically, we predict that higher construals will lead to reasons that are based on abstract (deontological) principles, whereas lower construals will lead to more concrete/contextual (utilitarian) factors being cited in reasons given (e.g., Körner & Volk, 2014; Xiao et al., 2015).

# The Current Research

The current research tests the prediction that manipulations designed to increase psychological distance will facilitate reason-giving in a moral dumbfounding task. We will test two different manipulations across two domains of psychological distance, temporal distance and abstract/concrete thinking manipulated using the “how-and-why” task. Our core prediction is that that under higher construals/increased psychological (temporal) distance, rates of reason-giving will be higher than in a control group or under reduced psychological distance. These findings will allow us to peer into the moral judgment process to better understand the cognitive processes that underpin moral dumbfounding, including those which may mitigate it.

We propose two studies to test the relationship between construal level and rates of reason-giving in a moral dumbfounding task. In Study 1 we will manipulate construal level using a temporal distance manipulation, participants will be asked to imagine the moral scenarios as occurring in the distant future (one year from now; increased temporal distance condition), the near future (tomorrow; reduced temporal distance condition), or without any specific instructions regarding a time frame (control). Study 2 will prime high versus low construal levels using the “how-and-why” task (taken from Freitas et al., 2004; Fujita et al., 2006). Participants in the abstract/high construal level group will be asked a target “why?” question with follow-up “why?” questions probing their responses. Participants in the concrete/low construal group will be presented with a “how?” question with a follow up “how?” questions probing their responses. Previous research has shown opposing effects of construal level on moral judgments. One possible explanation for these opposing effects is ambiguity (or trade-off salience). To account for we will present participants in both studies with four moral scenarios that can be characterized as either ambiguous (high trade-off salience) or unambiguous (low trade-off salience). The ambiguous scenarios are *Trolley* and *Heinz*, while the unambiguous scenarios are *Julie and Mark*, and *Jennifer*. The hypotheses for the proposed studies are summarized in Table 1.

**Table 1**

Overview of the hypotheses

|  |  |  |  |
| --- | --- | --- | --- |
| Hypothesis | Independent Variable(s) | Dependent Variable | Prediction(s) |
| H1 | Construal level | Reason-giving / dumbfounding / nothing-wrong | Higher construal increases reason-giving (H1a)  Higher construal decreases dumbfounding (H1b)  Higher construal decreases nothing-wrong (H1c) |
| H2 (competing) | Construal level | Moral judgment | Higher construal increases condemnation (H2a)  Higher construal decreases condemnation (H2b) |
| H3  (reconciling competing H2a & H2b) | Construal level ×  Ambiguity | Moral judgment | For unambiguous scenarios, higher construal increases condemnation.  For ambiguous scenarios, lower construal increases condemnation |
| H4  (exploratory) | Ambiguity | Reason-giving / dumbfounding | Increased reason-giving for ambiguous scenarios |
| H5  (exploratory) | Construal level × Ambiguity | Reason-giving / dumbfounding | Higher construal increases reason-giving for unambiguous scenarios only (H5a)  Higher construal increases reason-giving for ambiguous scenarios only (H5b)  Higher construal increases reason-giving for unambiguous scenarios, but *lower* construal increases reason-giving for ambiguous scenarios (H5c) |
| H6  (exploratory) | Construal level | Types of reasons provided | Higher construal leads to more abstract principles and deontological justifications  Lower construal leads to more concrete, contextual and utilitarian justifications |

We note two potential limitations associated with the hypotheses as presented above. First, it assumes that our manipulations will be successful in influencing construal level, and that these changes in construal level will have downstream effects on rates of reason-giving, and on moral judgments. However, the effectiveness and reliability of construal level manipulations has come under question in recent years, with multiple failed replications (for discussion, see Maier et al., 2022). A recent meta-analysis failed to find strong evidence for construal level theory effects when controlling for publication bias (Maier et al., 2022). Second, our hypotheses assume that the effects of two different manipulations will be comparable. While there is some evidence to support this assumption (e.g., Amaral & Jiao, 2023; Eyal et al., 2009; Lammers, 2012; Ledgerwood et al., 2010), it is also the case that different manipulations can lead to different effects (Alper, 2020; Žeželj & Jokić, 2014). The studies presented by Žeželj and Jokić particularly notable because they show inconsistencies in the direction of effects across manipulations, as well as showing null effects in come cases.

To mitigate these limitations, and ensure that meaningful conclusions can be taken from our findings (even in the case of null results), we have included manipulation checks in both studies. These means we can determine if any null findings are a consequence of the manipulation failing or if they provide evidence that the manipulation did not affect the measure(s) of interest. We will also investigate the effect of construal level on judgments made, in addition to our primary measure of interest (reason-giving), further helping us to make meaningful conclusions from any null results that we may observe.

Both Studies will be conducted in multiple sites. The methods across sites will be the same with the only difference being the sample (and the language of the materials). As such below we describe the overall methods for Study 1 and Study 2 below and only differentiate between locations when describing the participants.

