Appalachian
STATE UNIVERSITY

Department of Economics Working Paper

Number 26-02| February 2026

Consequential ethical dilemmas: The payoff-Trolley game.

David Dickinson
Appalachian State University

Department of Economics
Appalachian State University
Boone, NC 28608
Phone: (828) 262-2148
Fax: (828) 262-6105
www.business.appstate.edu/economics

Consequential ethical dilemmas: The payoff-Trolley game.

David L Dickinson

Affiliations: Appalachian State University (Economics and CERPA); IZA@LISER; ESI

ORCID 0000-0003-3899-0276 dickinsondl@appstate.edu

Abstract

This paper introduces a novel Payoff-Trolley dilemma task, where other participants' payoffs were at stake on the main or sidetrack in a trolley dilemma. Different scenarios varied the number of "others' payoffs" on the sidetrack and main track in ways that helped identify more clearly immoral choices of commission and omission, and utilitarian choices. Study participants had also made hypothetical Trolley dilemma choices in a separate study 4-5 years prior, allowing for a direct comparison of hypothetical and consequential moral dilemma choices. One key finding is that past hypothetical choices are statistically significant predictors of present consequential choices in the Payoff-Trolley task. Also, we find that one's degree of cognitive reflection is the most robust person-specific characteristic that predicts choices—higher cognitive reflection predicts more utilitarian choices, a reduced likelihood of immoral acts of commission and omission, and it impacts one's sensitivity to immoral choices for a given level of net-harm present in the scenario. These results hope to bridge a gap in our understanding of how choices in hypothetical moral dilemmas inform behaviors in consequential moral dilemmas.

Keywords: Moral choice, experiments, Trolley dilemma, dark personality, cognitive reflection

JEL codes: C9, D61, D91, I31

INTRODUCTION

The Trolley dilemma (Foot, 1967) has been of interest for decades as a way to study difficult moral trade-offs, examine factors that influence sacrificial decisions, and understand differences in what may be considered ethical behavior (Thomson, 1976, 1984; Spranca et al., 1991; Petrinovich et al., 1993; Green et al., 2001; Valdesolo and DeSteno, 2006; Shallow et al., 2011; Navarrete et al., 2012). Though variations abound, a classic Trolley dilemma assumes a runaway trolley is set to run over and kill 5 individuals on the trolley track. However, the decision maker may choose to pull a lever that diverts the trolley from its main track to a sidetrack. While this action saves the lives of the 5 individuals on the main track, it will unfortunately kill one individual who happens to be on the sidetrack. It is, of course, utilitarian to pull the lever, but deontological morals would suggest it is impermissible to *cause* harm by pulling the lever, because this action would actively cause the death of the individual on the sidetrack—a moral distinction is typically made between actively causing a death versus passively letting someone die (Kary, 1980).[1]

A common criticism is that the dilemma lacks realism (Bauman et al., 2014; Khazan, 2014). While true, and while some modifications to the trolley dilemma aimed to craft more realistic (if not still rare) scenarios (e.g., see Green et al., 2001, for a battery of moral dilemmas), many real-world choices scenarios already contain similarly difficult moral trade-offs. For example, should autonomous vehicles be programmed to make pure utilitarian choices to swerve or not when harm is inevitable (e.g., to different sets of pedestrians, or harm to pedestrian(s) or the vehicles occupants)? As a business application, should a small work unit be terminated to save a company and thus preserve a larger set of jobs? Or, in a medical setting, how should doctors make medical triage decisions, which amount to deciding who likely lives or dies?

So, Trolley-type dilemmas may not be as rare as first thought, but then the research efforts to systematically study behavior in such dilemmas have faced the problem of consequentiality. That is, the controlled experimental study of choices in Trolley dilemmas is almost inevitably hypothetical, but a core precept of designing a valid laboratory decision environment is that there be motivational relevance that makes one's decisions consequential in a way that mimics the incentives of the decision environment being studied (see Smith, 1982, on the precept of *saliency*). Because classic "Trolley" dilemmas involve life-and-death scenarios, the use of hypothetical Trolley dilemmas as a moral thought experiment is inevitable. Some researchers have attempted to mimic the life-and-death sacrificial dilemmas in mice (Bostyn et al., 2018), but these efforts have yielded mixed results and have not been without criticism (Plunkett and Green, 2019; Bostyn et al., 2025a). Still others have elicited choices over painful but medically safe electroshocks to human participants (Bostyn et al., 2025b). These same authors showed that choices in more standard hypothetical moral dilemmas had some predictive power regarding their decisions in the consequential electric-shock dilemmas (Bostyn et al., 2025b). Others have

[1] This distinction is also important in legal reasoning and application, though utilitarian choices that may be acceptable in tort law or regulatory policy would not be allowed in criminal law where deliberate killing is often punishable even if resulting in utilitarian outcome.

shown that hypothetical responses to Trolley-dilemmas predicted consequential anti-social behaviors in a distinct decision environment (Dickinson and Masclet, 2019).

The objective of the present paper is simple. We first introduce the consequential "Payoff-Trolley" dilemma, which replaces "people killed" in the hypothetical Trolley dilemma with "participants' monetary payoffs destroyed"—the environment is visualized as in the Trolley dilemma by showing images of money bags that represent an experimental payoff to study participants on trolley tracks. As has been done previously (see Dickinson and Masclet, 2019), the Trolley scenarios vary the number of others' payoffs, X, on the main track and those others' payoffs, Y, on the sidetrack such that varied (X:Y) dilemmas can highlight more unambiguously immoral behavior. For example, we define an act of *Immoral Omission* to be a choice to *not* pull the lever for any (X:Y) dilemma where X>Y. And an act of *Immoral Commission* would be a decision to pull the lever when X≤Y.

Another contribution of this paper is that we recruited from a previous set of online participants who had previously made decisions in hypothetical Trolley dilemmas—these previous choices occurred 4-5 years prior to the present study. Thus, we can test whether one's hypothetical choices predict their (future) consequential Payoff-Trolley choices. Finally, a third contribution is to examine the link between validated personality traits and characteristics like cognitive reflection and one's immoral (and/or utilitarian) choices in our consequential Payoff-Trolley setting, which is possible given that the participants in the previous study had completed a series of validated personality trait measures. These included the classic tetrad of dark personality traits (sadism, Machiavellianism, narcissism, psychopathy) and a more recent set of three light personality traits (humanism, faith in humanity, Kantianism).

Our results show that one's past hypothetical Trolley choices are a significant predictor of the likelihood that one makes utilitarian or immoral choices in the consequential Payoff-Trolley dilemma. That said, the data do not show the same link between dark or light personality traits and consequential Trolley choices as was found between these same traits and hypothetical choices. One's measured degree of cognitive reflection, however, is a significant predictor of both hypothetical and consequential Trolley choices, pointing to its potential value as a cognitive characteristic of individuals that can help better understand ethical decision-making.

METHODS

The study was preregistered on the Open Science Framework (doi.org/10.17605/OSF.IO/YKX5P) prior to the collection of data. Participants were recruited on the Prolific platform (Palan and Schitter, 2018; Peer et al., 2017), with preregistered plans to recruit 300 participants from a set of previous Prolific participants who took part in an earlier study reported in Dickinson (2025). Participants in this earlier study completed validated personality trait measures as well as a hypothetical set of 6 different life-and-death (i.e., typical) Trolley dilemma scenarios, which included (5:0), (1:1), (5:0), (5:5), (2:1), and (1,0) scenarios, where an (X,Y) scenario presents X individuals on the main track and Y individuals on the sidetrack.

The target sample size was based on an ex ante power analysis using G*Power 3.1.9.7, which found that an n=300 sample is sufficiently powered (power = .80) to detect small effect sizes (f^2 = .021) in a multiple regression evaluation of a single regression coefficient (i.e., one's past hypothetical Trolley choices).[2] This assumes an alpha level = .05, and a 1-tailed test of the directional hypothesis, as per our preregistration plans. Plans were to recruit n=300 participants from the original study drawing from those most and least "dark" in their validated personality traits, but many of our original participants were no longer active on Prolific 4-5 years later. In the end, we recruited from the full distribution of personality trait types reflected in those original participants, though approximately 80% of our sample enrolled was from the upper and lower quartile of the distribution of dark, relative to light, personality traits (note: this recruitment detail does not affect our calculated statistical power described above).

The Payoff-Trolley Task

The task replaces the standard "people on the tracks" with participant payoffs on the tracks (see full instructions in Appendix B). Every participant starts the task endowed with a $1 "bonus" payment that is separate from their $2 fixed Prolific study payment. Figure 1 showed two examples from among the 15 distinct Payoff-Trolley scenarios presented to the participants.[3] The Figure 1 upper panel shows a (5:1) dilemma, while the bottom panel shows the (1:3) dilemma. The set of 15 Payoff-Trolley (X:Y) Scenarios included all combinations of X, Y with X$\in$ [0,1,3,5] and Y $\in$ [0,1,3,5], except for X=Y=0. In each (X:Y) scenario, participants were asked to select "Do nothing" which meant the trolley would continue down the main track to run over and destroy payoff bags that may be on that main track, or "Pull the lever", which would divert the Trolley to the sidetrack where any payoff bag on the sidetrack would be destroyed. It was made clear to participants that a group of participants would be selected at random based on a scenario randomly selected to count for real bonus payoff determination—the selected scenario determined the size of the group necessary to have one decision-maker and others whose bonus payoffs were those on each of the two track(s). Among those in that anonymous group, one participant would be selected at random to be the decision-maker whose payoff bonus bag was safe (i.e., the decision-maker was guaranteed to receive the $1 bonus payment). The remaining participants in the group were randomly assigned to be those whose bonus payoff bags sat on either the main or sidetracks for that scenario, and the decision-maker's choice dictated whether those other participants received their bonus payoff or not.

[2] In the preregistration document, the detectable effect size was incorrectly listed as f^2 = .21 (as opposed to the correct f^2 = .021), which can be verified using the sample size, power, and alpha levels listed in the preregistration.

[3] As can be seen in Figure 1 (images here are unedited from what participants saw), the image inadvertently identified a $5 payoff in each payoff bag in the figure legend, though instructions clearly indicated throughout that each participant's Prolific bonus payoff was a $1 sum (consistently and in multiple locations in the instruction—see Appendix B for the full experiment instructions). None of the messages sent by participants to the experimenter via the Prolific messaging feature made any reference to confusion about bonus payment levels (for those whose payoff was not destroyed). As such, we feel it is unlikely this inconsistency influenced decisions at all, and it was clear that the choices elicited were consequential sacrificial choices over anonymous participant payoffs as described in the instructions.

Outcome measures from the Payoff-Trolley task were constructed to identify both utilitarian and immoral choices. Following Dickinson and Masclet (2019), we constructed variables to indicate immoral choices or two types: immoral acts of omission and immoral acts of commission. Of course, the classic (5:1) Trolley dilemma is precisely a dilemma because one can make a moral argument for pulling the lever or doing nothing, depending on one's view of morality (i.e., utilitarian vs. deontological). Such scenarios, therefore, cannot identify what we wish to classify as a more unambiguous immoral act. For our purposes, we defined a binary variable, *Omission Act,* equal to 1 for instances where the participant chose to do nothing in one or more (X:Y) scenarios where X > Y: the (1:0), (3:0), (5:0), (3:1), (5:1), and (5:3) scenarios. In contrast, we defined *Commission Act* equal to 1 for scenarios (X:Y) where $X \leq Y$ and the participant chose to pull the lever at least once—the (0:1), (0:3), (0:5), (1:3), (1:5), and (3:5) scenarios. A more loose definition of *Commission Act* included among the set of scenarios evaluated the (1:1), (3:3), and (5:5) dilemma scenarios, where the same number of payoffs are destroyed no matter what is done—it is generally recognized that individuals prefer passive compared active responsibility of a negative outcome (Descioli et al., 2011). Finally, we defined *Utilitarian Act* =1 for all (X:Y), $X \neq Y$ where the participant's choices (do nothing or pull the lever) in all such scenarios resulted in the least number of bonus payments being destroyed. An alternative indicator, *Sacrificial Utilitarian Act,* focused exclusively on scenarios where $X \neq Y$ *and* at least one payoff bag was on each track—the (1:3), (1:5), (3:5), (5:3), (5:1), and (3:1) scenarios.

Independent Control Measures

Personality characteristics available on the study participants were derived from their survey responses 4-5 years earlier, when recruited for the original study. These included short-form versions of dark personality measures (subclinical psychopathy, narcissism, Machiavellianism--Jones and Paulhus, 2014; everyday sadism--Plouffe et al., 2017), and a light triad of personality measures (Kantianism, faith in humanity, humanism: Kaufman et al., 2019). The Dark Tetrad $\in[1,5]$ of personality traits is operationalized here as the average of the measures of each dark personality trait, measured on a common 5-point scale (5 denotes the strongest measure of the that trait). Similarly, the Light Triad $\in[1,5]$ is the average measure from each of the individual measures scored on a common 5-point scale. For purposes of our participant recruitment, we preregistered plans to preferentially recruit from the tails of the distribution of *NetLight* = (*Light Triad – Dark Tetrad*) measure. Original study measures also provided additional control variables we used in our analysis: basic demographics (age, sex, student status), a measure of social value orientation (Murphy et al., 2011), and a 6-item cognitive reflection task (CRT) to assess thinking style (Primi et al., 2016).

The current study elicited one's subjective state-level sleepiness (the Karolinska Sleepiness scale) just prior to decision-making and, after the Payoff-Trolley task, their self-reported effort and motivation in completing the task on a Likert scale.[4] Finally, the present study elicited a 7-point scale response to how strongly one currently feels on each of following 10 emotion dimensions: Happy, Enthusiastic, Interested, Determined, Proud, Irritated, Distressed, Ashamed,

[4] The subjective sleepiness scale (Åkerstedt and Gillberg, 1990) has been shown valid in measuring subjective sleepiness and has been shown to predict physiological and behavioral manifestations of sleepiness (Kaida, 2006).

Angry, Sad. Emotion elicitation occurred at two points during the survey: at baseline before, and again after completing the Payoff-Trolley task, which allows for testing of one the preregistered hypothesis aimed at replicating a finding in the earlier Dickinson (2025) study.

HYPOTHESES

The following hypotheses were preregistered prior to data collection. Other hypotheses tested from the data are considered exploratory in nature, including our examination of one's sensitivity to the size of the net harm/benefit from action.

> H1a: Dark Personality Traits will predict more immoral acts of omission and commission in the Payoff-Trolley dilemma
>
> H1b: Light Personality Traits will predict fewer immoral acts of omission and commission in the Payoff-Trolley dilemma
>
> H2: Light Personality Traits will predict more Utilitarian choices in the Payoff-Trolley dilemma
>
> H3: Dark personality types will report increased positive mood after making consequential Trolley dilemma choices (compared to Light personality types).
>
> H4: Immoral choices in hypothetical Trolley dilemma scenarios will predict immoral choices in the consequential Payoff-Trolley dilemma scenarios

Hypotheses H1-H3 are replication tests of the results found in Dickinson (2025), which administered a hypothetical life-and-death Trolley Dilemma task, with the novel consequential Payoff-Trolley task. The last hypothesis, H4, examines the predictive validity of one's hypothetical choices made 4-5 years prior on one's current consequential choices.

RESULTS

A total of n=301 participants completed the study. In summarizing the prevalence of choice outcomes in Figure 2, we defined several variables of interest. *Immoral Omission Participants* looks at the number of participants who committed at least one immoral act of omission from the set of relevant scenarios (i.e., did not pull the level when doing so would have implied fewer bonus payoffs destroyed, such as the (3:1) dilemma). Similarly, we define *Immoral Commission Participant* based on the stricter or the looser definition given in the Methods section, where at least one such immoral act of commission from among the relevant scenarios classified one as an *Immoral Commission Participant*. Finally, *Utilitarian Participants* are defined as those who made the utilitarian choice in *all* $X \neq Y$ scenarios or, in the case of sacrificial dilemmas only, we excluded from consideration those scenarios where the utilitarian choice would lead to zero

payoffs destroyed. The reader will note from this categorization that an individual can possibly be both a utilitarian and one who commits an immoral act of certain types. We intend these classifications to present an initial view of the prevalence of different choice tendencies.

Hypotheses 1a and 1b

We have separated the related hypotheses H1a and H1b because the personality trait data reveal that the same individual can possess both dark and light traits, which are treated as separate constructs (Kaufman et al., 2019), and so we favor not combining them for analysis but rather we separate H1a and H1b—when we make reference to participants as being "more dark" or "more light", however, this references the relative strength of the light compared to dark traits. Hypotheses H1a and H1b were tested using means tests as well as multi-variate non-linear Probit estimations, where the dependent variable captures whether an individual made at least one immoral choice of given type (from among all trials that could identify such a choice).

