



Contextual Framing of Loss Impacts Harm Avoidance during Risky Spatial Decisions

Kevin Jarbo^{1,2}, David Colaço^{2,3}, and Timothy Verstynen^{1,2}

¹Department of Psychology

²Center for the Neural Basis of Cognition

Carnegie Mellon University

Pittsburgh, PA 15213

³Department of History & Philosophy of Science

University of Pittsburgh

Pittsburgh, PA 15213

Abbreviated Title Contextual framing of loss impacts harm avoidance

Keywords Contextual framing, loss aversion, spatial decision-making, moral decision-making

Corresponding/Senior Author

Timothy Verstynen

Carnegie Mellon University

Department of Psychology

Baker Hall 342E

Pittsburgh, PA 15213

timothyv@cmu.edu

Total Words 7234

Figures 4

Tables 1

Funding This project and manuscript were made possible through the support of a grant from the John Templeton Foundation, Award #48365. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

Conflicts of Interest The authors have no conflicts of interest to report.

Acknowledgements The authors thank Prof. Marlene Berhmann and Roberta Klatzky for their consultation and commentary on previous drafts of this manuscript. Thanks also to Dr. Wesley Buckwalter, Dr. Erik Peterson, Prof. Felipe De Brigard, and Prof. Walter Sinnott-Armstrong for their commentary on previous drafts. We also thank undergraduate research assistant Cristina Bañuelos for facilitating data collection.

Author Biosketches

Kevin Jarbo is currently a President's Postdoctoral Fellow in the Department of Social and Decision Sciences at Carnegie Mellon University. His research focuses on avoidant behavioral responses during risky spatial decisions and the neural circuitry that supports decision-making processes.

David Colaço is currently a postdoctoral fellow in philosophy and neuroscience at Mississippi State University. His research focuses on the philosophy of science and the epistemology of experiment in neuroscience and psychology.

Timothy Verstynen is currently an associate professor in the Department of Psychology and Carnegie Mellon Neuroscience Institute at Carnegie Mellon University. His research focuses on how humans plan actions, adapt actions with experience, and how learning and decision-making arise from brain circuits.

Abstract

While the same decision to act can occur in multiple contexts, how these contexts differentially influence behavior is not well understood. In this paper, we investigate whether contextual framing affects individuals' behavior in spatial decision-making. While previous research suggests that individuals' judgments are sensitive to contextual (and particularly moral) factors of a scenario, no work has addressed whether this effect extends to spatial decisions. To investigate the impact of context on perceptual sensorimotor behavior, we superimposed two moral dilemmas (which we call *help* and *harm*) on a spatial decision-making paradigm (Jarbo, Flemming, & Verstynen, 2017). The basic task required participants select a target area while avoiding an overlapping non-target area. While the visuospatial task was constant, the moral context was changed when participants had to execute either a drone missile strike on enemies in the harm context or deliver ammunition to allies in the help context. Participants more strongly avoided losses in the harm context, reflected by a greater selection bias away from the non-target (i.e., allies) on drone strike trials. These findings suggest that the contextual framing of a subjective perceived loss on a spatial decision can drive avoidant motor execution behavior.

Introduction

Imagine you are challenged to throw a dart and hit the center of a dartboard. Now, imagine that your (mean-spirited) challenger puts a photograph of your child over the bullseye and tells you to aim for them. Suddenly, throwing the dart is imbued with contextual meaning, even if how to perform the challenge did not change: after all, hitting the center remains the goal. Such a case raises the question: when spatial decisions are presented with a narrative that contextualizes them, how does this affect the spatial decisions we make? To answer this question, we introduce a paradigm in which we contextually frame risky spatial decision-making scenarios, to determine whether participants' decisions are affected by the context, though the task is otherwise identical.

Recent research (Jarbo et al., 2017) suggests that, in spatial decision-making tasks where there is high sensory variance, distraction, and both penalty and reward parameters, individuals make decisions that are not well captured by existing accounts of decision-making, such as prospect theory (Tversky & Kahneman, 1978, 1992) or probabilistic decision theoretic models (Trommershäuser, Maloney, & Landy, 2003a, 2008). Specifically, participants biased their selections away from a spatial region in which selections incur a penalty, but they showed no selection bias (i.e., selections were not significantly different from the target center) in no-penalty conditions with identical perceptual signals. Importantly, since the target was experimentally determined to be the optimal selection location across all conditions, participants should have selected the target center in order to maximize expected gain on the task. Given that the task stimuli were the same across conditions, selection bias away from the penalizing non-target indicated that participants may have subjectively judged the penalty, or *loss*, associated with the non-target to be aversive. In particular, the trend of participants' decisions indicates that

the presence of sensory variance and penalty factors on action outcomes influences participants' judgments of visuospatial stimuli, resulting in spatial selection behavior that is biased away from regions of a target that maximize gain (i.e., optimize task performance) in order to minimize loss. It is worth emphasizing that loss is an aspect of the task that is presented as a numerical score in this task.