**Table 2**

Location and label for each study in each location

|  |  |  |
| --- | --- | --- |
| Location | Study 1 (*N* = 685) | Study 2 (*N* = 281) |
| Ireland | Study 1a | Study 2a |
| Poland | Study 1b | Study 2b |
| UK | Study 1c | Study 2c |
| Chile | Study 1d | Study 2d |
| USA | Study 1e | Study 2e |
| Belgium | Study 1f | Study 2f |

# Study 1a-e

## Method

### Participants and Design

Study 1 will be a mixed between-within-subjects design. The first dependent variable is rates of reason-giving/dumbfounding, assessed using two measures: (i) the primary measure will be responses to the critical slide which has 3 response options: 1: reason-giving; 2: nothing-wrong; 3: dumbfounded response - admission; (ii) participants who declare they can provide a reason will be asked to provide a reasons with an open-ended response question, we will code these open-ended responses for reasons/dumbfounded responses (defined as unsupported declarations or tautological responses). The second dependent variable is moral judgment, measured on a 7-point Likert scale (where 1 = *morally wrong*; 4 = *neutral*; 7 = *morally right*). The between-subjects independent variable is temporal distance, with three levels: control, reduced temporal distance, and increased temporal distance. Temporal distance will be manipulated by varying the instructions provided to participants. In the control condition, participants will not be provided with any instructions about a time frame; in the increased temporal distance condition, participants will be asked to imagine the target scenario occurring in the distant future (a year from now); in the reduced distance condition, participants will be asked to imagine the scenario occurring in the near future (tomorrow). The within-subjects independent variable is ambiguity with two levels: ambiguous and unambiguous. Ambiguity will be manipulated through different scenarios. We will use the same four scenarios as McHugh, McGann, et al. (2023). The two unambiguous scenarios do not include any salient pay-offs (*Julie and Mark* depicts an act of consensual incest while *Jennifer* engages in an act of cannibalism). In contrast, the ambiguous scenarios involve salient trade-offs (in *Trolley* the worth of one life is compared against the worth of five lives; in *Heinz* there is a trade-off between the rights of the Druggist and the rights of Heinz and his wife).

The most basic test of our hypothesis will involve a series of four 3 × 3 chi-squared tests. Our dependent variable is responses in a dumbfounding task (with three levels: 1: reason-giving; 2: nothing-wrong; 3: dumbfounded response) and our independent variable is temporal distance (with three levels: 1: control; 2: increased; 3 decreased). For each scenario we will conduct a 3 × 3 chi-squared test to examine if there is a relationship between temporal distance and rates of reason-giving. We will conduct follow-up mixed-effects logistic-regressions to test for an overall effect across scenarios while controlling for within-participant variability and differences between scenarios. This follow-up analysis will be based on a larger number of observations than the initial chi-squared tests and as a result will be of higher statistical power (Kumle et al., 2021). We therefore base our sample size calculations on the analysis of lower statistical power (the chi-squared tests). In Table 3 we report the required sample size to detect small (*V* = .05), medium (*V* = .15) , and large (*V* = .25) effects with 90% power. We set out target sample for Study 1 at *N* = 685, sufficient to detect a medium effect with 90% power. This target sample of *N* = 685 will be collected at each site across Studies 1a-1f (i.e., Study 1a target *n* = 685 in Ireland, Study 1b target *n* = 685 in Poland etc.)

**Table 3**

*Power analyses and sample size calculations*



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  | Study 1: Temporal 3 groups (control, distant, near) | Study 2: how-and-why task 2 groups (abstract concrete) |
| Large | *V* |  | 0.25 | 0.35 |
|  | *N* |  | 246 | 101 |
| Medium | *V* |  | 0.15 | 0.21 |
|  | *N* |  | 685 | 281 |
| Small | *V* |  | 0.05 | 0.07 |
|  | *N* |  | 6162 | 2531 |

*Note.* *V* = Cramer’s V, effect size for proposed chi-squared test



Participants in Study 1a will be recruited through a combination of convenience and snowball sampling in Ireland. Some participants may be recruited using the SONA research participation system where college students will receive course credit in return for their participation. Where resources allow, additional participants may be recruited through online research participation systems (e.g., Prolific / Lucid / CloudResearch / MTurk) if necessary.

Participants in Study 1b…

Participants in Study 1c…

Participants in Study 1d…

Participants in Study 1e…

Participants in Study 1f…

### Procedure and Materials

Data will be collected using an online survey programmed in Qualtrics. On agreeing to take part in the survey, participants will be randomly assigned to one of three experimental groups. At this point participants will be presented with an instruction page that will contain the experimental manipulation (taken from Žeželj & Jokić, 2014)

In the control condition, participants will be presented with instructions that read: “*In the following you will read several situations of people making decisions. After you read each scenario you will be asked a series of questions, including making a judgment. Read carefully and answer the questions below each description.*” The control condition will not contain any additional instructions appended to the scenarios.

In the increased temporal distance condition, participants will be presented with the following instructions:

**In the next couple of minutes try to think about a day a year from now: how it will look like, what will happen, what you will do/feel...**

In the following you will read several situations of people making decisions. After you read each scenario you will be asked a series of questions, including making a judgment.