For means tests , binary measures *MoreDark* and *MoreLight* were constructed based on a median split in the data. Two-sample tests of proportions show that those *MoreDark* = 1 actually had a *lower* proportion of participants making immoral acts of omission (2.6% versus 6.8% of participants, z = -1.73, p = .0833 for a 2-tailed test), which would be opposite our preregistered hypothesis H1a. No other significant differences in proportions were found in testing H1a and all tests for H1b focusing on those higher or lower in light personality traits were statistically insignificant (p >.10 in all instances). We also conducted multi-variate Probit estimations that would control for additional person-specific characteristics that may impact choices in the Payoff-Trolley dilemma. The marginal effects are reported in Table 1. The highlighted cells focus on preregistered hypotheses tests, and we can see in Table 1 that the data fail to support the preregistered hypotheses H1a and H1b. The one instance of significance in column 1 indicates that those higher in validated dark personality traits were *less* likely to have made an immoral act of omission in the Payoff-Trolley scenarios, which supports the same finding in the unconditional tests of proportions.[5]

Hypothesis 2

Hypothesis 2 examines whether those higher in the Light Triad personality measure are more utilitarian in their choices than those lower in Light Triad measures—this finding was reported in Dickinson (2025). Two-sample proportions tests were used as an uncontrolled test of H2 using the median split of the light triad measure to generate a *MoreLight* = 1 and *MoreLight* =0 samples. We tested proportions differences both using the full set of Payoff-Trolley scenarios where there was an identifiable utilitarian choice (12 of the 15 total scenarios) as well as for those sacrificial dilemmas (i.e., where neither X nor Y was equal to 0). Results of these tests showed no significant differences in the proportion of participants who made utilitarian choices (p > .10 for both tests). Though not preregistered, we also conducted similar tests of Utilitarian

[5] If examining the proportion of immoral choices made, which is more of an intensive-margin analysis, we find that neither the dark tetrad nor the light triad predict the proportion of immoral choices of any type defined.

choices for the *MoreDark* (= 0 and 1) samples. For the set of 6 scenarios involving sacrificial Payoff-Trolley dilemmas, a higher proportion of *MoreDark* = 1 participants were found to make fully utilitarian choices (90.26%), as compared to those *MoreDark* = 0 participants (82.99%), though this was only marginally significantly different ($p = .06$).[6] Table 2 shows Probit estimation results (marginal effects and their significance) that support these proportions tests. Light personality trait measures do not significantly affect the proportion of participants deemed utilitarian, which means the data fail to support H2, although the data show some limited support for concluding that dark personality traits predict utilitarian behavior in the consequential sacrificial Payoff-Trolley dilemma (see Table 2 column 4).

Hypothesis 3

This hypothesis sought to examine a previous (and surprising) finding that dark personality types reported an improved emotion state after completing the hypothetical life-and-death Trolley dilemma task. Here, for the consequential version of the Trolley dilemma, we elicited emotion states both at baseline before and again after completion of the Trolley dilemma scenarios. A regression-based approach allows us to not only control for person-specific demographic characteristics but also control for whether reported changes in emotion are related to immoral or utilitarian choice tendencies. Because H3 was based on findings in Dickinson (2025) (note: these findings were relegated to the Appendix A in that paper), we replicate the approach of constructing a composite measure of one's net emotion change, which is less sensitive to variation and interpretation by the participants of each individual emotion dimension. The emotion dimensions of Happy, Excited and Satisfied were averaged along their common 1-7 response scales to create an average measure of positive emotion. Similarly, Angry, Irritated, Regret, and Disgust were averaged into a combined negative emotion measure, and the difference (Positive minus Negative emotions) is considered one's self-reported *Net Positive Emotion*—we considered that Surprised, Confused, and Alert were not clearly positive or negative emotions and so they were omitted from these compositive measures. This measure, which lies within the [-6 ,+6] interval, was constructed for one's emotion state at baseline (pre-Trolley task) and post-Trolley task. We then take the change in *Net Positive Emotion* (post-Trolley minus pre-Trolley measure) to be the *Net Positive Emotion Change* measure used for the analysis. Participants higher in dark personality traits (*MoreDark* = 1) self-reported a significantly lower *Net Positive Emotion* at baseline (*Net Positive Emotion* = 1.81) compared to those lower in dark personality traits (*Net Positive Emotion* = 2.36 (t-test t=2.51, p =.013).

[6] This test of utilitarian choices among more dark personality types was of interest because previous findings in the literature had reported that dark personality types are more utilitarian in hypothetical moral dilemmas (see Bartels and Pizarro, 2011). However, findings in Dickinson (2025) found a strong effect of the Light Triad on Utilitarian choice—thus, we preregistered this as an ex-ante hypothesis. Though of marginal significance, the direction of the difference in our present data from consequential choices would be more aligned with findings in the literature connecting utilitarian choice frequency to dark traits—the argument usually revolve around how dark personality types may lack the emotion that would typically drive deontological reactions that make pulling the lever in the Trolley dilemma more difficult.

To test whether dark traits promoted a positive change in emotion state after completing the Payoff-Trolley, regressions were run using the *Dark Tetrad* and, separately, the *Light Triad* as key independent variables in the analysis that included other control measures. Results are reported in Table 3. For these regressions, the only variable that approaches significance in predicting one's reported net emotion change is the indicator for whether one committed an immoral act of omission—the direction of the positive coefficient estimated indicates qualitatively that such participants reported an increase in their net positive emotion, although only at the $p = .079$ (when using *Dark Tetrad*) or the $p = .088$ (when using *Light Triad*) controls. To be fair, the analysis in Dickinson (2025) preregistered an emotion-related hypothesis specific to the dimensions of subclinical *Psychopathy* and *Machiavellianism*, though our findings are similar if examining individual specific dark traits as regressors rather than the combined *Dark Tetrad*. In short, these validated personality traits do not predict self-reported emotion changes. Though there is some evidence that those who made consequential immoral acts of omission reported an increase in positive emotion state, this effect is general and not based on personality type.

Hypothesis 4

The final preregistered hypothesis seeks to understand the predictive validity of past hypothetical Trolley choices on current consequential Payoff-Trolley Choices. Table 4 shows results of Probit estimations of the different key outcome measures describing immoral and utilitarian classifications, and for each model estimated we included as the key independent measure one's classification from the previous study over 4 years ago. The previous study from which our participants were drawn involved hypothetical Trolley choices from a more limited set of Trolley scenarios. Nevertheless, we classified participants based on past choices in the same manner as for our present choice classifications. Thus, a positive coefficient estimate (and marginal effect) on the indicator variable for one's past choices would support H4. Results from columns (1) and (4) indicate some support for the hypothesis that past hypothetical Trolley choices predict a higher probability of making a comparable current choice in the domains of immoral acts of omission and utilitarian choices. It is also worth noting H4 was directional but two-tailed test significance levels are highlighted in Table 4. If one evaluates significance using the appropriate one-tailed test (as was preregistered for such directional hypotheses) then the shaded cells result in column (1) is significant at $p = .035$ level and the column (4) finding is significant at the $p = .003$ level. Additionally, the one-tailed significance of these preregistered hypotheses would also conclude that the column 2 result is marginally significant ($p = .072$), and the column (5) effect is marginally significant at $p = .058$). Thus, results in Table 4 are generally supportive of Hypothesis 4.

Exploratory panel data analysis

Finally, we also analyzed the Payoff-Trolley choices as a panel data set to better understand the characteristics of a given Trolley dilemma scenario that predict a higher likelihood of action (i.e.,

a choice to pull the lever). Each participant made choices in 15 Payoff-Trolley scenarios. Models estimated were nonlinear Probit with robust standard errors clustered by participant. We estimated models without controls, with the set of participant-specific controls used previously (separating the models that include the *Dark Tetrad* versus *Light Triad* measures), and we add in columns (4) and (5) in Table 5 to include a comprehensive set of controls that interact the Main-Sidetrack payoff bag difference and the participant-specific characteristics. For this, we simplify the measure of *Main-Sidetrack* payoff bag difference into a singular variable, rather than the separate measures for positive, negative, and zero differences used in columns (1) – (3) of Table 5. Otherwise, we would be faced with a rather burdensome set of interaction terms.

Table 5 results in columns (1) through (3) highlight two things. First, participants are responsive to the difference in payoff bags on the two sets of tracks when deciding whether to pull the lever, and the differences are in the expected directions. Secondly, there is a heightened responsiveness to the negative Main-Side deviations and the Main-Side no difference scenarios compared to the Main-Side positive deviation scenarios (the magnitude of the effect for positive versus negative Main-Side deviations is different at the $p < .05$ level for the linear coefficient restriction test). Specifically, more payoff bags on the sidetrack compared to main track (or an equal number on each set of tracks) predicts a significantly lesser likelihood of pulling the lever as compared to how much the same payoff bags difference promotes pulling the lever when the greater number of payoff bags is on the main tracks. This likely reflects the status quo effect whereby it is somewhat costly to deviate from the status quo and take action, even when it is the utilitarian outcome.

To examine whether one's sensitivity to the size of the difference represented in a Payoff-Trolley dilemma may interact with participant-specific personality or cognitive characteristics, we simplify the measure of Main-Side payoff bags differences to a singular measure called *MSD*, which ranges from -5 to +5, and we interact this with person-specific control measures in the columns (4) and (5) estimation results of Table 5. Interestingly, the estimation results in columns (4) and (5) show that the *MSD* is not estimated to have a main effect impact on the likelihood of pulling the lever. Rather, the effect is entirely dependent on one's *CRTscore* measure. To clarify how one's CRT score interacts with attributes of the Payoff-Trolley scenario, a final examination is reported in Table 6. Here, a simplified version of the specifications in the previous Table 5, columns (4) and (5) is presented in column (1) of Table 6—the simplifications removed the interaction terms that were statistically insignificant in the Table 5 estimations. Then, column (2) of Table 6 uses the more specific independent variables that separate the different comparisons of payoff bags on the main and sidetracks in the Payoff-Trolley scenario. The qualitative ordering of the coefficient estimates on these Main-Sidetracks control variables in Table 6 (column (6)) is comparable to columns (1)-(3) of Table 5. However, the sign and significance of the variables interacting the Main-Sidetrack payoff difference with one's CRT score indicates a lower marginal effect on the probability of pulling the lever when there are equal numbers of payoff bags on each track (though is it only marginally significant: *p*

< .10 for the 2-tailed test). And the marginal effect on pulling the lever is significantly lower when there are more participants' payoff bags on the side compared to main tracks ($p < .05$ for the 2-tailed test). Figure 3 graphically depicts forecasted interaction effects and predicted probabilities of pulling the lever for different Main-Side payoff bag differences for stylized participants with CRT score = 2 (low) compared to CRT score = 5 (high).[7] The result can be interpreted in two ways. First, this finding would be consistent with the conclusion that higher CRT score participants are more sensitive to being utilitarian when inaction aligns with utilitarianism. A second interpretation might be that higher CRT participants are significantly more averse to committing immoral acts of commission—both in the case of equally harmful passive versus active harm (i.e., X = Y scenarios), and when the result of pulling the lever actively increases net harm.

DISCUSSION AND CONCLUSIONS

One aim of this paper was to test hypotheses focused on the link between validated personality traits and choices in a consequential Payoff-Trolley dilemma. Another objective was to examine how one's past choice in hypothetical life-and-death Trolley dilemmas relate to their present choices in consequential dilemmas. The data showed little support for our preregistered hypotheses regarding dark or light personality traits and choices in the Payoff-Trolley dilemma. Though the data failed to support those hypotheses, there was some evidence that past choices from (several years prior) in hypothetical Trolley dilemmas predicted choices in the Payoff-Trolley dilemma. In other words, our findings support the claim that immoral or utilitarian choices in the classic hypothetical Trolley dilemma can predict immoral and utilitarian choices in the consequential Payoff-Trolley dilemma.

It may surprise the reader that dark personality traits predict hypothetical immoral Trolley choices, and immoral Trolley choices predict analogous consequential Payoff-Trolley immorality, yet dark traits in those same individuals do not predict consequential Payoff-Trolley immorality. These findings are consistent with the conclusion that non-personality determinants of hypothetical Trolley immorality are the main channel through which hypothetical choices map to consequential choices. It is also possible that unmeasured attributes of one's personality may be the primary drivers of consequential immoral choices, and these unmeasured attributes correlate to some extent with dark/light personality trait measures. The earlier research on hypothetical Trolley choices in these same participants focused its hypotheses on validated dark and light personality traits measures, but data were also collected on participant's validated short-version Big-5 personality trait measures (Gosling et al., 2003). There were some instances where Big-5 trait measures had some predictive power regarding choices in hypothetical Trolley

[7] Here, only those coefficient estimates that were statistically significant at $p < .10$ or better were included in the forecast—the constant term in the Probit estimation was not statistically different from zero ($p > .10$). For *Age*, we evaluated the marginal effect at the mean sample age of 48 years of age.

scenarios (see Dickinson, 2025). However, we conducted further analysis to include these measures as additional co-variates in our present analysis, and results showed little systematic evidence that Big-5 traits predicted consequential Payoff-Trolley choices (and our main findings are similar with their inclusion)—see Appendix Tables A1-A3 for these estimation results.

One measure collected previously on these original participants (Dickinson, 2025) that was not included as a co-variate in that previous study was the 6-item measure of cognitive reflection (*CRTscore*), which had significant and robust predictive power regarding consequential Payoff-Trolley choices in the present study. We revisited the estimations from that earlier study and found that the inclusion of *CRTscore* as a covariate does not affect the findings reported in Dickinson (2025), but *CRTscore* does provide additional and significant explanatory power regarding hypothetical Trolley choice outcomes. Specifically, higher *CRTscore* is associated with significantly lower likelihood of immoral (hypothetical) acts of omission and commission, *and* a significantly lower likelihood of being classified as a utilitarian (these findings on this separate data set are available on request). In the present data set, we find that *CRTscore* robustly predicts a lower likelihood of immoral acts of omission and commission but a *greater* likelihood of being classified as a utilitarian decision-maker. Thus, attributes captured in one's *CRTscore* may be what links one's hypothetical Trolley choices to their consequential Payoff-Trolley choices, rather than the link being one's dark or light personality trait measures or one's Big-5 characteristics. This may help explain why hypothetical Trolley choices predict consequential Payoff-Trolley choices for immoral choices. However, this connection is less clear regarding utilitarian choices given that *CRTscore* predicts a lesser likelihood of making hypothetical utilitarian choicen but a greater likelihood of consequential utilitarian choices. Yet another possibility is that one's personality traits may moderate the estimated *CRTscore* effect, which is also an area for future research.[8]

While future research is needed to more fully sort out the connections between one's hypothetical versus consequential choices in the domain of moral dilemmas, the present paper still offers some useful take-away messages. First, we generally find little support for the hypothesis that the same personality trait measures that predict hypothetical immoral and utilitarian choices also predict consequential choices in the novel Payoff-Trolley game. Given the prevalence of the classic Trolley dilemma as a thought experiment for a better understanding of choices in the face of difficult moral trade-offs, it is of significant interest to document that key person-specific characteristics predicting consequential choices in analogous dilemmas may differ in important ways. Our data also support the conclusion that those making immoral or utilitarian choices in hypothetical dilemmas are more likely to make similar choices in

[8] Analogous probit estimations to those in Table 6 (column (2)) were conducted on the subset of those participants more and less dark in validated personality traits. These results are shown in Appendix Table A4, Figure A1 and A2, and they are suggestive that the less dark personality trait individuals may be more or less sensitive to immoral acts of commission that more dark personality trait individuals, depending on whether the less dark type is low or high in CRT score, respectively (compare forecasts in Figure A1 and A2).

consequential settings. Our exploratory analysis finds that one's tendency towards cognitive reflection may be an important factor in understanding the link between hypothetical and consequential choices in the domain of moral dilemmas. Exploratory analysis also highlighted that *CRTscore* may be important in revealing those who will be most sensitive to immoral acts of commission for a given set of trade-offs in the specific dilemma.

Of course, the limitations of this research be acknowledged. While they are consequential, our Payoff-Trolley dilemmas are not high-stakes, and so replication with varied stake sizes would help identify the generality of our findings. There may also exist a variety of factors that contribute to how one views others' payoff losses. Or it may matter whether one's action is further removed from direct responsibility. For example, researchers in hypothetical dilemmas have found that one preference for utilitarian outcomes in autonomous vehicle scenarios may depend on key situational details such as the characteristics of those who may be saved or harmed (Awad et al., 2018). And choices in the "footbridge" version of the Trolley dilemma (where an individual is sacrificed by being pushed onto the track, rather than by diverting a Trolley via the lever pull) reveal that the emotional engagement generated with a more personal action reduces one's willingness to make a Utilitarian choice. How one views a moral dilemma depends not just on the end-consequences but also on the means used, the side effects of the means, and more (see Mikhael, 2007, on this point, and on its important legal implications).[9]

This paper introduced the consequential *Payoff-Trolley* version of a classical moral dilemma that has long served as a standard and valuable thought experiment in moral philosophy and business ethics. Our findings underscore the value of a better understanding of how individuals weigh the various situational elements of hypothetical moral dilemmas—such an understanding offers important insights into one's likely behavior in related consequential choice. Exploratory analysis helped highlight the importance of cognitive reflection in understanding the likelihood with which one may choose to pull the Trolley lever. We also documented the asymmetry that may exist in one's sensitivity to net-harm of the dilemma, which depends on whether it the outcome is achieved via action or inaction, and how this differs based on one's measure of cognitive reflection—this is an asymmetry reflects the asymmetry in how different individuals view an immoral act of omission versus commission. Whether such sensitivity differences may also be moderated by personality (such as the dark or light tendencies) represents testable hypotheses that can be pursued with further research aimed at deepening our understanding of what affects choices in the important area of moral dilemma.