**Try to imagine each scenario is going to happen a year from now.** Read carefully and answer the questions below each description.

At the end of each scenario, participants will be reminded to “*Imagine this event is going to happen in a year from now.*”

The decreased temporal distance conditions will include the following instructions:

**In the next couple of minutes try to think about tomorrow: how it will look like, what will happen, what you will do/feel...**

In the following you will read several situations of people making decisions. After you read each scenario you will be asked a series of questions, including making a judgment.

**Try to imagine each scenario is going to tomorrow.** Read carefully and answer the questions below each description.

At the end of each scenario participants will be reminded to “*Imagine this event is going to happen tomorrow.*”

As a manipulation check, after each scenario, we will ask participants to rate on a seven-point Likert scale “how far into the future did you imagine this event?”1 = *very close/very soon*, 7 = *very far/very distant* (taken from Gamliel et al., 2017)

Following these initial instructions, participants will be presented with a series of four moral scenarios (*Heinz*, *Trolley*, *Jennifer*, and *Julie and Mark*) previously used in the study of moral dumbfounding (McHugh et al., 2017), and the influences on moral dumbfounding (McHugh, McGann, et al., 2023). These scenarios will be used to manipulate ambiguity, with *Heinz* and *Trolley* as ambiguous scenarios, and *Julie and Mark* and *Jennifer* as unambiguous scenarios. Participants will be asked to rate on a 7-point Likert scale how right or wrong the behavior described is (where 1 = *morally wrong*; 4 = *neutral*; 7 = *morally right*), as well as their confidence in this judgment (where 1 = *not at all confident*; 7 = *extremely confident*). Participants will then be presented with a series of counter-arguments, which refute commonly used justifications for rating the behavior as “wrong” (see supplementary materials for full text of scenarios and all counter-arguments). After each counter-argument, participants will be asked if they view the behavior as wrong (yes/no), and if they have a reason for this judgment (yes/no/unsure).

Following the counter-arguments and related questions, dumbfounding will be measured using the critical slide (McHugh et al., 2017). This contains a statement defending the behavior and a question as to how the behavior could be wrong (e.g., “Jennifer’s actions did not harm anyone, or negatively affect anyone. How can there be anything wrong with what she did?”). There are three possible answer options (presented in randomized order): (a) “It’s wrong, and I can provide a valid reason” (reasons-giving); (b) “It’s wrong, but I can’t think of a reason” (dumbfounding: an admission of not having reasons); (c) “There is nothing wrong” (nothing-wrong). Participants who select (a) will be prompted to type a reason on the next page. Finally, participants will rate the behavior (again where 1 = *morally wrong*; 4 = *neutral*; 7 = *morally right*) and their confidence (where 1 = *not at all confident*; 7 = *extremely confident*) again. We will also record participants’ response times, which may be used in future secondary and exploratory analyses.

Our primary measure of reason-giving/dumbfounding will be based on the responses to this critical slide – participants who claim they have reasons will be treated as if they can provide reasons, while only participants who select the admission of not having reasons will be classified as having provided a dumbfounded response. We note that this is a conservative measure of dumbfounded responding (for discussion see McHugh et al., 2017; McHugh, McGann, et al., 2023), that likely fails to account for people who claim to be able to provide reasons but cannot provide reasons when asked to do so.

In the original dumbfounding studies (Haidt et al., 2000) and in the McHugh et al. (2017) replication studies participants responses were coded for responses that may be evidence of dumbfounding. McHugh et al., (2017) identified two classes of such responses: (i) unsupported declarations, e.g., “*It’s just wrong*” (also reported by Haidt et al., 2000), and (ii) tautological responses, where participants just restated the behavior as justification for their judgment, e.g., “*because it’s incest*”. There remains some debate whether these responses are truly evidence of dumbfounding (e.g., Royzman et al., 2015) and this has resulted in the conservative measure (selecting the admission of not having reasons on the critical slide) being more widely adopted.

It is possible that unsupported declarations/tautological responses do present a failure to provide reasons, and more critically for the current research, it is possible that the rates of these kinds of responses could be influenced by our experimental manipulation, e.g., in line with our hypothesis that increased distance will increase rates of reason-giving, it follows that participants in the increased distance condition may be more likely to provide more “substantive” reasons and fewer unsupported declarations/tautological responses. For this reason, we will employ an additional secondary measure of reason-giving/dumbfounding. Participants who indicate that they can provide a reason will be invited to provide their reason in an open-ended text box. We will code these open-ended responses for unsupported declarations and tautological responses, and such responses will be taken as dumbfounded responding (rather than reason-giving). These responses (unsupported declarations/tautological responses) are the only responses that we will classify as dumbfounding, and as such, we anticipate that the majority of responses will be acceptable as reasons, e.g., participants may invoke religion (“*because God doesn’t permit it*”), the law (“*because it’s illegal*”), or social norms (“*because it is unacceptable in our society*”) This secondary measure will incorporate participants’ responses to the critical slide with coded open-ended, such that some participants’ reason-giving responses will be re-classified as dumbfounded (or even nothing-wrong) following coding. As with the primary measure, this second measure will have three levels, reason-giving, dumbfounding, nothing-wrong.

A third measure will differentiate between different kinds of dumbfounded responding, separating participants who admitted to not having reasons (on the critical slide) from participants who failed to provide reasons (dumbfounded after coding). This measure will have four levels: reason-giving, dumbfounding-before-coding, dumbfounding-after-coding, nothing-wrong. Based on previous research (McHugh et al., 2017) we anticipate these proportions will be relatively low, potentially too low to draw meaningful conclusions, and as such any analyses using this third measure will be exploratory.

Two attention check tasks will be 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 is misleading, such that if participants skim or only read some of the text, they will likely provide an incorrect response (see supplementary materials). Participants who fail both attention checks will be excluded from the analysis. Participants will also be asked to report basic demographic information including age, gender, nationality, religion, political ideology, and place of residence.

### Analysis Plan

The main analyses for each Study 1a-1f will be conducted separately. Findings will be aggregated using an internal meta-analysis. We will additionally conduct exploratory pooled analyses similar to that described below (where we also control for study site). This pooled analysis will have a larger sample size and thus may be sufficient powered to detect smaller effects, however our conclusions will primarily be based on the results of the separate analyses and associated meta-analysis.

To test the effectiveness of our manipulation in influencing participants’ perceptions of how far into the future the events took place we will conduct a linear-mixed-effects model. This will include responses to the manipulation check question as the outcome measure. Temporal distance and scenario will be included as included as fixed factors, and participant ID will be included as a random factor. To examine potential differences between specific conditions we will conduct Tukey’s post-hoc pairwise comparisons.

We will conduct a series of chi-squared tests to test for an association between temporal distancing (near vs far vs control) and dumbfounded responding for each scenario separately (*Julie and Mark*, *Jennifer*, *Trolley*, *Heinz*), and for each measure of reason-giving/dumbfounding. The first set of chi-squared tests will take raw responses to the critical slide (reason-giving, dumbfounded, nothing-wrong) as the outcome measure resulting in 3 × 3 chi-squared tests. A second set of chi-squared tests will be based on our secondary measure of reason-giving/dumbfounding that additionally includes the coded open-ended responses from participants who provided reasons. Again, we will conduct a series of 3 × 3 chi-squared tests to test the association between temporal distance and rates of reason-giving/dumbfounding. Finally a series of exploratory 3 × 4 chi-squared tests will examine our third measure that differentiates between dumbfounding before and after coding.

Following-our scenario-by-scenario tests, we will test for an overall effect of temporal distance on reason-giving/dumbfounding using a series of mixed-effects multinomial logistic regressions. As above we will run separate analyses for each measure of reason-giving/dumbfounding: (i) raw responses to the critical slide (3 levels) (ii) incorporating the coded response from participants who provided reasons (3 levels) (iii) differentiating between dumbfounding before coding and dumbfounding after coding (4 levels). For the mixed-effects multinomial logistic regressions we will construct four different models, Models 1 and 2 will directly test our hypothesis that construal level will influence reason-giving, Model 3 will be exploratory and will include additional measures taken as covariates, and Model 4 will test the prediction that construal level might interact with ambiguity (trade-off salience). Model fit comparisons will be based on AIC and BIC statistics and all three models will be reported. Model 1 will include reason-giving/dumbfounding as the outcome measure and temporal distancing as a fixed effect, with random effects for participant ID. Model 2 will include reason-giving/dumbfounding as the outcome measure, temporal distancing, scenario, and a temporal distance × scenario interaction will be included as fixed effects, and random effects will be included for participant ID. Model 3 will be similar to Model 2 but will additionally include participants’ the other questions (initial/revised judgment, initial/revised confidence, confusion, irritation, based on reason, based on gut feeling) as fixed effects. Model 4 will include reason-giving/dumbfounding as the outcome measure, temporal distance, ambiguity/scenario type (ambiguous vs unambiguous), and a temporal distance × ambiguity interaction will be included as fixed effects and random effects will be included for participant ID.

One potential challenge to interpreting our results comes the inclusion of the “nothing-wrong” response option on our measure of reason-giving/dumbfounding. It is possible that temporal distancing may influence valence of judgment, without influencing participants’ ability to provide reasons. In this case we would see higher rates of selecting nothing-wrong, and lower rates of selecting the reason-giving response. However, attempting to interpret these lower rates of reason-giving would be confounded by valence of judgment. In order to account for this, we will run the above analyses twice: first, we will run the analyses on the full sample, second, we will remove participants wo selected nothing-wrong and re-run the analyses.