[9] Others have argued that moral dilemmas, such as the Trolley dilemma, do not fully capture concepts like utilitarianism if one does not separately examine the altruistic component of utilitarianism (i.e., to achieve the least net harm) from a permissive attitude towards harm (Kahane et al., 2018). For this reason, we included (X:Y) scenarios that removed this confound that is inherently present in traditional (X:Y), X>0, Y>0 dilemmas to better identify unambiguous immoral choices that favor the permissive attitude towards harm.

Table 1: Tests of Hypotheses 1a and 1b

Probit estimations: Coefficient estimates shown with standard errors in parentheses

	Dependent Variable is binary indicator for a participant's choice designation listed at the top of each column					
	(1)	(2)	(3)	(4)	(5)	(6)
INDEPENDENT VARIABLES	Immoral Omission	Immoral Commission (Loose define)	Immoral Commission (Strict define)	Immoral Omission	Immoral Commission (Loose define)	Immoral Omission (Strict define)
Dark Tetrad	-0.79**	0.15	-0.24			
	(0.37)	(0.19)	(0.25)			
Light Triad				0.07	-0.17	-0.05
				(0.24)	(0.15)	(0.19)
Age (years)	0.01	0.03***	0.01*	0.02*	0.03***	0.02**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Female (=1)	-0.57*	-0.13	-0.24	-0.33	-0.15	-0.16
	(0.33)	(0.20)	(0.25)	(0.30)	(0.20)	(0.24)
Ksleepy$\in$ [1,9]	0.08	-0.01	-0.00	0.06	-0.02	-0.02
	(0.08)	(0.06)	(0.07)	(0.08)	(0.06)	(0.07)
SVO-ratio	0.45	0.06	-0.05	0.77	0.06	0.10
	(0.60)	(0.38)	(0.46)	(0.59)	(0.37)	(0.45)
CRT score $\in$ [0,6]	-0.19**	-0.26***	-0.20***	-0.17**	-0.25***	-0.19***
	(0.08)	(0.05)	(0.06)	(0.07)	(0.05)	(0.06)
Observations	293	293	293	293	293	293
Pseudo-R^2	0.148	0.151	0.0923	0.103	0.154	0.0876

Notes: *$p < .10$, **$p < .05$, ***$p < .01$ for the two-tailed test of significance on the coefficient estimate. A one-tailed test is appropriate for any preregistered directional hypothesis, and we shade the coefficient estimates and standard errors that serve to test the preregistered hypotheses. Dependent variables are indicator variables = 1 if the participant made one or more immoral choices defined by that column. SVO-ratio is the measure of social value orientation where higher values of this ratio indicate a more prosocial individual. Because 8 participants had missing demographic data in their Prolific profile, the sample used here for analysis with control measures was n=293 rather than the n=301 total number of participants.

Table 2: Tests of Hypothesis 2

Probit estimations: Coefficient estimates shown with standard errors in parentheses

	Dependent Variable is binary indicator for a participant's utilitarian choice designation listed at the top of each column			
INDEPENDENT VARIABLES	(1) Utilitarian Participant	(2) Utilitarian Participant (sacrificial dilemmas only)	(3) Utilitarian Participant	(4) Utilitarian Participant (sacrificial dilemmas only)
Light Triad	-0.05	-0.19		
	(0.17)	(0.18)		
Dark Tetrad			0.30	0.44*
			(0.21)	(0.23)
Age (years)	-0.02**	-0.01*	-0.01**	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Female (=1)	0.12	0.23	0.20	0.33
	(0.20)	(0.21)	(0.21)	(0.22)
Ksleepy$\in$ [1,9]	0.07	0.10	0.06	0.10
	(0.06)	(0.07)	(0.06)	(0.07)
SVO-ratio	-0.48	-0.38	-0.36	-0.28
	(0.38)	(0.39)	(0.38)	(0.39)
CRT score $\in$ [0,6]	0.14***	0.15***	0.14***	0.15***
	(0.05)	(0.05)	(0.05)	(0.05)
Observations	293	293	293	293
Pseudo-R^2	0.0787	0.0887	0.0866	0.101

Notes: *$p < .10$, **$p < .05$, ***$p < .01$ for the two-tailed test of significance on the coefficient estimate. A one-tailed test is appropriate for any preregistered directional hypothesis, and we shade the coefficient estimates and standard errors that serve to test the preregistered hypotheses. Dependent measures are indicator variables = 1 if the participant made utilitarian choices in all X≠Y dilemmas (columns 1 and 3) or in all sacrificial dilemmas (X≠Y, X≠0, Y≠0: columns 2 and 4). Other regressor are defined previously (in the text or in the note to Table 1).

Table 3: Test of Hypothesis 3

OLS regressions: Coefficient (standard error) shown

	Dependent variable in each estimation is the ***Net Positive Emotion Change*** *= (Positive Emotion - Negative Emotion compositive measure)*	
	(1)	(2)
INDEPENDENT VARIABLES	NetPosMoodChange	NetPosMoodChange
Light Triad		0.12
		(0.10)
Dark Tetrad	0.06	
	(0.12)	
Age (years)	0.00	0.00
	(0.00)	(0.00)
Female (=1)	0.04	-0.00
	(0.13)	(0.13)
Ksleepy∈ [1,9]	0.02	0.03
	(0.04)	(0.04)
SVO-ratio	-0.05	-0.14
	(0.23)	(0.23)
CRT score ∈ [0,6]	-0.05	-0.05
	(0.03)	(0.03)
Immoral Omission (=1)	0.58*	0.56*
	(0.33)	(0.33)
Immoral Commission (=1)	-0.08	-0.05
	(0.17)	(0.17)
Utilitarian (=1)	0.08	0.09
	(0.21)	(0.21)
Constant	-0.07	-0.32
	(0.49)	(0.52)
Observations	293	293
R-squared	0.03	0.03

Notes: *$p < .10$, **$p < .05$, ***$p < .01$ for the two-tailed test. A one-tailed test is appropriate for any preregistered directional hypothesis, and we shade the coefficient estimates and standard errors that serve to test the preregistered hypothesis. The dependent variable is the *Net Positive Emotion Change* measures defined in the text. Other regressor are defined previously (in the text or in the note to Table 1).

Table 4: Test of Hypothesis 4

Probit estimations: Coefficient estimates shown with standard errors in parentheses

	Dependent Variable is binary indicator for a participant's choice designation listed at the top of each column				
	(1)	(2)	(3)	(4)	(5)
INDEPENDENT VARIABLES	Immoral Omission	Immoral Commission (Loose define)	Immoral Commission (Strict define)	Utilitarian Participant	Utilitarian Participant (sacrificial dilemmas)
Immoral Omission (=1) In hypothetical Trolley	1.08*				
	(0.58)				
Immoral Commission (=1) In hypothetical Trolley		0.43	0.34		
		(0.29)	(0.33)		
Utilitarian (=1) In hypothetical Trolley				0.56***	
				(0.21)	
Utilitarian (=1) In hypothetical Sacrificial Trolley					0.35
					(0.22)
Age (years)	0.02**	0.02***	0.01*	-0.02***	-0.01**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Female (=1)	-0.22	-0.16	-0.14	0.14	0.22
	(0.30)	(0.20)	(0.24)	(0.20)	(0.21)
Ksleepy $\in$ [1,9]	0.03	-0.01	-0.01	0.08	0.12*
	(0.08)	(0.06)	(0.07)	(0.06)	(0.07)
SVO-ratio (Prosociality)	0.80	-0.05	0.05	-0.47	-0.43
	(0.60)	(0.36)	(0.45)	(0.37)	(0.38)
CRT score $\in$ [0,6]	-0.14*	-0.24***	-0.18***	0.16***	0.16***
	(0.07)	(0.05)	(0.06)	(0.05)	(0.05)
Observations	293	293	293	293	293
Pseudo-R^2	0.131	0.156	0.0931	0.107	0.0943

Notes: *$p < .10$, **$p < .05$, ***$p < .01$ for the two-tailed test of significance on the coefficient estimate. A one-tailed test is appropriate for any preregistered directional hypothesis, and we shade the coefficient estimates and standard errors that serve to test the preregistered hypotheses. Dependent measures are defined in Tables 1 and 2.

Table 5: Panel data analysis of lever-pull choice

Probit estimations: Coefficient estimates shown with standard errors in parentheses

	Dependent Variable is binary indicator *Pull Lever* = 1 if participant chose to pull the level in that trial (otherwise *Pull Lever* = 0)				
INDEPENDENT VARIABLES	(1)	(2)	(3)	(4)	(5)
Main-Sidetracks positive Dif	0.48***	0.48***	0.48***		
	(0.05)	(0.05)	(0.05)		
Main-Sidetracks negative Dif	-0.74***	-0.75***	-0.75***		
	(0.11)	(0.11)	(0.11)		
Main=Sidetracks Zero Dif	-1.42***	-1.46***	-1.46***		
	(0.13)	(0.13)	(0.13)		
Main-Sidetrack Dif (MSD)				0.16	0.53
				(0.37)	(0.39)
Age (years)		0.01**	0.00**	0.01**	0.01**
		(0.00)	(0.00)	(0.00)	(0.00)
Female (=1)		0.07	0.06	0.05	0.07
		(0.08)	(0.08)	(0.09)	(0.09)
Ksleepy$\in$ [1,9]		0.00	-0.00	-0.02	-0.03
		(0.02)	(0.02)	(0.03)	(0.03)
SVO-ratio		-0.18	-0.19	-0.10	-0.07
		(0.15)	(0.16)	(0.23)	(0.22)
CRT score $\in$ [0,6]		-0.06***	-0.06***	-0.09***	-0.09***
		(0.02)	(0.02)	(0.02)	(0.02)
Dark Tetrad		0.11		-0.01	
		(0.08)		(0.10)	
Light Triad			-0.09		-0.11
			(0.08)		(0.08)
MSD * Age				-0.00	-0.01
				(0.00)	(0.00)
MSD * Female				0.04	-0.03
				(0.10)	(0.11)
MSD * Ksleepy				0.04	0.05
				(0.03)	(0.03)
MSD * SVOratio				-0.08	-0.17
				(0.25)	(0.26)
MSD * CRT score				0.07***	0.07***
				(0.03)	(0.03)
MSD * Dark Tetrad				0.20	
				(0.13)	
MSD * Light Triad					0.04
					(0.09)
Observations	4,515	4,395	4,395	4,395	4,395
Pseudo-R^2	0.631	0.635	0.635	0.600	0.596

Notes: *$p < .10$, **$p < .05$, ***$p < .01$ for the two-tailed test of significance on the coefficient estimates.

Table 6: Panel data analysis of lever-pull choice

Probit estimations: Coefficient estimates shown with standard errors in parentheses

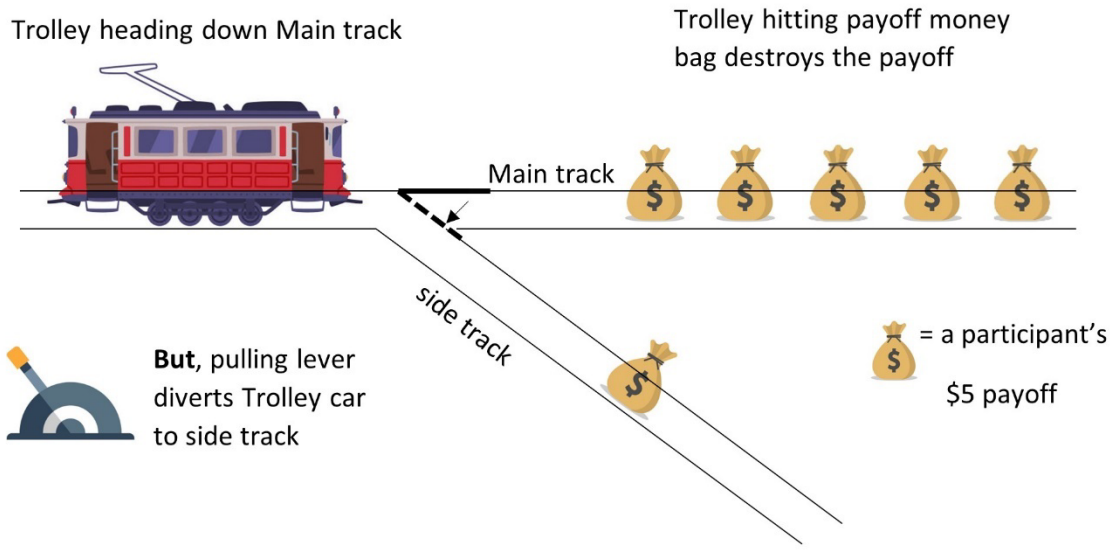
Dependent Variable is binary indicator *Pull Lever* = 1 if participant chose to pull the level in that trial (otherwise *Pull Lever* = 0)

INDEPENDENT VARIABLES	(1)	(2)
Main-Sidetrack Dif (MSD)	0.45***	
	(0.07)	
Main-Sidetracks positive Dif		0.41***
		(0.07)
Main-Sidetracks negative Dif		-0.50***
		(0.12)
Main=Sidetracks Zero Dif		-0.75***
		(0.19)
Age (years)	0.00**	0.01**
	(0.00)	(0.00)
Female (=1)	0.04	0.04
	(0.07)	(0.08)
Ksleepy$\in$ [1,9]	0.00	0.01
	(0.02)	(0.02)
SVO-ratio	-0.20	-0.24
	(0.14)	(0.16)
CRT score $\in$ [0,6]	-0.09***	0.04
	(0.02)	(0.05)
(Main-Side positive Dif) * CRT score		0.02
		(0.02)
(Main-Side negative Dif) * CRT score		-0.10*
		(0.05)
(Main-Side Zero Dif) * CRT score		-0.25***
		(0.07)
MSD * CRT score	0.07***	
	(0.03)	
Observations	4,395	4,395
Pseudo-R^2	0.586	0.649

Notes: *$p < .10$, **$p < .05$, ***$p < .01$ for the two-tailed test of significance on the coefficient estimates.