Following our main analysis to test our primary hypothesis, we will conduct follow-up analysis investigating the effect of temporal distance on other measured variables of interest. Specifically we will test for differences in the following measures: initial judgment, revised judgment, initial confidence, revised confidence, confusion, irritation, how much judgment was based on reason, and how much judgment was based on gut feeling. Each of these measures is recorded on a 7-point Likert scale. For each measure we will conduct a linear-mixed-effects model, with the measure of interest as the outcome variable; temporal distance, scenario, and a temporal distance × scenario interaction will be included as fixed effects, and participant ID will be included as random effects. We will conduct post-hoc Tukey’s pairwise comparison to examine any differences between specific conditions. Sample code for the analyses described above (using a simulated dataset) is available in the supplementary materials.

## Results and Discussion

# Study 2

## Method

### Participants and Design

Study 2 will be a mixed-between-within-subjects design. As in Study 1, the first dependent variable is rates of reason-giving/dumbfounding, again assessed using two measures: (i) responses to the critical slide and (ii) the coded open-ended responses incorporated with responses to the critical slide. Moral judgment is the second dependent variable, again, measured on a 7-point Likert scale. The between-subjects independent variable is construal level manipulated using a mindset manipulation, with two levels, higher level construal (abstract mindset) and low level construal (concrete mindset), this will be manipulated using the “how-and-why” task (Freitas et al., 2004). The within-subjects variable is ambiguity, again manipulated by presenting participants with four scenarios, where two are ambiguous (high trade-off salience) and two are unambiguous (low trade-off salience).

Similar to Study 1, the most basic test of our hypothesis will involve a series of 2 × 3 chi-squared tests. Our dependent variable is responses to the dumbfounding task (three levels: 1: reason-giving; 2: nothing-wrong; 3: dumbfounded response), and construal level/mindset is our independent variable with two levels (abstract vs concrete). For each scenario we will conduct a 2 × 3 chi-squared test for each scenario to test for a relationship between construal level/mindset and rates of reason-giving. Follow-up mixed-effects logistic-regressions will test for an overall effect across scenarios while controlling for within-participant variability and differences between scenarios. As with Study 1 this follow-up analysis will be of higher statistical power (due to larger number of observations, see Kumle et al., 2021), so we base our sample size calculations on our lower powered chi-squared tests. In Table 3 we report the required sample size to detect small (*V* = .07), medium (*V* = .21) , and large (*V* = .35) effects with 90% power. We set out target sample for Study 2 at *N* = 281, sufficient to detect a medium effect with 90% power. This target sample of *N* = 281 will be collected at each site across Studies 2a-2f (i.e., Study 2a target *n* = 281 in Ireland, Study 2b target *n* = 281 in Poland etc.).

Participants in Study 2a will be recruited through the same channels as in Study 1a, combining convenience and snowball sampling in Ireland, supplemented by the SONA research participation system and online research participation systems (e.g., Prolific / Lucid / CloudResearch / MTurk) where necessary.

Participants in Study 2b…

Participants in Study 2c…

Participants in Study 2d…

Participants in Study 2e…

Participants in Study 2f…

### Procedure and Materials

As with Study 1, data will be collected using an online survey programmed in Qualtrics. Participants who agree to take part will be randomly assigned to one of two experimental groups. Construal level/mindset will be manipulated using the “how-and-why” task (Freitas et al., 2004). Participants in the concrete (lower level construal) condition will complete the *how* task, and participants in the abstract (higher level construal) condition will complete the *why* task.

In the concrete (lower level construal) condition participants will read a brief preamble describing how a person could think about the process of *how* we can achieve our goals (with some illustrative examples see Supplementary materials for full text). The final sentence will read “For this thought exercise, please consider the following activity: ‘improving and maintaining one’s physical health’”. Participants will then be asked to “Please list three ways in which you can improve your health:”. There will be three empty text boxes which participants can enter their responses into. Under each text box participants will be asked to rate on a 5-point Likert scale (1 = *a little*, 5 = *very, very much*) in response to the question “how much will engaging in this activity improve and maintain your health?”. On the next page participants will be presented with a series of five vertically aligned boxes connected with downward arrows labelled “How?”, accompanied by the following instructions:

“Starting with the goal outlined in the top box, please identify how you might achieve this and enter your answer into the second box. Next, think about the response you just provided and identify how you might achieve this; enter your answer in the box. Work your way through the boxes until you have entered a response into all boxes. Your first response should identify how you might improve and maintain your health.”

As participants progress, they will be prompted with their previous response and asked how they would achieve this; e.g., if their response in the first box was “Go exercise”, the prompt would read “You entered ‘Go exercise’. How would you ‘Go exercise’?”. This process will repeat until all boxes are completed.

In the abstract (higher level construal) condition participants will read a similar preamble describing how a person could consider the reasons why they do things (again with some examples, see Supplementary materials for full text). The final sentence will read “For this thought exercise, please consider the following activity: ‘improving and maintaining one’s physical health’”. Participants will then be asked to “Please list three ways in which improving and maintaining your physical health could help you meet important life goals”, with three empty text poxes for participant responses. Under each box participants will be asked to rate “How much will improving and maintaining your health help you meet this important goal?” on a 5-point Likert scale (1 = *a little,* 5 = *very, very much*). On the next page, participants will be presented with a series of five vertically aligned boxes connected with upward arrows labelled “Why?”, and the following instructions:

“Starting with the goal outlined in the bottom box, please identify why you might want to achieve this and enter your answer into the second box. Next, think about the response you just provided and identify why you might want achieve this; enter your answer in the next box. Work your way through the boxes until you have entered a response into all boxes. Your first response should identify why you might want to improve and maintain your health”

As participants progress, they will be prompted with their previous response and asked why they want to achieve this, e.g., if their response was “To do well in school” the prompt would read “You entered ‘to do well in school’. Why do you want to ‘do well in school’?”. This process will repeat until all boxes are completed.

As a manipulation check, two judges who are unaware of the condition will rate the abstractness of responses to the second part of the how-and-why task. Any response that is identified as a subordinate means to the original statement will be assigned a score of -1, while any response that is a superordinate end of the original statement will be assigned a score of +1 (responses not meeting either criterion will be assigned a score of 0). Ratings for the four responses will be summed to create construal level index with a potential range of -4 to +4 where higher scores represent higher levels of construal (see Fujita et al., 2006).

The procedure for the dumbfounding task will be the same as in Study 1. We will present participants with four scenarios (*Heinz*, *Trolley*, *Jennifer*, and *Julie and Mark*). After each scenario, participants will rate the behavior (1 = *morally wrong*; 4 = *neutral*; 7 = *morally right*), and their confidence in this judgment (1 = *not at all confident*; 7 = *extremely confident*). Participants will be presented with the same series of counter-arguments and follow-up questions asking if they view the behavior as wrong (yes/no), and if they have a reason for their judgment (yes/no/unsure). Next participants will respond to the critical slide, and any participants who select reason-giving will be prompted to type a reason on the next page. Participants will then rate the behavior and their confidence again. Responses times will be recorded for potential future secondary and exploratory analyses.

As with Study 1, we will take two measures of reason-giving/dumbfounding. Our first measure will be based on responses to the critical slide only. Our second measure will additionally include the coded responses provided by participants who indicated that they can provide a reason; we will code these responses for unsupported declarations (*it’s just wrong*) and tautological responses (*because it’s incest*). Our second measure will reclassify participants providing such responses will be reclassified as dumbfounded (rather than reason-giving).

The same two attention checks as in Study 1 will be used in Study 2, and participants who fail both will be excluded from analysis. We also ask participants to provide basic demographic information including age, gender, nationality, religion, political ideology, and place of residence.

### Analysis Plan

The analysis plan for Study 2 will closely resemble that of Study 1. The design and hypotheses are similar. The differences between the studies is the manipulation of construal level that was used. In Study 1 we manipulated temporal distance and this independent variable had three levels, control, increased, and decreased. In Study 2 we manipulated mindset/construal level directly leading to an independent variable with two levels, abstract vs concrete.

As with Study 1, the main analyses for each Study 2a-2f will be conducted separately, with aggregated results collated using internal meta-analysis. Again, we will conduct some exploratory pooled analyses (controlling for study site), but our conclusions will be based on the separate analyses and meta-analyses.

As noted above, our manipulation check will involve independent coders coding the responses to the “how-and-why” task to generate an index with a potential range of -4 to +4 where higher scores represent higher levels of construal (described above). We will conduct an independent samples t-test to test for differences in this construal level index depending on experimental condition.

As in Study 1 our first test of our hypothesis will involve a series of chi-squared tests for each scenario separately (*Julie and Mark*, *Jennifer*, *Trolley*, *Heinz*). The first set of 2 × 3 chi-squared tests will examine raw responses to the critical slide, the second set of 2 × 3 chi-squared tests will examine the secondary measure of reason-giving/dumbfounding (when coded open-ended responses are included). A final set of exploratory 2 × 4 chi-squared tests will examine the third measure of reason-giving/dumbfounding that differentiates between dumbfounding before and after coding.

Next we will test for an overall effect of construal level on rates of reason-giving/dumbfounding using a series of mixed-effects multinomial logistic regression (with separate analyses for each of the three measures of reason-giving/dumbfounding: raw critical slide responses, including coded responses, differentiating between dumbfounding before/after coding). As in Study 1 we will construct four models for each outcome measure, with Models 1 and 2 directly testing the hypothesis, an exploratory Model 3 including additional measures, and Model 4 will examine the role of ambiguity (trade-off salience). Model 1 will include construal level as a fixed effect and random effects for participant ID, Model 2 will additionally include scenario and a construal level × scenario interaction as fixed effects, while Model 3 will additionally include initial/revised judgment, initial/revised confidence, confusion, irritation, based on reason, based on gut feeling, as fixed effects. In line with Study 1, we will conduct these analyses twice, first, on the full sample, and second, on the sample with participants who selected nothing-wrong excluded (to account for the potentially confounding influence of valence on interpreting our results).

As in Study 1 we will conduct follow-up analyses investigating the effect of construal level on the other measures taken: initial judgment, revised judgment, initial confidence, revised confidence, confusion, irritation, how much judgment was based on reason, and how much judgment was based on gut feeling. This will involve a series of linear-mixed-effects models with the measure of interest as the outcome variable; construal level, scenario, and a construal level × scenario interaction will be included as fixed effects, and participant ID will be included as random effects. We will conduct post-hoc Tukey’s pairwise comparison to examine any differences between specific conditions. Sample code for the analyses described above (using a simulated dataset) is available in the supplementary materials.