FIGURE 1: Payoff-Trolley scenarios

(5:1) Payoff-Trolley scenario



(1:3) Payoff-Trolley scenario

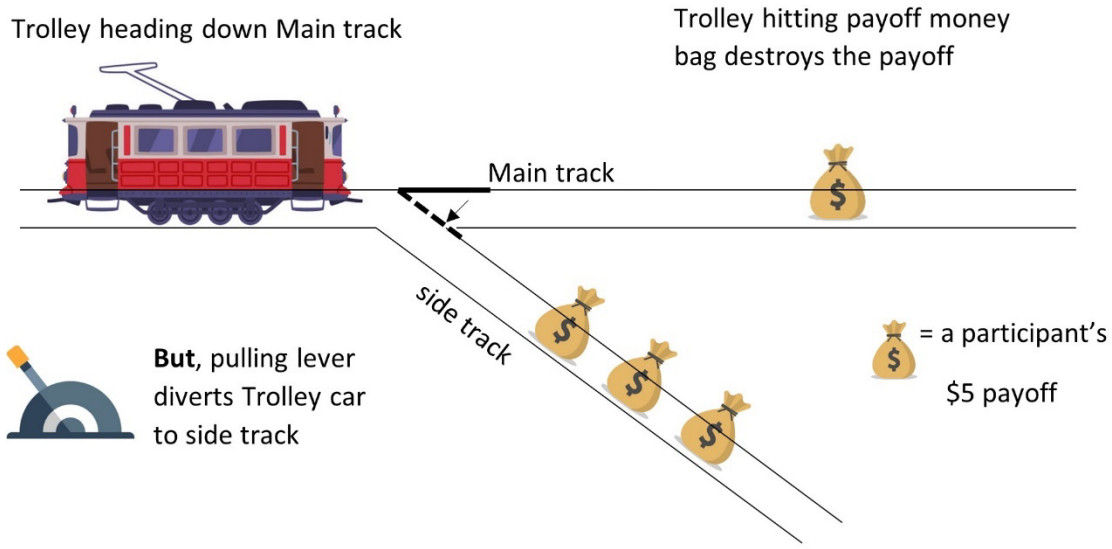
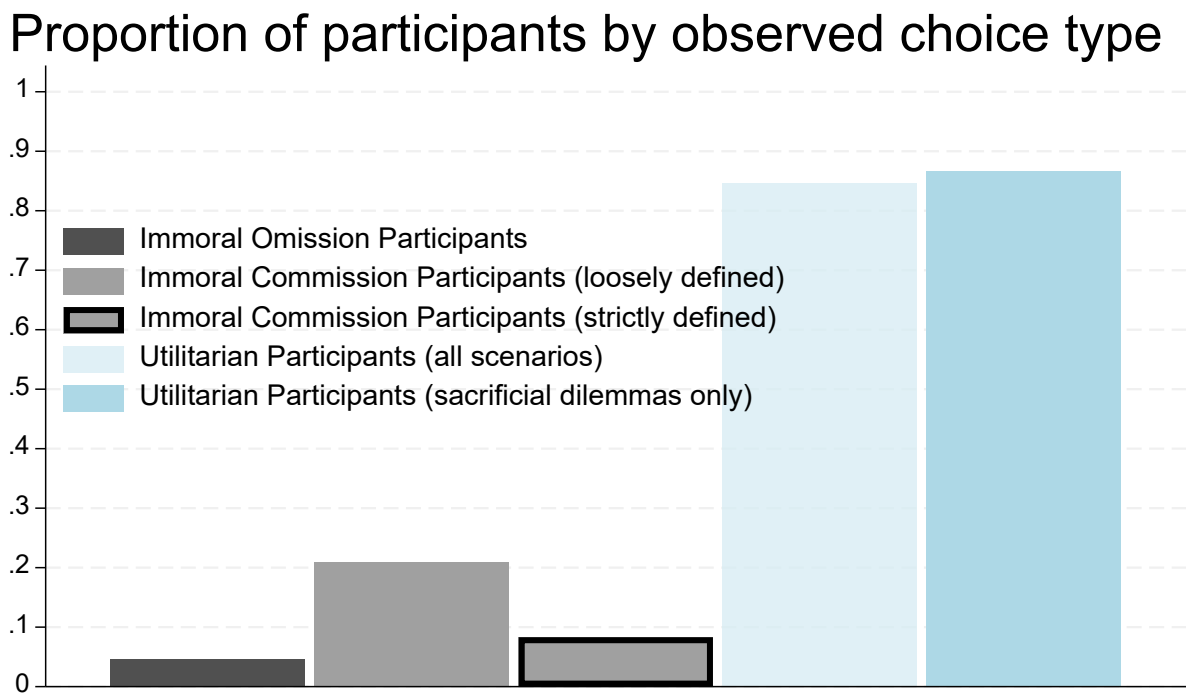
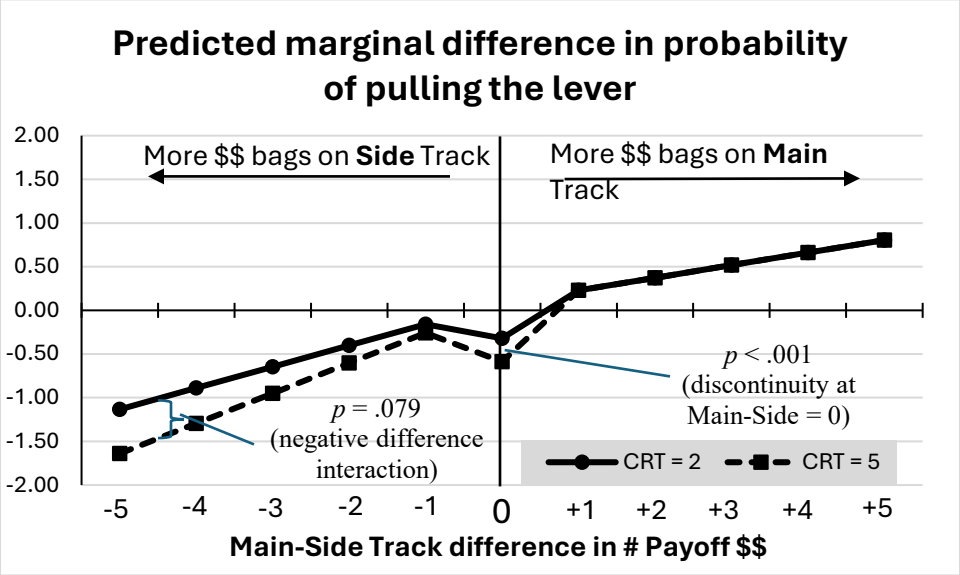


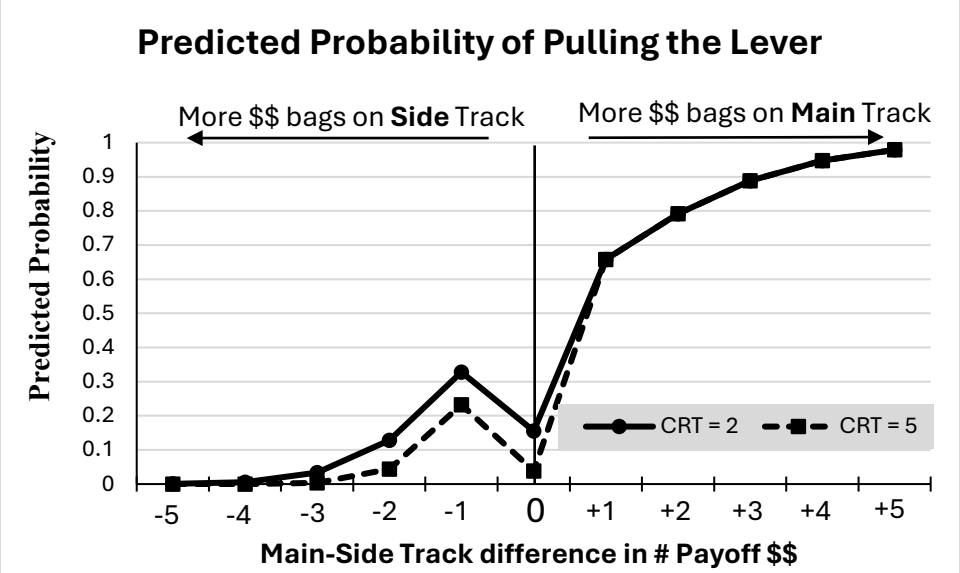
FIGURE 2



Note: Classification as one who is immoral (by commission or omission) is defined above as committing *at least one* immoral commission or omission act from the set of 15 Payoff-Trolley scenarios. Also, we note that a larger number of trials are considered eligible for classifying an individual as an "immoral commission participant", and so across-category classifications above should be viewed with some caution. A utilitarian designation implies one made non-weakly utilitarian across all 12 (X:Y) scenarios where $X \neq Y$ (otherwise, everyone would be labeled a utilitarian if it meant a utilitarian choice need only be made in at least 1 of the relevant 12 $X \neq Y$ scenarios.

FIGURE 3: Exploratory Analysis—the impact of CRT on Trolley cost sensitivity





Notes: Upper panel forecasts are based on the Probit model marginal effects (dF/dx effects) associated with the coefficient estimates reported in Table 6, column (2). The lower panel translates the Table 6 (column (2)) coefficient estimates into predicted probabilities for a stylized CRTscore = 2 and CRTscore = 5 individual.

REFERENCES

Abbink, K., & Herrmann, B. (2011). The moral costs of nastiness. *Economic Inquiry*, *49*(2), 631-633.

Abbink, K., & Sadrieh, A. (2009). The pleasure of being nasty. *Economics Letters*, *105*(3), 306-308.

Ahlenius, H., & Tännsjö, T. (2012). Chinese and Westerners respond differently to the trolley dilemmas. *Journal of Cognition and Culture*, *12*(3-4), 195-201.

Åkerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience*, *52*(1-2), 29-37.

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., ... & Rahwan, I. (2018). The moral machine experiment. *Nature*, *563*(7729), 59-64.

Bartels, D. M., & Pizarro, D. A. (2011). The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition*, *121*(1), 154-161.

Bauman, C. W., McGraw, A. P., Bartels, D. M., & Warren, C. (2014). Revisiting external validity: Concerns about trolley problems and other sacrificial dilemmas in moral psychology. *Social and Personality Psychology Compass*, *8*(9), 536-554.

Bostyn, D. H., Dierckx, K., Van Assche, J., Haesevoets, T., Valcke, B., De Souter, L., ... & Roets, A. (2025a). Hypothetical Judgment and Real-Life Decision-Making in Sacrificial Dilemmas: A Replication of "of Mice, Men and Trolleys". *Collabra: Psychology*, *11*(1), 143715.

Bostyn, D. H., Gouwy, M. C., De Craene, E., Vanmechelen, C., Scheirlinckx, J., Tissot, T. T., ... & Roets, A. (2025b). Beyond hypothetical trolleys: Moral choices and motivations in a real-life sacrificial dilemma. *Journal of Personality and Social Psychology*, *129*(5), 834-849.

Bostyn, D. H., Sevenhant, S., & Roets, A. (2018). Of mice, men, and trolleys: Hypothetical judgment versus real-life behavior in trolley-style moral dilemmas. *Psychological Science*, *29*(7), 1084-1093.

DeScioli, P., Bruening, R., & Kurzban, R. (2011). The omission effect in moral cognition: Toward a functional explanation. *Evolution and Human Behavior*, *32*(3), 204-215.

Dickinson, D. L. (2025). Dark versus light personality types and moral choice. *European Economic Review*, *178*(September), 105092.

Dickinson, D. L., & Masclet, D. (2019). Using ethical dilemmas to predict antisocial choices with real payoff consequences: An experimental study. *Journal of Economic Behavior & Organization*, *166*(October), 195-215.

Foot, P. (1967). The problem of abortion and the doctrine of double effect. *Oxford Review*, *5*, 5-15.

Gosling, S. D., Rentfrow, P. J., & Swann Jr, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*(6), 504-528.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*(5537), 2105-2108.

Jones, D. N., & Paulhus, D. L. (2014). Introducing the short dark triad (SD3) a brief measure of dark personality traits. *Assessment*, *21*(1), 28-41.

Kahane, G., Everett, J. A., Earp, B. D., Caviola, L., Faber, N. S., Crockett, M. J., & Savulescu, J. (2018). Beyond sacrificial harm: A two-dimensional model of utilitarian psychology. *Psychological Review*, *125*(2), 131-164.

Kaida, K., Takahashi, M., Åkerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., & Fukasawa, K. (2006). Validation of the Karolinska sleepiness scale against performance and EEG variables. *Clinical Neurophysiology*, *117*(7), 1574-1581.

Kary, C. (1980). A moral distinction between killing and letting die. *The Journal of Medicine and Philosophy*, *5*(4), 326-332.

Kaufman, S. B., Yaden, D. B., Hyde, E., & Tsukayama, E. (2019). The light vs. dark triad of personality: Contrasting two very different profiles of human nature. *Frontiers in Psychology, 10*, 467.

Khazan, O. (2014). Is one of the most popular psychology experiments worthless? *The Atlantic*, *July 24*.

Lim, H. S. M., & Taeihagh, A. (2019). Algorithmic decision-making in AVs: Understanding ethical and technical concerns for smart cities. *Sustainability*, *11*(20), 5791.

Mikhail, J. (2007). Universal moral grammar: Theory, evidence and the future. *Trends in Cognitive Sciences*, *11*(4), 143-152.

Murphy, R. O., Ackermann, K. A., & Handgraaf, M. J. (2011). Measuring social value orientation. *Judgment and Decision Making*, *6*(8), 771-781.

Navarrete, C. D., McDonald, M. M., Mott, M. L., & Asher, B. (2012). Virtual morality: Emotion and action in a simulated three-dimensional "trolley problem". *Emotion*, *12*(2), 364.

Palan, S., & Schitter, C. (2018). Prolific. ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22-27.

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology, 70*, 153-163.

Petrinovich, L., O'Neill, P., & Jorgensen, M. (1993). An empirical study of moral intuitions: Toward an evolutionary ethics. *Journal of Personality and Social Psychology*, *64*(3), 467.

Pletti, C., Lotto, L., Buodo, G., & Sarlo, M. (2017). It's immoral, but I'd do it! Psychopathy traits affect decision-making in sacrificial dilemmas and in everyday moral situations. *British Journal of Psychology*, *108*(2), 351-368.

Plouffe, R. A., Saklofske, D. H., & Smith, M. M. (2017). The assessment of sadistic personality: Preliminary psychometric evidence for a new measure. *Personality and Individual Differences, 104*, 166-171.

Plunkett, D., & Greene, J. D. (2019). Overlooked evidence and a misunderstanding of what trolley dilemmas do best: Commentary on Bostyn, Sevenhant, and Roets (2018). *Psychological Science*, *30*(9), 1389-1391.

Primi, C., Morsanyi, K., Chiesi, F., Donati, M. A., & Hamilton, J. (2016). The development and testing of a new version of the cognitive reflection test applying item response theory (IRT). *Journal of Behavioral Decision Making*, *29*(5), 453-469.

Shallow, C., Iliev, R., & Medin, D. (2011). Trolley problems in context. *Judgment and Decision Making*, *6*(7), 593-601.

Smith, V. L. (1982). Microeconomic systems as an experimental science. *The American Economic Review*, *72*(5), 923-955.

Spranca, M., Minsk, E., & Baron, J. (1991). Omission and commission in judgment and choice. *Journal of Experimental Social Psychology*, *27*(1), 76-105.

Thomson, J. J. (1976). Killing, letting die, and the trolley problem. *The Monist*, 59(2), 204-217.

Thomson, J. J. (1984). The trolley problem. *Yale Law Journal*, *94*, 1395-1415.

Valdesolo, P., & DeSteno, D. (2006). Manipulations of emotional context shape moral judgment. *Psychological Science 17*(6), 476-477.

Zizzo, D. J. (2003). Money burning and rank egalitarianism with random dictators. *Economics Letters*, *81*(2), 263-266.

Zizzo, D. J., & Oswald, A. J. (2001). Are people willing to pay to reduce others' incomes?. *Annales d'Economie et de Statistique*, 39-65.

APPENDIX A: Additional estimation results

Table A1: Tests of Hypotheses 1a and 1b

Probit estimations: Coefficient estimates shown with standard errors in parentheses

Dependent Variable is binary indicator for a participant's choice designation listed at the top of each column

INDEPENDENT VARIABLES	(1) Immoral Omission	(2) Immoral Commission (Loose define)	(3) Immoral Commission (Strict define)	(4) Immoral Omission	(5) Immoral Commission (Loose define)	(6) Immoral Omission (Strict define)
Dark Tetrad	-0.63	0.18	-0.17			
	(0.43)	(0.22)	(0.29)			
Light Triad				-0.06	-0.24	-0.05
				(0.32)	(0.20)	(0.24)
Age (years)	0.01	0.03***	0.01*	0.01	0.02***	0.02**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Female (=1)	-0.56	-0.13	-0.17	-0.37	-0.15	-0.13
	(0.35)	(0.21)	(0.26)	(0.33)	(0.20)	(0.25)
Ksleepy∈ [1,9]	0.07	-0.01	-0.03	0.05	-0.00	-0.04
	(0.09)	(0.06)	(0.08)	(0.09)	(0.06)	(0.08)
SVO-ratio	0.56	0.08	-0.08	0.76	0.09	-0.01
	(0.64)	(0.38)	(0.47)	(0.64)	(0.38)	(0.47)
CRT score ∈ [0,6]	-0.21**	-0.25***	-0.20***	-0.19**	-0.25***	-0.20***
	(0.08)	(0.05)	(0.06)	(0.08)	(0.05)	(0.06)
Extraversion	-0.03	0.03	-0.06	-0.04	0.05	-0.06
	(0.09)	(0.06)	(0.07)	(0.09)	(0.06)	(0.07)
Agreeableness	0.18	0.02	0.04	0.30	0.04	0.08
	(0.18)	(0.09)	(0.11)	(0.19)	(0.09)	(0.11)
Conscientiousness	-0.09	0.02	-0.11	-0.08	0.02	-0.10
	(0.13)	(0.08)	(0.10)	(0.13)	(0.08)	(0.10)
Emot Stability	0.12	-0.01	0.03	0.11	-0.00	0.02
	(0.11)	(0.07)	(0.09)	(0.11)	(0.07)	(0.09)
Openness	-0.24**	-0.08	-0.07	-0.26**	-0.07	-0.07
	(0.12)	(0.08)	(0.09)	(0.12)	(0.08)	(0.09)
Observations	293	293	293	293	293	293
Pseudo-R^2	0.212	0.156	0.110	0.193	0.159	0.109

Notes: *$p < .10$, **$p < .05$, ***$p < .01$ for the two-tailed test of significance on the coefficient estimate. A one-tailed test is appropriate for any preregistered directional hypothesis, and we shade the coefficient estimates and standard errors that serve to test the preregistered hypotheses. Dependent variables are indicator variables = 1 if the participant made one or more immoral choices defined by that column. SVO-ratio is the measure of social value orientation where higher values of this ratio indicate a more prosocial individual. Because 8 participants had missing demographic data in their Prolific profile, the sample used here for analysis with control measures was n=293 rather than the n=301 total number of participants.