## Results and Discussion

# General Discussion

# Disclosures

All measures, manipulations, and data/participant exclusions are reported in the manuscript or its Supplementary Material.

# Open Practices

All (pilot) data, analysis scripts, and materials are available from the OSF at <https://osf.io/3fuer/?view_only=c2b02ef663fc4a12a2c3a143c21d9776>. Full data from the main study will be made available on completion of the study.

# References

Agerström, J., & Björklund, F. (2009a). Moral Concerns Are Greater for Temporally Distant Events and Are Moderated by Value Strength. *Social Cognition*, *27*(2), 261–282. https://doi.org/10.1521/soco.2009.27.2.261

Agerström, J., & Björklund, F. (2009b). Temporal Distance and Moral Concerns: Future Morally Questionable Behavior is Perceived as More Wrong and Evokes Stronger Prosocial Intentions. *Basic and Applied Social Psychology*, *31*(1), 49–59. https://doi.org/10.1080/01973530802659885

Agerström, J., Björklund, F., & Carlsson, R. (2013). Look at Yourself!: Visual Perspective Influences Moral Judgment by Level of Mental Construal. *Social Psychology*, *44*(1), 42–46. https://doi.org/10.1027/1864-9335/a000100

Alper, S. (2020). Explaining the Complex Effect of Construal Level on Moral and Political Attitudes. *Current Directions in Psychological Science*, *29*(2), 115–120. https://doi.org/10.1177/0963721419896362

Amaral, N. B., & Jiao, J. (2023). Responses to Ethical Scenarios: The Impact of Trade-Off Salience on Competing Construal Level Effects. *Journal of Business Ethics*, *183*(3), 745–762. https://doi.org/10.1007/s10551-021-04995-x

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

Bar-anan, Y., Liberman, N., & Trope, Y. (2006). The Association Between Psychological Distance and Construal Level: Evidence From an Implicit Association Test. *Journal of Experimental Psychology: General*, *135*(4), 609–622.

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

Byrd, N., & Conway, P. (2019). Not all who ponder count costs: Arithmetic reflection predicts utilitarian tendencies, but logical reflection predicts both deontological and utilitarian tendencies. *Cognition*, *192*, 103995. https://doi.org/10.1016/j.cognition.2019.06.007

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

Cushman, F. A., Young, L., & Greene, J. D. (2010). Multi-system Moral Psychology. In J. M. Doris (Ed.), *The Moral Psychology Handbook* (pp. 47–71). Oxford University Press.

De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition*, *106*(3), 1248–1299. https://doi.org/10.1016/j.cognition.2007.06.002

Evans, J. St. B. T., & Stanovich, K. E. (2013). Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspectives on Psychological Science*, *8*(3), 223–241. https://doi.org/10.1177/1745691612460685

Eyal, T., Liberman, N., & Trope, Y. (2008). Judging near and distant virtue and vice. *Journal of Experimental Social Psychology*, *44*(4), 1204–1209. https://doi.org/10.1016/j.jesp.2008.03.012

Eyal, T., Liberman, N., & Trope, Y. (2014). Thinking of why a transgression occurred may drawn attention to extenuating circumstances: A comment on Žeželj & Jokić replication. *Social Psychology*, *45*(4), 329–331.

Eyal, T., Sagristano, M. D., Trope, Y., Liberman, N., & Chaiken, S. (2009). When values matter: Expressing values in behavioral intentions for the near vs. distant future. *Journal of Experimental Social Psychology*, *45*(1), 35–43. https://doi.org/10.1016/j.jesp.2008.07.023

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*, *87*(2), 177–189. https://doi.org/10.1037/0022-3514.87.2.177

Freitas, A. L., Gollwitzer, P., & Trope, Y. (2004). The influence of abstract and concrete mindsets on anticipating and guiding others’ self-regulatory efforts. *Journal of Experimental Social Psychology*, *40*(6), 739–752. https://doi.org/10.1016/j.jesp.2004.04.003

Fujita, K., Trope, Y., Liberman, N., & Levin-Sagi, M. (2006). Construal Levels and Self-Control. *Journal of Personality and Social Psychology*, *90*(3), 351–367. https://doi.org/10.1037/0022-3514.90.3.351

Gamliel, E., Kreiner, H., & McElroy, T. (2017). The effect of construal level on unethical behavior. *The Journal of Social Psychology*, *157*(2), 211–222. https://doi.org/10.1080/00224545.2016.1208139

Gong, H., & Medin, D. L. (2012). Construal levels and moral judgment: Some complications. *Judgment and Decision Making*, *7*(5), 628–638. https://doi.org/10.1017/S1930297500006343

Green, A. (2025). Moral dumbfounding and imaginative resistance. *Philosophical Psychology*, *38*(5), 2146–2164. https://doi.org/10.1080/09515089.2023.2299374

Greene, J. D. (2008). The Secret Joke of Kant’s Soul. In W. Sinnott-Armstrong (Ed.), *Moral Psychology Volume 3: The neurosciences of morality: Emotion, brain disorders, and development* (pp. 35–79). The MIT Press.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, *108*(4), 814–834. https://doi.org/10.1037/0033-295X.108.4.814

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

*Health (Family Planning) Bill, 1978: Committee Stage (Resumed).: Hearing on Vol. 315 No. 1 before the Dáil Éireann*, DáIl Éireann (1979).

*Higher Education (Freedom of Speech) Bill - Hansard - UK Parliament: Hearing before the Commons Chamber*, UK Parliament (2021). https://hansard.parliament.uk/Commons/2021-07-12/debates/3E5A48AD-72E6-420A-910A-9F1863983743/HigherEducation(FreedomOfSpeech)Billhighlight=morally+wrong

Körner, A., & Volk, S. (2014). Concrete and abstract ways to deontology: Cognitive capacity moderates construal level effects on moral judgments. *Journal of Experimental Social Psychology*, *55*, 139–145. https://doi.org/10.1016/j.jesp.2014.07.002

Kreps, T. A., & Monin, B. (2014). Core Values Versus Common Sense: Consequentialist Views Appear Less Rooted in Morality. *Personality and Social Psychology Bulletin*, *40*(11), 1529–1542. https://doi.org/10.1177/0146167214551154

Kumle, L., Võ, M. L.-H., & Draschkow, D. (2021). Estimating power in (generalized) linear mixed models: An open introduction and tutorial in R. *Behavior Research Methods*, *53*(6), 2528–2543. https://doi.org/10.3758/s13428-021-01546-0

Lammers, J. (2012). Abstraction increases hypocrisy. *Journal of Experimental Social Psychology*, *48*(2), 475–480. https://doi.org/10.1016/j.jesp.2011.07.006

Ledgerwood, A., Trope, Y., & Chaiken, S. (2010). Flexibility now, consistency later: Psychological distance and construal shape evaluative responding. *Journal of Personality and Social Psychology*, *99*(1), 32–51. https://doi.org/10.1037/a0019843

Maier, M., Bartoš, F., Oh, M., Wagenmakers, E.-J., Shanks, D., & Harris, A. (2022). *Adjusting for Publication Bias Reveals That Evidence for and Size of Construal Level Theory Effects is Substantially Overestimated*. OSF. https://doi.org/10.31234/osf.io/r8nyu

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

McHugh, C., Zhang, R., Karnatak, T., Lamba, N., & Khokhlova, O. (2023). Just wrong? Or just WEIRD? Investigating the prevalence of moral dumbfounding in non-Western samples. *Memory & Cognition*, *51*(5), 1043–1060. https://doi.org/10.3758/s13421-022-01386-z

Monin, B., Pizarro, D. A., & Beer, J. S. (2007). Deciding versus Reacting: Conceptions of Moral Judgment and the Reason-Affect Debate. *Review of General Psychology*, *11*(2), 99–111. https://doi.org/10.1037/1089-2680.11.2.99

*Parents have sacred role in raising their children: Hearing on H1198 before the House of Representatives*, United States House of Representatives (2023). https://www.congress.gov/congressional-record/volume-169/issue-44/house-section/article/H1198-3

Powers, J. P., & LaBar, K. S. (2019). Regulating emotion through distancing: A taxonomy, neurocognitive model, and supporting meta-analysis. *Neuroscience & Biobehavioral Reviews*, *96*, 155–173. https://doi.org/10.1016/j.neubiorev.2018.04.023

Prinz, J. J. (2005). Passionate Thoughts: The Emotional Embodiment of Moral Concepts. In D. Pecher & R. A. Zwaan (Eds.), *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thinking* (pp. 93–114). Cambridge University Press.

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.

Sim, P. (2016, January 26). MSPs throw out incest petition. *BBC News*. http://www.bbc.com/news/uk-scotland-scotland-politics-35401195

Skitka, L. J. (2010). The Psychology of Moral Conviction. *Social and Personality Psychology Compass*, *4*(4), 267–281. https://doi.org/10.1111/j.1751-9004.2010.00254.x

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance., Construal-Level Theory of Psychological Distance. *Psychological Review, Psychological Review*, *117, 117*(2, 2), 440, 440–463. https://doi.org/10.1037/a0018963,%252010.1037/a0018963

Xiao, W., Wu, Q., Yang, Q., Zhou, L., Jiang, Y., Zhang, J., Miao, D., & Peng, J. (2015). Moral Hypocrisy on the Basis of Construal Level: To Be a Utilitarian Personal Decision Maker or to Be a Moral Advisor? *PLOS ONE*, *10*(2), e0117540. https://doi.org/10.1371/journal.pone.0117540

Žeželj, I. L., & Jokić, B. R. (2014). Replication of Experiments Evaluating Impact of Psychological Distance on Moral Judgment. *Social Psychology*. https://econtent.hogrefe.com/doi/10.1027/1864-9335/a000188