Table A2: Tests of Hypothesis 2

Probit estimations: Coefficient estimates shown with standard errors in parentheses

	Dependent Variable is binary indicator for a participant's utilitarian choice designation listed at the top of each column			
INDEPENDENT VARIABLES	(1) Utilitarian Participant	(2) Utilitarian Participant (sacrificial dilemmas only)	(3) Utilitarian Participant	(4) Utilitarian Participant (sacrificial dilemmas only)
Light Triad	0.01	-0.17		
	(0.21)	(0.22)		
Dark Tetrad			0.25	0.39
			(0.24)	(0.26)
Age (years)	-0.01**	-0.01	-0.01*	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Female (=1)	0.11	0.24	0.17	0.32
	(0.21)	(0.22)	(0.22)	(0.23)
Ksleepy$\in$ [1,9]	0.07	0.10	0.06	0.09
	(0.07)	(0.07)	(0.07)	(0.07)
SVO-ratio	-0.50	-0.39	-0.42	-0.34
	(0.39)	(0.40)	(0.39)	(0.40)
CRT score $\in$ [0,6]	0.14***	0.15***	0.14***	0.15***
	(0.05)	(0.05)	(0.05)	(0.05)
Extraversion	0.01	0.01	0.01	-0.01
	(0.06)	(0.06)	(0.06)	(0.06)
Agreeableness	-0.07	-0.07	-0.03	-0.04
	(0.10)	(0.11)	(0.10)	(0.10)
Conscientiousness	-0.04	-0.07	-0.03	-0.07
	(0.09)	(0.10)	(0.09)	(0.10)
Emot Stability	-0.06	-0.04	-0.07	-0.05
	(0.07)	(0.08)	(0.08)	(0.08)
Openness	0.11	0.16*	0.11	0.15*
	(0.08)	(0.08)	(0.08)	(0.08)
Observations	293	293	293	293
Pseudo-R^2	0.0949	0.113	0.0993	0.120

Notes: *$p < .10$, **$p < .05$, ***$p < .01$ for the two-tailed test of significance on the coefficient estimate. A one-tailed test is appropriate for any preregistered directional hypothesis, and we shade the coefficient estimates and standard errors that serve to test the preregistered hypotheses. Dependent measures are indicator variables = 1 if the participant made utilitarian choices in all X≠Y dilemmas (columns 1 and 3) or in all sacrificial dilemmas (X≠Y, X≠0, Y≠0: columns 2 and 4). Other regressor are defined previously (in the text or in the note to Table 1).

Table A3: Test of Hypothesis 4

Probit estimations: Coefficient estimates shown with standard errors in parentheses

	Dependent Variable is binary indicator for a participant's choice designation listed at the top of each column				
	(1)	(2)	(3)	(4)	(5)
INDEPENDENT VARIABLES	Immoral Omission	Immoral Commission (Loose define)	Immoral Commission (Strict define)	Utilitarian	Utilitarian (sacrificial dilemmas)
Immoral Omission (=1) in hypothetical Trolley	1.42**				
	(0.67)				
Immoral Commission (=1) in hyp. Trolley		0.46	0.40		
		(0.30)	(0.34)		
Utilitarian (=1) in hypothetical Trolley				0.62***	
				(0.21)	
Utilitarian (=1) in hyp. sacrificial dilemmas					0.40*
					(0.23)
Age (years)	0.02	0.02***	0.02*	-0.02**	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Female (=1)	-0.30	-0.14	-0.11	0.15	0.26
	(0.33)	(0.20)	(0.25)	(0.21)	(0.22)
Ksleepy$\in$ [1,9]	0.01	-0.00	-0.04	0.08	0.11
	(0.10)	(0.06)	(0.08)	(0.07)	(0.07)
SVO-ratio	0.67	-0.00	-0.05	-0.45	-0.38
	(0.65)	(0.37)	(0.47)	(0.39)	(0.40)
CRT score $\in$ [0,6]	-0.16**	-0.24***	-0.19***	0.17***	0.17***
	(0.08)	(0.05)	(0.06)	(0.05)	(0.05)
Extraversion	-0.04	0.03	-0.07	0.00	-0.00
	(0.09)	(0.06)	(0.07)	(0.06)	(0.06)
Agreeableness	0.29*	-0.02	0.07	-0.09	-0.12
	(0.17)	(0.08)	(0.10)	(0.09)	(0.09)
Conscientiousness	-0.08	0.03	-0.10	-0.03	-0.07
	(0.13)	(0.08)	(0.10)	(0.09)	(0.10)
Emot Stability	0.09	0.01	0.03	-0.06	-0.04
	(0.11)	(0.07)	(0.09)	(0.08)	(0.08)
Openness	-0.31**	-0.08	-0.09	0.14*	0.17**
	(0.13)	(0.08)	(0.09)	(0.08)	(0.09)
Observations	293	293	293	293	293
Pseudo-R^2	0.229	0.161	0.116	0.128	0.124

Notes: *$p < .10$, **$p < .05$, ***$p < .01$ for the two-tailed test of significance on the coefficient estimate. A one-tailed test is appropriate for any preregistered directional hypothesis, and we shade the coefficient estimates and standard errors that serve to test the preregistered hypotheses.

Table A4: Panel data analysis of lever-pull choice—by More vs. Less Dark types

Probit estimations: Coefficient estimates shown with standard errors in parentheses

Dependent Variable is binary indicator *Pull Lever* = 1 if participant chose to pull the level in that trial (otherwise *Pull Lever* = 0)

	***More Dark* = 1**	***More Dark* = 0**
INDEPENDENT VARIABLES	(1)	(2)
Main-Side tracks positive Dif	0.41**	0.35***
	(0.16)	(0.08)
Main-Side tracks negative Dif	-0.99***	-0.28***
	(0.31)	(0.10)
Main=Side tracks Zero Dif	-1.20***	-0.54**
	(0.38)	(0.22)
Age (years)	0.01	0.01**
	(0.00)	(0.00)
Female (=1)	0.02	0.08
	(0.11)	(0.15)
Ksleepy$\in$ [1,9]	0.03	-0.01
	(0.04)	(0.03)
SVO-ratio	-0.04	-0.45*
	(0.19)	(0.27)
CRT score $\in$ [0,6]	-0.09	0.15***
	(0.07)	(0.05)
(Main-Side positive Dif) * CRT score	0.04	-0.00
	(0.04)	(0.03)
(Main-Side negative Dif) * CRT score	0.06	-0.24***
	(0.07)	(0.06)
(Main-Side Zero Dif) * CRT score	-0.10	-0.39***
	(0.09)	(0.08)
Observations	2,250	2,145
Pseudo-R^2	0.680	0.641

Notes: *$p < .10$, **$p < .05$, ***$p < .01$ for the two-tailed test of significance on the coefficient estimates. The unconditional mean level choices to pull the lever is similar in comparing choices from the *More Dark* = 1 subsample (mean = .414 ± .493) to those in the *More Dark* = 0 subsample (mean = .411 ± .492).

FIGURE A1: Table A4 forecasts for **low CRT participants**

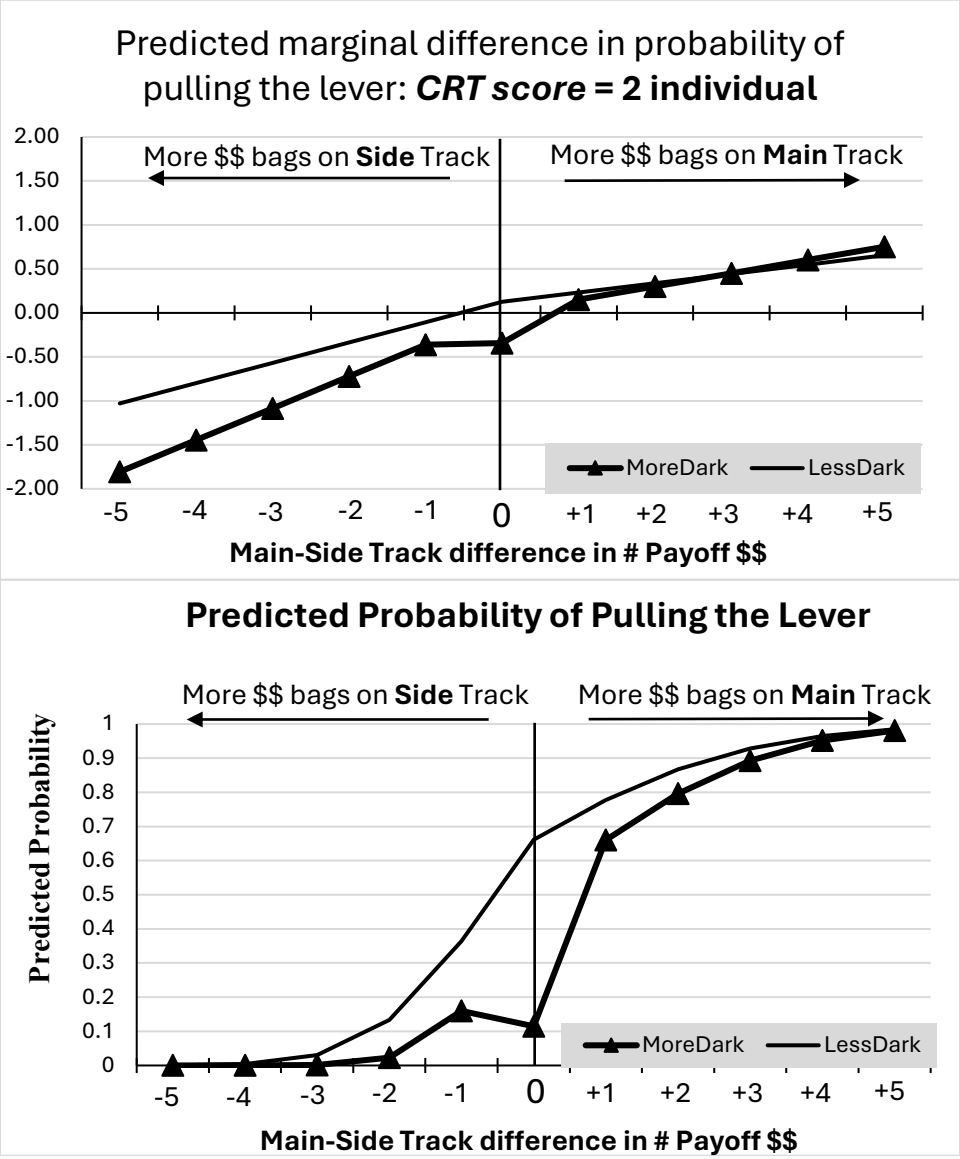
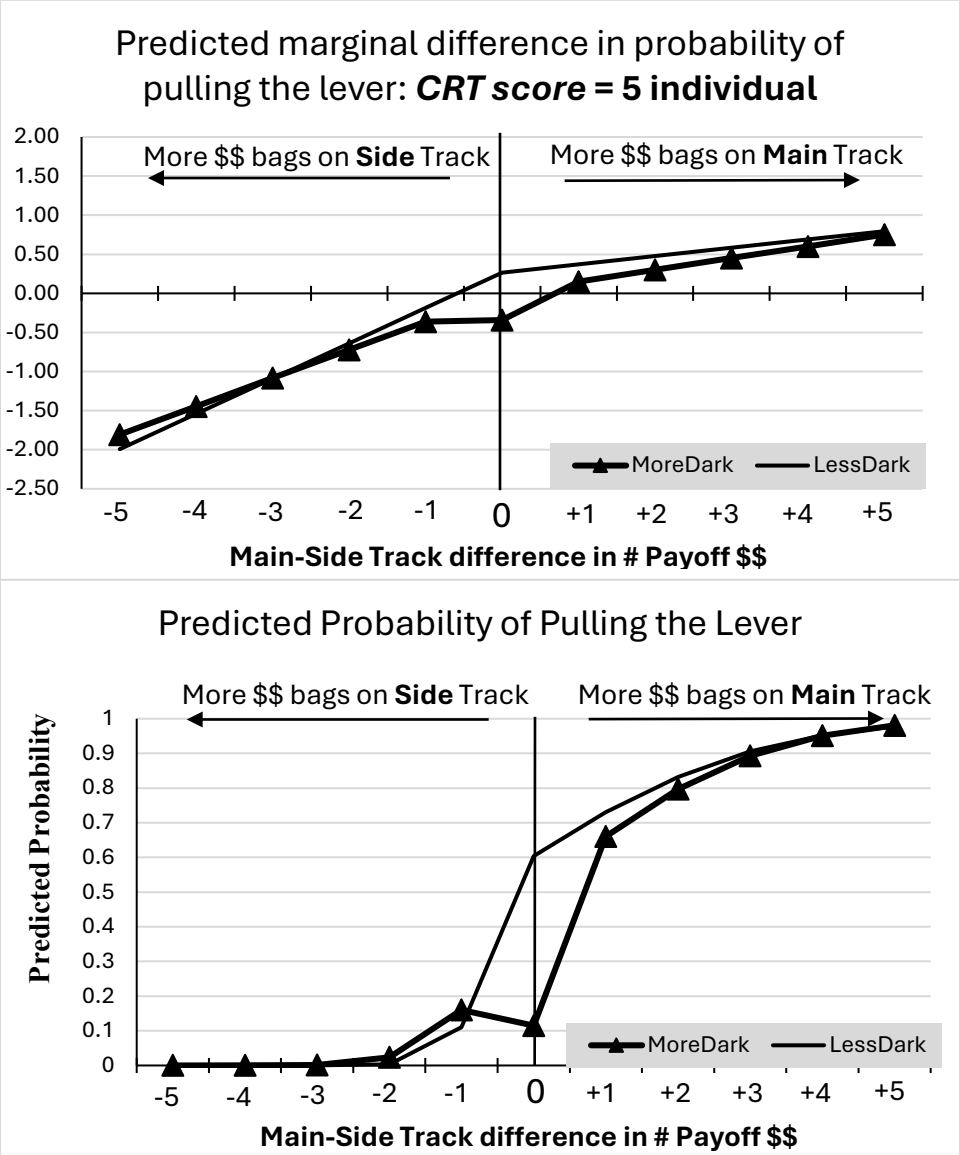


FIGURE A2: Table A4 forecasts for **high CRT participants**



Appendix B: Experiment Instructions

Note: document is a pdf export of the Qualtrics survey used to administer the experiment.

Consent Block Prolific

Politics (follow-up short survey)

Informed Consent:
You are being asked to complete this online survey that includes some mood/emotion questions and an incentivized task where you will make decisions about trade-offs among payoffs to others.

Participation in this online survey is completely voluntary, your responses to this survey will remain completely confidential, the data will be securely stored, your name will not be recorded anywhere on this survey. The only identifier we will record will be your Prolific ID, which we as researchers cannot link to personally identifiable data of yours.

This survey is estimated to take 13 minutes to complete and your payment for successful survey completion will be $2.00. Plus, **you will have the chance to earn a bonus payment of $1.00 based** on responses made by participants in the incentivized decision task within the survey.

There may be multiple attention-check questions within the survey to ensure that you are being attentive and reading each question prior to responding. Successful survey completion includes passing the attention check question. **Failure to answer attention-check questions correctly may jeopardize receiving your payment** for this Prolific study.

There are no known risks associated with this study beyond those associated with everyday life. Although this study will not benefit you personally, its results will help our understanding of how people make decisions about politics.

For additional information related to this questionnaire, contact Dr. David Dickinson, Department of Economics, Appalachian State University, at dickinsondl@appstate.edu. Appalachian State University's Institutional Review Board has determined this study to be exempt from review by the IRB Administration (828-262-4060).

- ○ **I Consent** and wish to continue with this study
- ○ **I do not consent** to participating and **do not wish to continue**

Prolific ID

Before you start, **please switch off phone/ e-mail/ music** so that you can focus on this study. Thank you!

Please carefully enter your Prolific ID

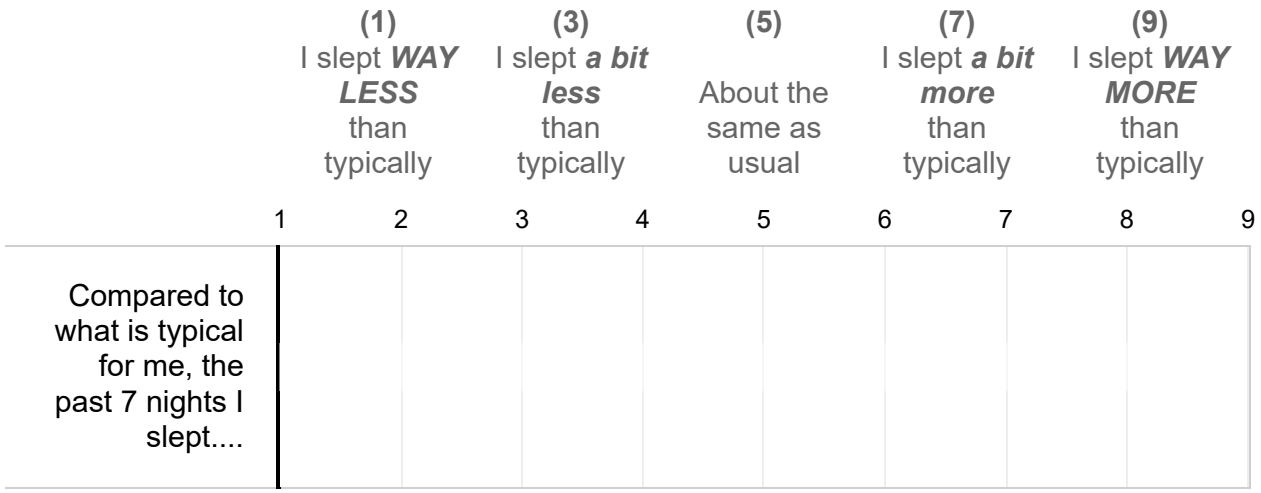
${e://Field/PROLIFIC_PID}

Self Report Sleep Mood Block NEW 3 questions

Karolinska Sleepiness Scale
Please mark the number that best corresponds to how sleepy you feel **right now**. You may mark any number, but mark only one number.

- ○ 1. Extremely alert
- ○ 2.
- ○ 3. Alert
- ○ 4.
- ○ 5. Neither alert nor sleepy
- ○ 6.
- ○ 7. Sleepy--but no difficulty remaining awake
- ○ 8.
- ○ 9. Extremely sleepy--fighting sleep

Using the scale below, please indicate **how much less or more you slept**, on average, **over the previous week compared to what is typical for you**.



Now, we'd like to get your baseline ratings on some mood/emotion states.

Please rate how strongly you feel each of these emotions ***right now***.

Right now I feel.....

	Zero level of this emotion **(1)**	**(2)**	**(3)**	Mid-Range level of this emotion **(4)**	**(5)**	**(6)**	Maximum level of this emotion **(7)**
Happy	○	○	○	○	○	○	○
Excited	○	○	○	○	○	○	○

	Zero level of this emotion **(1)**	**(2)**	**(3)**	Mid-Range level of this emotion **(4)**	**(5)**	**(6)**	Maximum level of this emotion **(7)**
Surprised	○	○	○	○	○	○	○
Satisfied	○	○	○	○	○	○	○
Angry	○	○	○	○	○	○	○
Irritated	○	○	○	○	○	○	○
Confused	○	○	○	○	○	○	○
Regret	○	○	○	○	○	○	○
Disgust	○	○	○	○	○	○	○
Alert	○	○	○	○	○	○	○

Trolley Block (with graphics)

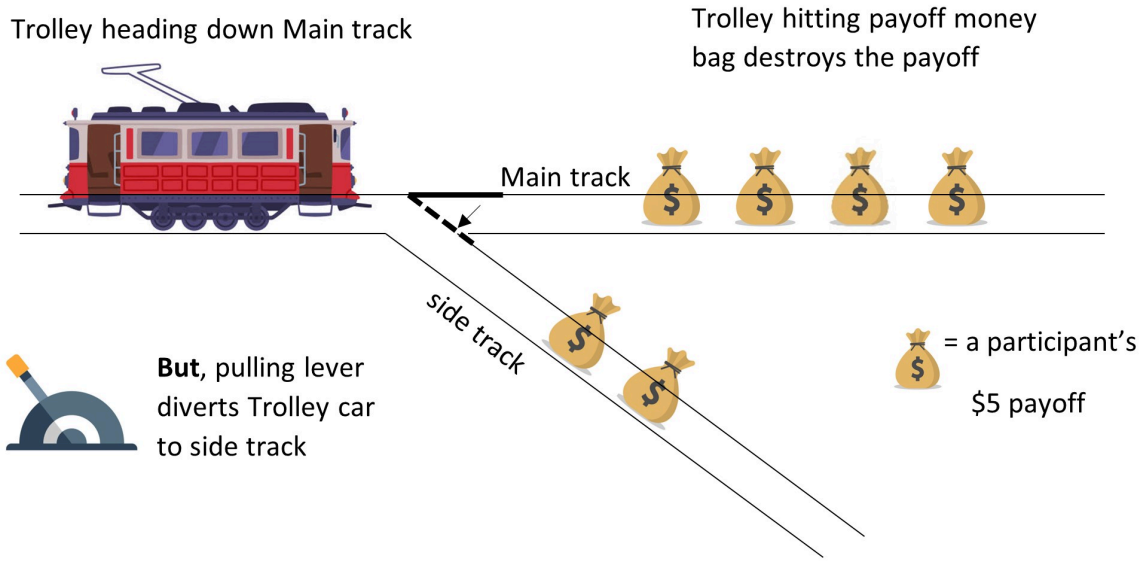
The next set of pages asks about choices you would make in different scenarios. There are no correct or incorrect answers, but we want you to please tell us what your preferred choice is in each scenario (and you have to make a choice in each scenario). **The "payoff" you receive as a result of a randomly chosen trial from this task will determine your Prolific bonus payment for this study.**

In this choice task, **every participant in an experiment group is given a $1 payoff--a "group" for this task is comprised of other Prolific participants.** The payoffs of some of the other (randomly selected) participants in your group are placed on a train track. Each $1 payoff of another participant is represented by a money bag with a "$" inside of the bag in the graphic for each decision scenario (an example is given below)

In each decision scenario, **imagine a trolley or train car that is out of control and heading down the main train track where it will run over and destroy the $1 payoffs of however many individuals' whose payoffs are on that track**. In this case, these participants will lose their $1 payoff for this task.

The decision-maker in this task has the option to take action and save the payoff of those participants by pulling a lever that will divert the train to a side track. **Unfortunately, the payoffs of any participant that happens to be on the side track will then be destroyed.** You will be asked to make choices as the decision-maker in this task.

The decision-maker's $1 payoff is *not* affected at all by the choice he/she makes this task. Here's an example of one such scenario below, where you are the decision maker. In this example, if you choose to do nothing, the payoffs of 4 other participants (i.e., the money bags on the main track) will be destroyed. On the other hand, if you choose to pull the lever, this will save the payoffs on the main track of those 4 participants but the payoffs of two other participants will be destroyed (i.e., the money bags on the side track). You will be asked to make decisions in several such scenarios that will present you with different numbers of others' payoffs on the main and side tracks.



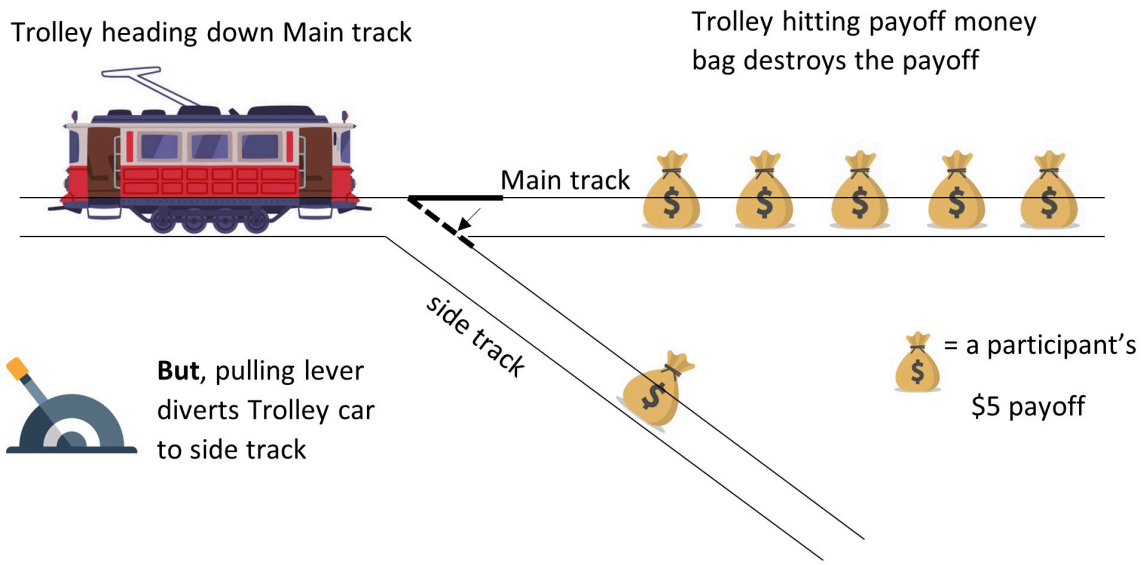
After all decisions have been made by all participants in the study, we will then **randomly group small sets of participants together and select one of the scenarios at random to be used to determine bonus payments for that grouping of Prolific participants.** Specifically, one participant will be randomly selected to be the decision making for your group, and there is an equally likely chance that the decision-maker could be anyone in your grouping. The other participants in your grouping will be randomly assigned to have their bonus payment placed either on the main or side-track for the randomly selected scenario. Then, the choice of the decision-maker in that scenario will be used to determine whether a $1 bonus payment goes to other participants in the group, or whether the $1 bonus payment of that participant is destroyed. Remember, the $1 of the decision-maker is not at risk in these decision scenarios.

Because there is an equally likely chance that you may be the decision-maker in your grouping of participants, please consider each choice carefully because it will dictate real bonus payments for the other participants.

Again, the different scenarios you will face on the following pages will vary the number of individuals' payoffs on the main track versus the side track, and we assume that you know exactly how many are on each track (as shown in the image for that scenario) when you must make your choice as decision-maker. Any one of the scenarios you face may be the one randomly chosen to determine bonus payments, and you may be the decision-maker in that chosen scenario, so **each decision you make may be consequential in determining others' bonus payments** in this task.

Please advance to the next page when you are ready to start making your decisions in these scenarios.

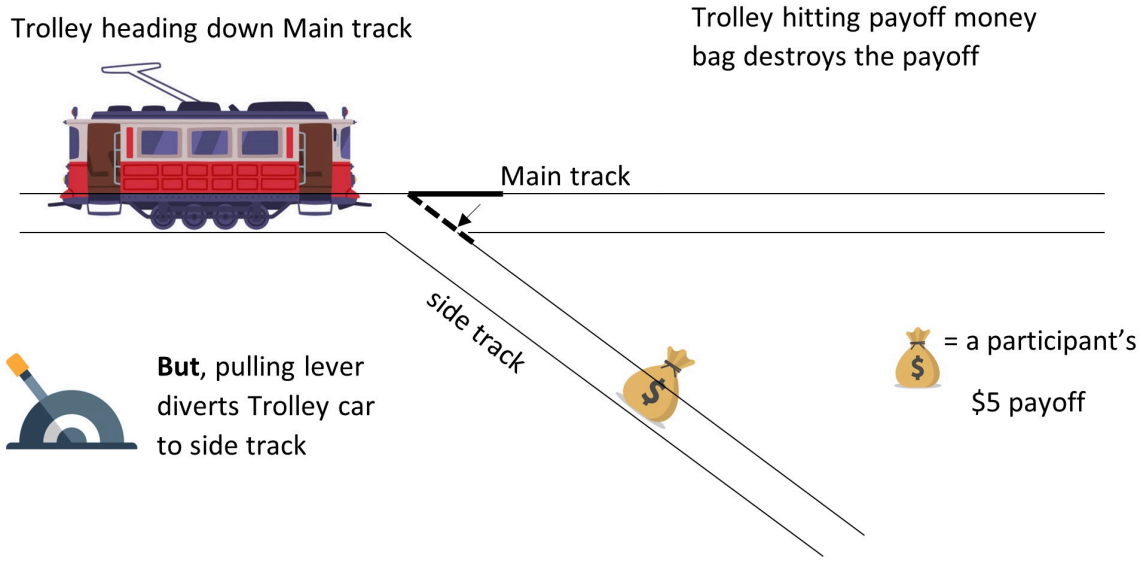
Suppose **the payoffs of 5 participants are on the main train track**. These payoffs will be run over and destroyed by the train *if you do nothing* (i.e., these participants will not receive a payoff from this task). **The payoff of 1 (different) participant is on the side track**, and his/her payoff will be run over and destroyed if you if you pull the lever to divert the train.



What do you choose to do?

- ○ **Do nothing** (train continues on main track)
- ○ **Pull the lever** (train is diverted to the side track)

Suppose **nobody's payoff is on the main train track**. Therefore, nobody's payoffs will be run over and destroyed by the train *if you do nothing*. **The payoff of 1 participant is on the side track**, and his/her payoff will be run over and destroyed if you if you pull the lever to divert the train.

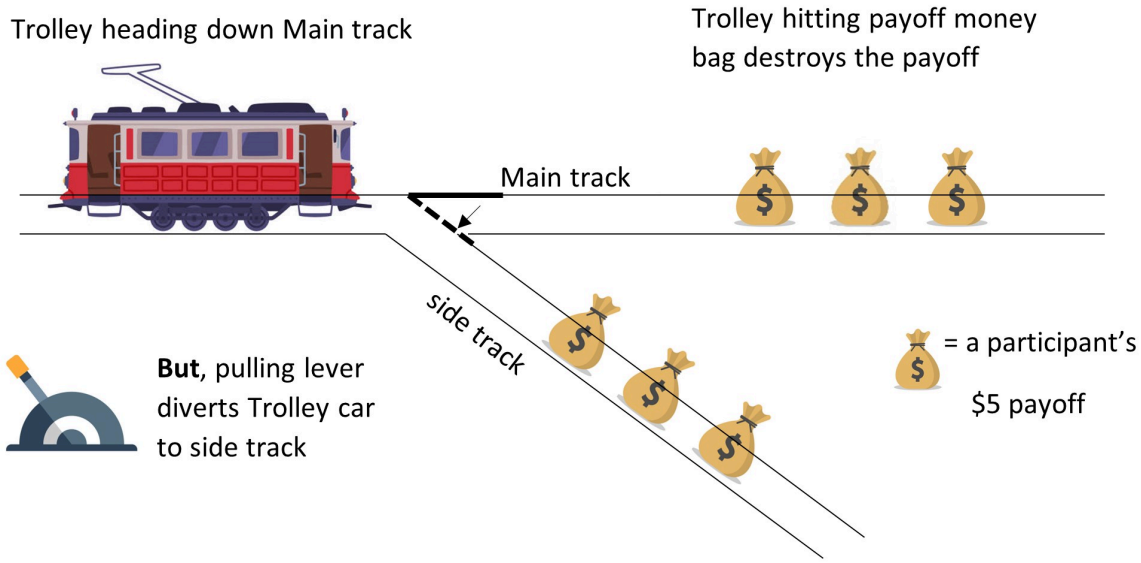


What do you choose to do?

- ○ **Do nothing** (train continues on main track)
- ○ **Pull the lever** (train is diverted to the side track)

Suppose **the payoffs of 3 participants are on the main train track**. These payoffs will be run over and destroyed by the train *if you do nothing* (i.e., these participants will not receive a payoff from this task).

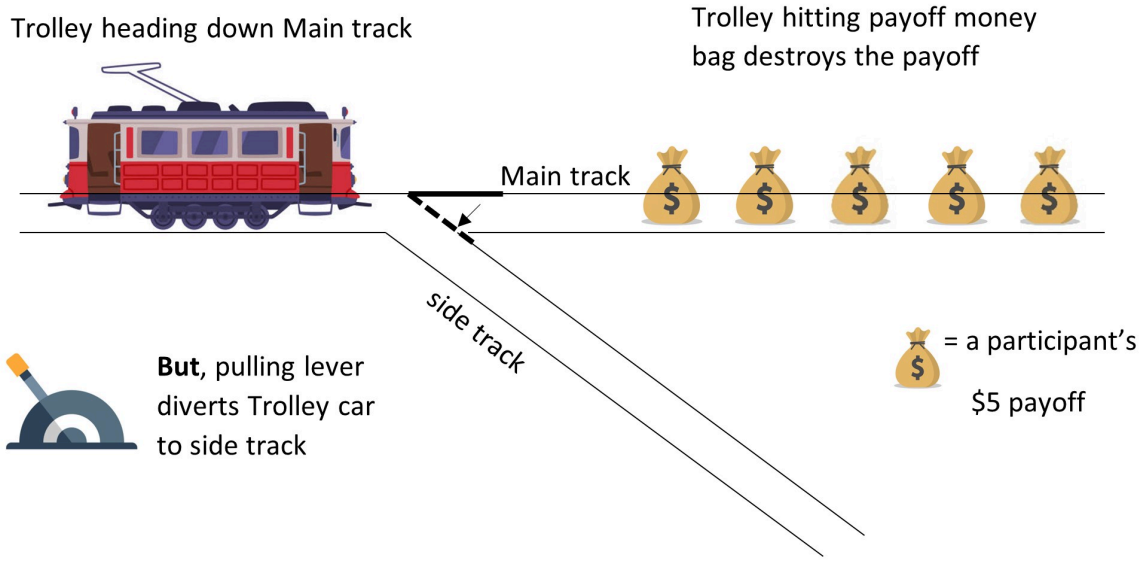
The payoffs of 3 (different) participants are on the side track, and their payoffs will be run over and destroyed if you if you pull the lever to divert the train.



What do you choose to do?

- ○ **Do nothing** (train continues on main track)
- ○ **Pull the lever** (train is diverted to the side track)

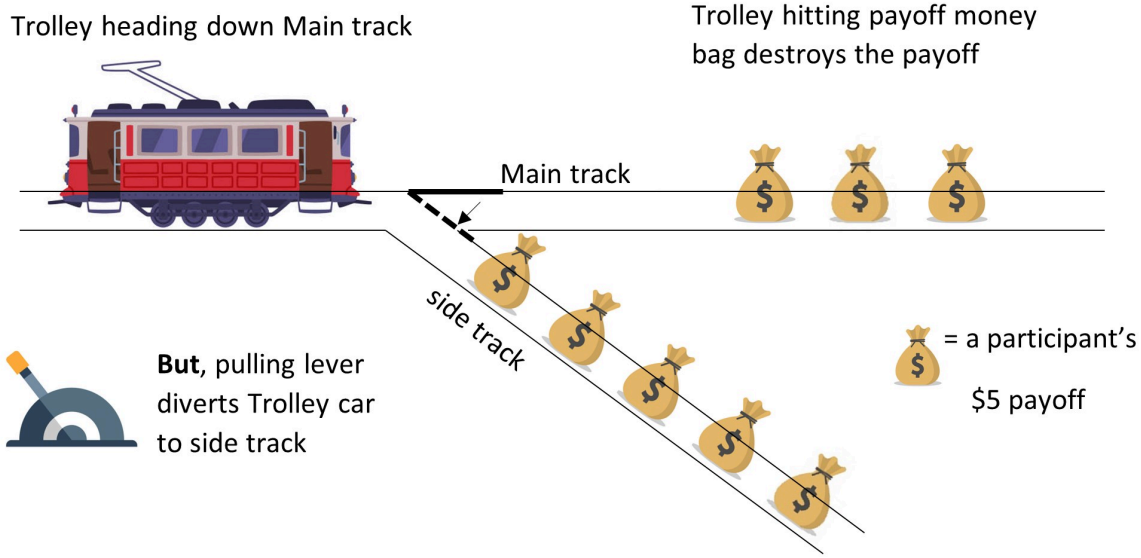
Suppose **the payoffs of 5 participants are on the main train track**. These payoffs will be run over and destroyed by the train *if you do nothing* (i.e., these participants will not receive a payoff from this task). **Nobody's payoff is on the side track**, and so no one's payoff will be run over and destroyed if you if you pull the lever to divert the train.



What do you choose to do?

- ○ **Do nothing** (train continues on main track)
- ○ **Pull the lever** (train is diverted to the side track)

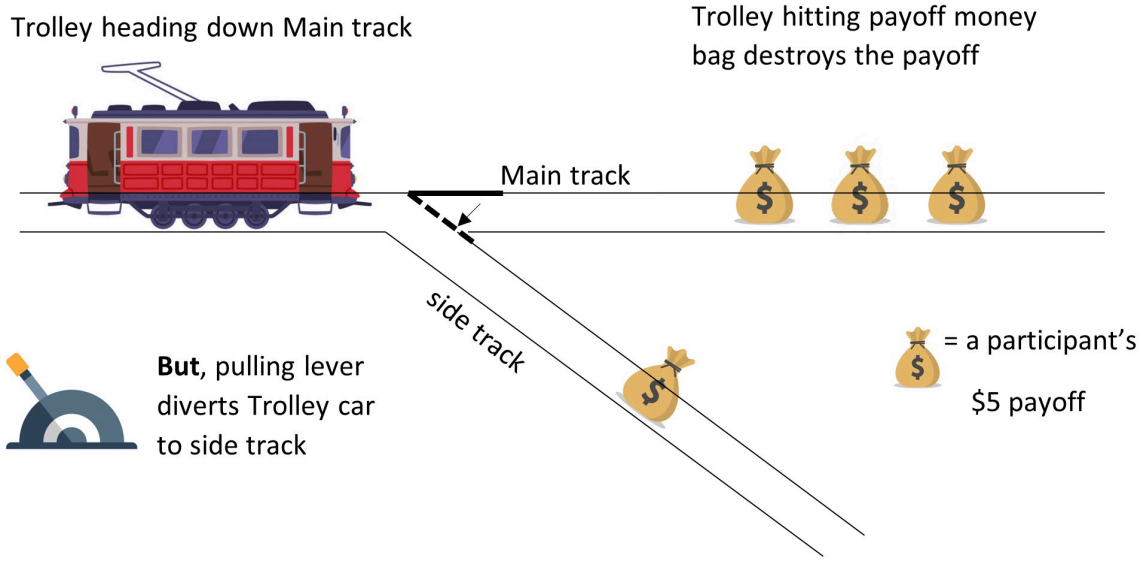
Suppose **the payoffs of 3 participants are on the main train track**. These payoffs will be run over and destroyed by the train *if you do nothing* (i.e., these participants will not receive a payoff from this task). **The payoffs of 5 (different) participants are on the side track**, and their payoffs will be run over and destroyed if you if you pull the lever to divert the train.



What do you choose to do?

- ○ **Do nothing** (train continues on main track)
- ○ **Pull the lever** (train is diverted to the side track)

Suppose **the payoffs of 3 participants are on the main train track**. These payoffs will be run over and destroyed by the train *if you do nothing* (i.e., these participants will not receive a payoff from this task). **The payoff of 1 (different) participant is on the side track**, and his/her payoff will be run over and destroyed if you if you pull the lever to divert the train.

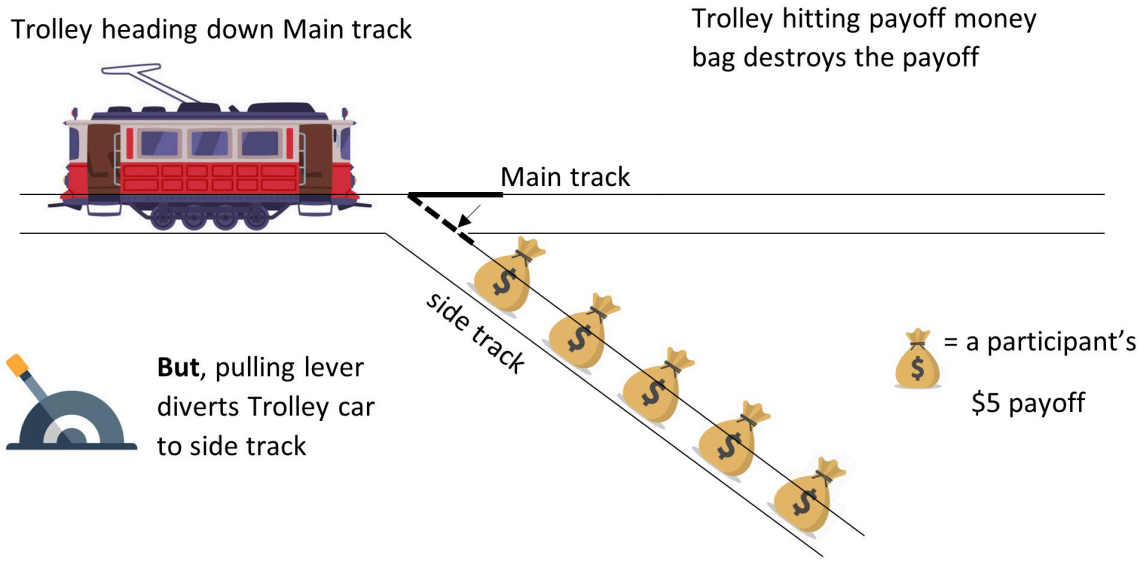


What do you choose to do?

○ **Do nothing** (train continues on main track)

○ **Pull the lever** (train is diverted to the side track)

Suppose **nobody's payoff is on the main train track**. Therefore, nobody's payoffs will be run over and destroyed by the train *if you do nothing*. **The payoffs of 5 participants are on the side track**, and their payoffs will be run over and destroyed if you if you pull the lever to divert the train.

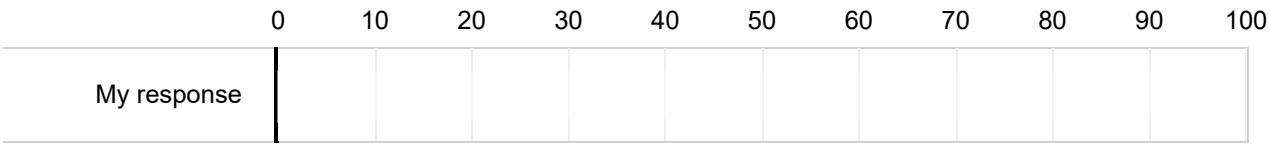


What do you choose to do?

○ **Do nothing** (train continues on main track)

○ **Pull the lever** (train is diverted to the side track)

In order for the results of this survey to be valid, **it is essential that you remain attentive for each of the trials in this study**. So we know that you are still paying attention, please place the slider below on the answer to (33+12)=? Thank you for your continued attention to each of the remaining trials in this study.

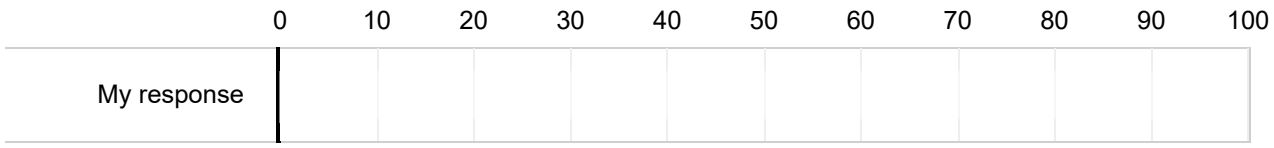


WARNING.....you just failed an attention check question, but we will give you another chance at this.

Please note, ***there may be another attention check question in the survey to ensure that you are reading each question***, and if you fail any other attention checks within this survey, you will be routed to the end of the survey and not compensated.

Here's the question one more time.....

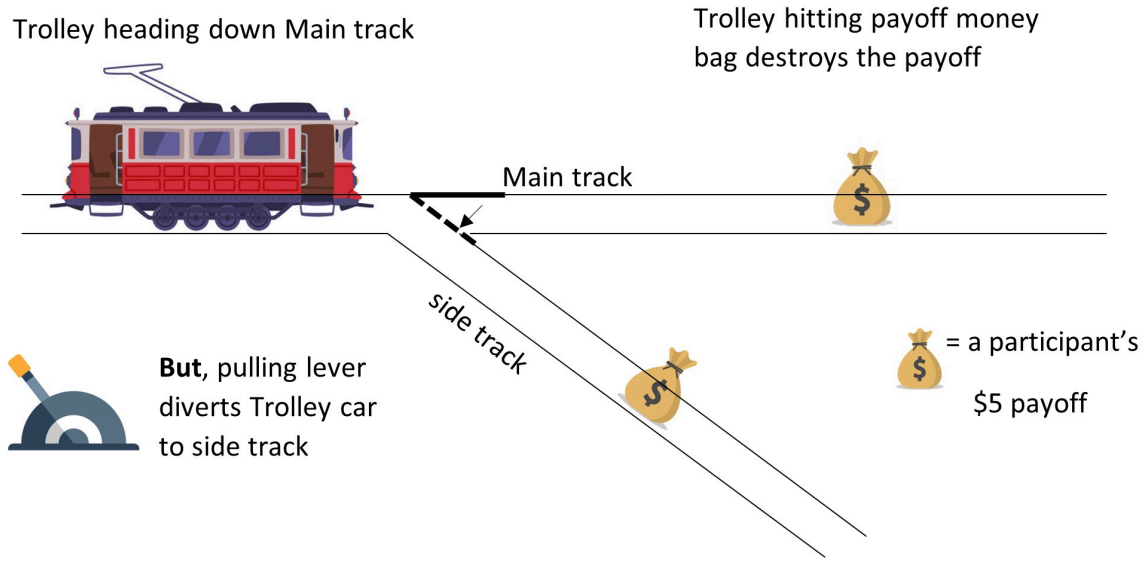
As described earlier, we are interested in factors that influence the decisions you might make. In order for the results of this survey to be valid, **it is essential that you read all the instructions and questions carefully**. So we know that you have read these instructions, please place the slider below on the answer to (33+12)=? Thank you for taking the time to read these instructions.



You have now unfortunately failed multiple attention check questions and so we are ending your survey.
(We hope you understand because we cannot trust that the data in such cases where it seems participants are not reading questions carefully in our survey.)

Please return your submission on Prolific by selection the "Stop without completing' button. If you do not return the submission, then we will have to "reject" payment for your survey, which may damage your Prolific participant standing.

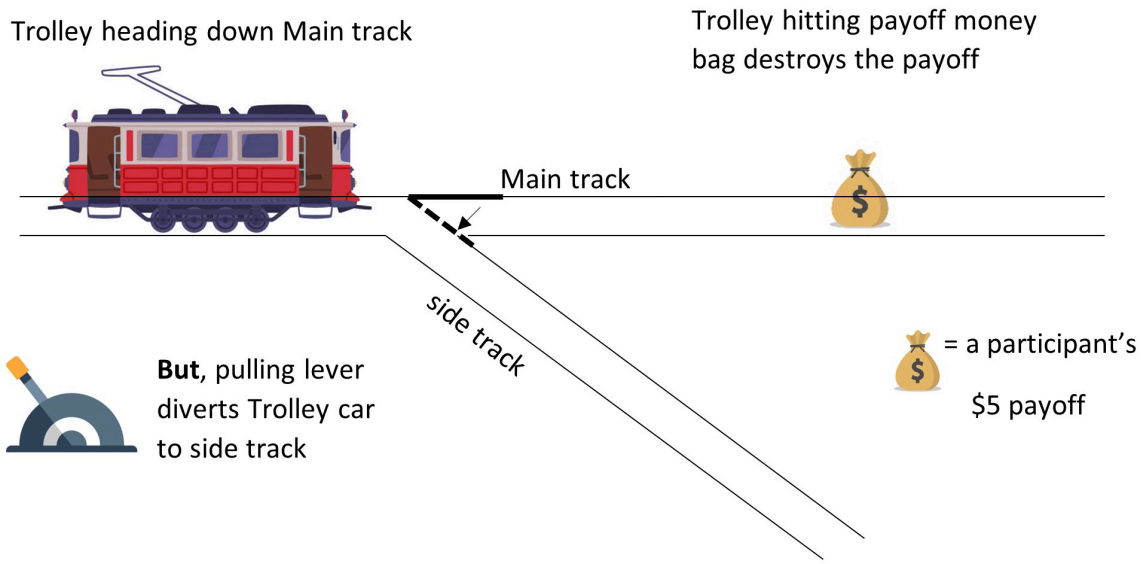
Suppose **the payoff of 1 participant is on the main train track**. His/her payoff will be run over and destroyed by the train *if you do nothing* (i.e., that participant will not receive a payoff from this task). **The payoff of 1 (different) participant is on the side track**, and his/her payoff will be run over and destroyed if you if you pull the lever to divert the train.



What do you choose to do?

- ○ **Do nothing** (train continues on main track)
- ○ **Pull the lever** (train is diverted to the side track)

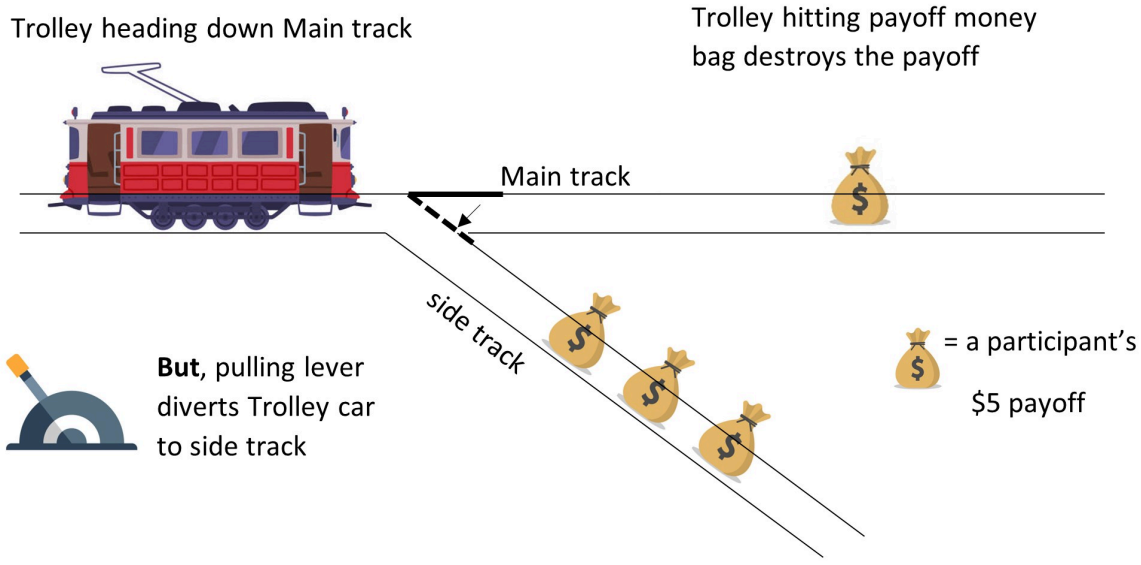
Suppose **the payoff of 1 participant is on the main train track**. His/her payoff will be run over and destroyed by the train *if you do nothing* (i.e., that participant will not receive a payoff from this task). **Nobody's payoff is on the side track**, and so no one's payoff will be run over and destroyed if you if you pull the lever to divert the train.



What do you choose to do?

- ○ **Do nothing** (train continues on main track)
- ○ **Pull the lever** (train is diverted to the side track)

Suppose **the payoffs of 1 participant is on the main train track**. His/her payoff will be run over and destroyed by the train *if you do nothing* (i.e., that participant will not receive a payoff from this task). **The payoffs of 3 (different) participants are on the side track**, and their payoffs will be run over and destroyed if you if you pull the lever to divert the train.

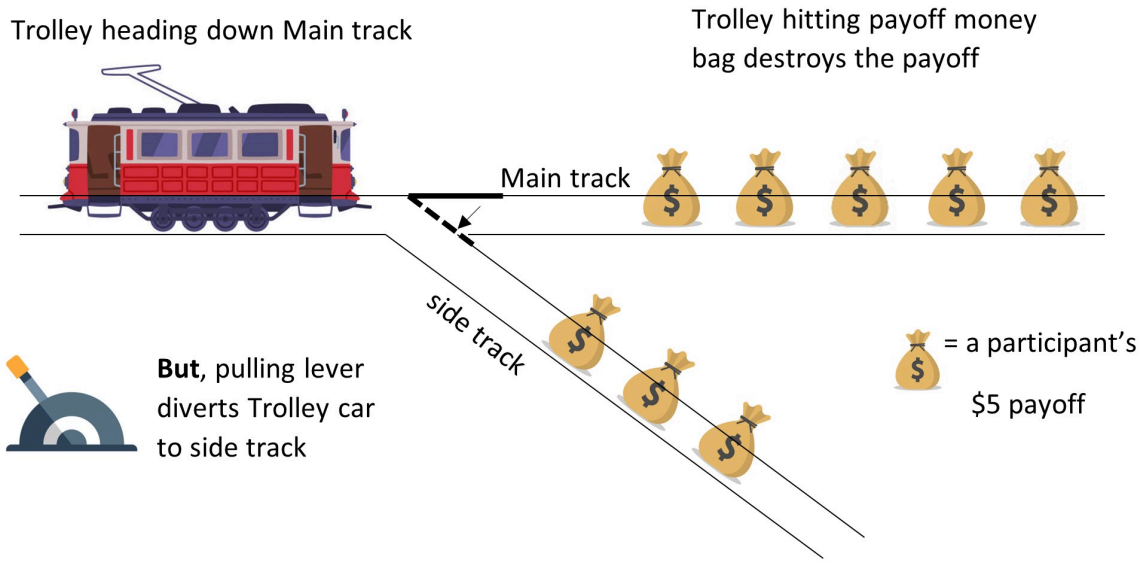


What do you choose to do?

○ **Do nothing** (train continues on main track)

○ **Pull the lever** (train is diverted to the side track)

Suppose **the payoffs of 5 participants are on the main train track**. These payoffs will be run over and destroyed by the train *if you do nothing* (i.e., these participants will not receive a payoff from this task). **The payoffs of 3 different participants are on the side track**, and their payoffs will be run over and destroyed if you if you pull the lever to divert the train.

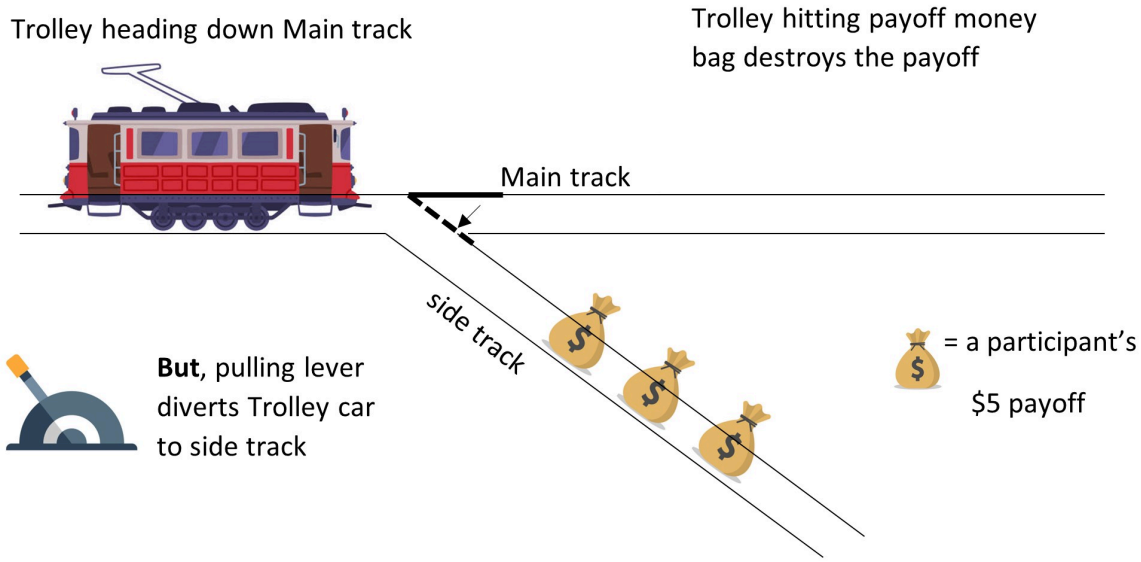


What do you choose to do?

○ **Do nothing** (train continues on main track)

○ **Pull the lever** (train is diverted to the side track)

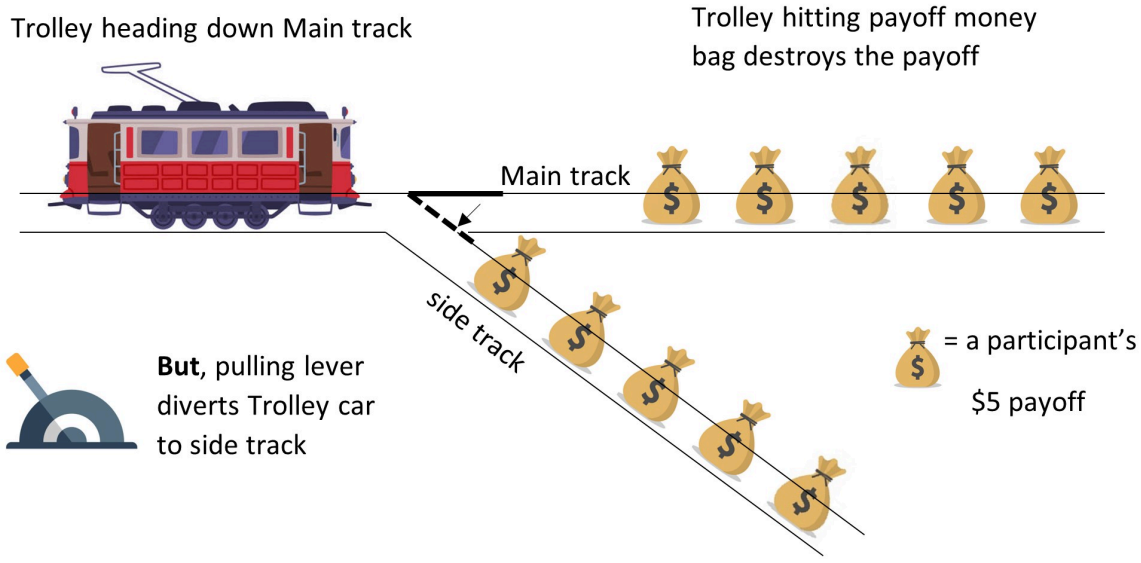
Suppose **nobody's payoff is on the main train track**. Therefore, nobody's payoffs will be run over and destroyed by the train *if you do nothing*. **The payoffs of 3 participants are on the side track**, and their payoffs will be run over and destroyed if you if you pull the lever to divert the train.



What do you choose to do?

- ○ **Do nothing** (train continues on main track)
- ○ **Pull the lever** (train is diverted to the side track)

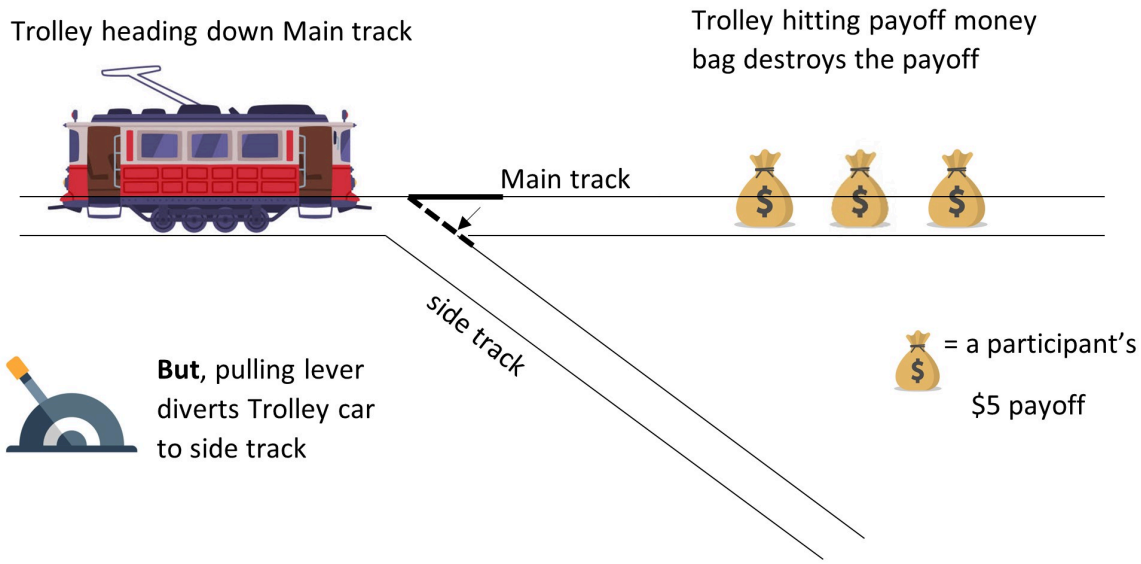
Suppose **the payoffs of 5 participants are on the main train track**. These payoffs will be run over and destroyed by the train *if you do nothing* (i.e., these participants will not receive a payoff from this task). **The payoffs of 5 different participants are on the side track**, and their payoffs will be run over and destroyed if you if you pull the lever to divert the train.



What do you choose to do?

- ○ **Do nothing** (train continues on main track)
- ○ **Pull the lever** (train is diverted to the side track)

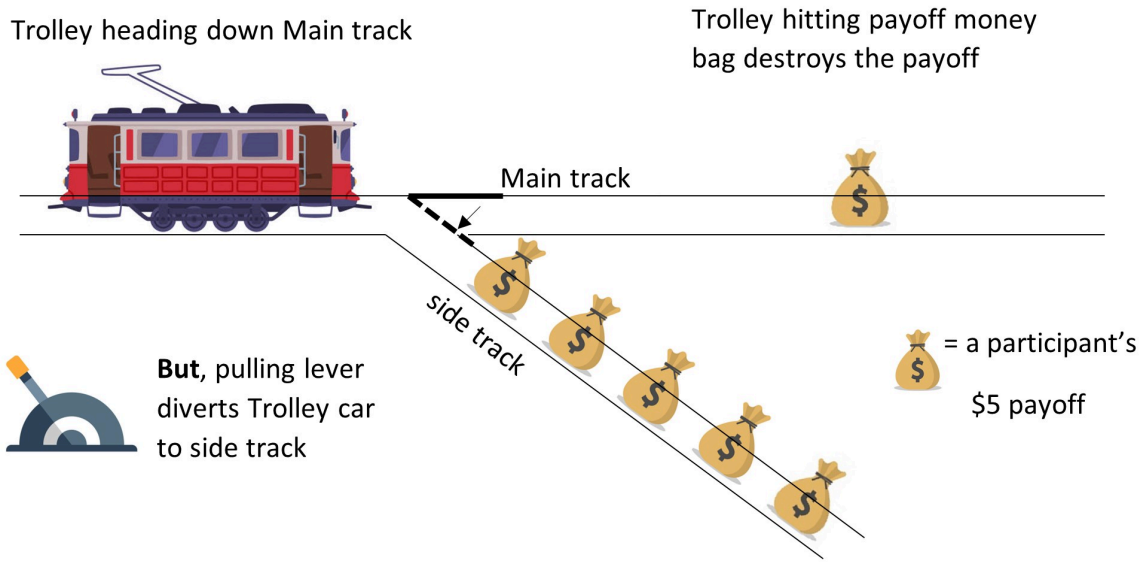
Suppose **the payoffs of 3 participants are on the main train track**. These payoffs will be run over and destroyed by the train *if you do nothing* (i.e., these participants will not receive a payoff from this task). **Nobody's payoff is on the side track**, and so no one's payoff will be run over and destroyed if you if you pull the lever to divert the train.



What do you choose to do?

- ○ **Do nothing** (train continues on main track)
- ○ **Pull the lever** (train is diverted to the side track)

Suppose **the payoffs of 1 participant is on the main train track**. His/her payoff will be run over and destroyed by the train *if you do nothing* (i.e., that participant will not receive a payoff from this task). **The payoffs of 5 (different) participants are on the side track**, and their payoffs will be run over and destroyed if you if you pull the lever to divert the train.



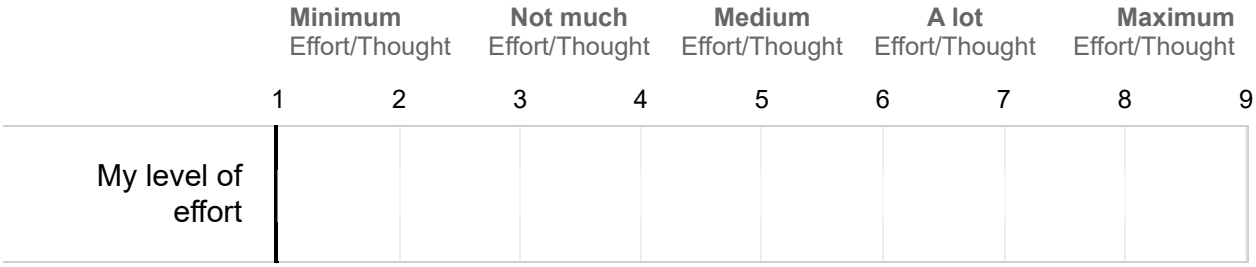
What do you choose to do?

○ **Do nothing** (train continues on main track)

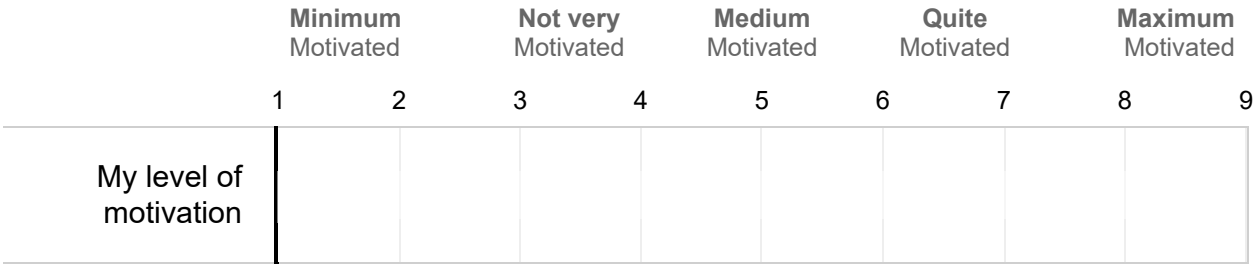
○ **Pull the lever** (train is diverted to the side track)

Motivation and Interest new Mood (Trolley)

On the scale below please indicate **how much effort/thought** you put into the Trolley scenarios task you just completed.



On the scale below please indicate **how motivated you were** in doing the Trolley scenarios task you just completed.



Please rate how strongly you feel each of these emotions ***right now***.

Right now I feel.....

	Zero level of this emotion **(1)**	**(2)**	**(3)**	Mid-Range level of this emotion **(4)**	**(5)**	**(6)**	Maximum level of this emotion **(7)**
Happy	○	○	○	○	○	○	○
Excited	○	○	○	○	○	○	○
Surprised	○	○	○	○	○	○	○
Satisfied	○	○	○	○	○	○	○
Angry	○	○	○	○	○	○	○
Irritated	○	○	○	○	○	○	○
Confused	○	○	○	○	○	○	○
Regret	○	○	○	○	○	○	○

	Zero level of this emotion (1)	(2)	(3)	Mid-Range level of this emotion (4)	(5)	(6)	Maximum level of this emotion (7)
Disgust	○	○	○	○	○	○	○
Alert	○	○	○	○	○	○	○

FINISH

Please click below to record and finalize your responses in this decision task survey.

○ **Record and Finish**