In this paper, we investigate whether individuals' selection bias, and therefore their subjective aversion to this loss, can be manipulated by imbuing this task with contextual meaning, akin to the dart throwing example above. This investigation is motivated by two trends in research on human reasoning. First, a body of literature suggests that the way in which a decision is presented influences a participant's response. Framing an otherwise equivalent decision as two different kinds of losses has been shown to elicit distinct choice behavior (Kahneman & Tversky, 1979, 1984). Thus, the contextual framing of the decisions impacts whether an individual chooses an option that is more likely to maximize expected gain, suggesting that individuals may have a subjective preference to avoid one kind of loss over another. Second, previous research suggests that the manipulation of moral factors of a case influences participants' judgments in this case. Notably in one experiment (Knobe, 2003a), Knobe investigated whether or not participants judged a CEO's effect on the environment to be intentional when he acted in order to increase his company's profits. Knobe found that participants judged the CEO to be acting intentionally when he harms the environment but not when he helped it. In this experiment, only the word "harm" was changed to "help" in the narrative that the participants received. Despite what might initially seem to be a minor change of wording related to moral consequences, the change of "harm" to "help" resulted in reported

effects of changes in judgment of intentionality, causal responsibility (Knobe 2007), and knowledge (Beebe & Buckwalter 2010). Importantly in these studies, the “harm” or “help” cases were otherwise identical. For both of these trends in research, vignettes were used, but no paradigm involved spatial decision-making.

Thus, our aim is informed by these reports of the impact of introducing contextual narratives in judgment and decision-making scenarios. While these previous task designs are different from our own in non-trivial ways - meaning that visuospatial decisions may not be impacted by the presence and content of these narratives - these tasks nonetheless share several elements with our visuospatial task, including both the involvement of a narrative and the modulation of the content of this narrative. Together, these elements of our task allow us to investigate their influences on the decisions of participants. Further, some decision-making researchers have suggested that visuospatial motor decisions are akin to economic decisions (Wu, Delgado, & Maloney, 2009), the latter of which *are* susceptible to contextual framing. While we remain neutral on the veracity of this claim, it nonetheless provides motivation for us to investigate visuospatial decisions. Because of these similarities, we conjecture that visuospatial decisions may likewise be susceptible to contextual framing, and we suggest that integrating the insights from the framing of judgments and economic decisions may provide greater elucidation of the nature of visuospatial decision-making.

With our study, we aim to determine whether the contextual framing of a visuospatial decision-making task affects the decisions participants make in this task, even though the context does *not* change which decision is most optimal in this task. Thus, when we refer to *contextual framing*, we have in mind the idea that an otherwise simple perceptual decision-making task,

where the underlying sensory and motor decisions do not change, is imbued with meaning by the introduction of a narrative that contextualizes the task, its components, and the participant's action (see Discussion for concerns regarding whether or not this counts as *framing*). For example, a task where participants must click on dots can be contextually framed by a narrative where the dots are taken to represent people and the act of clicking is taken to represent attacking them. These narratives include *moral factors*, requiring the participant to make a decision where the 'reward' and 'penalty' relate to moral outcome, and differ from one another in terms of the *valence* (i.e., "goodness" or "badness") of the moral factors. Overall, then, we aim to investigate how a moral narrative can change participants' decision-making *outcomes*, e.g., the spatial locations that they select, in the same visuospatial task.

In our experimental task, that we call "Drone Strike," we used a wartime scenario to develop two moral dilemmas that provided contextual frames for the same risky spatial decision. Namely, the target represented enemies to be neutralized by a drone strike or allies to whom ammunition needed to be delivered. In penalty conditions, the non-target represented either nearby allies to be avoided by a drone strike ("harm" context) or enemies to be avoided on ammunition deliveries ("help" context). The "harm" context contextualizes loss (i.e., ally casualties) in a morally different way when compared to the "help" context (i.e., ammunition intercepted by enemies). Importantly, the sensory signals are identical between the "help" and "harm" conditions; only the contextualization of the spatial decision changes.

We specifically address the hypothesis that if risky spatial decision-making behavior is impacted by the subjective aversion to potential loss, then selection bias away from the penalizing non-target in the context of harm (i.e., ally casualties) will be significantly greater

than in the help context (i.e., ammunition interception by enemies). In other words, we expect there to be a difference in selection bias between help-context and harm-context versions of our task, where the task's parameters, are equal in terms of riskiness (i.e., there is high variance) and loss (i.e., there is a numerically-represented penalty). What changes, we suggest, is not the metric of riskiness or loss, but rather what meaning they are imbued with by the narratives.

In addition to our main set of hypotheses, we also analyze the effects of context (harm versus help), cost (no-penalty versus penalty), and target variance (low versus high) on other measures of performance, including selection variability, reaction time, movement time, maximum movement velocity, and average movement velocity. Together, these results more fully characterize avoidant selection behavior during risky spatial decisions.

Methods

Participants

All participants were screened for normal or corrected-to-normal vision and right-handedness. We used ColorBrewer (<https://colorbrewer2.org>) to select colorblind-safe stimulus colors, and verbally confirmed with participants during the instructional period that they were able to discriminate between the stimulus colors used in the task. The participant pool consisted of undergraduate and graduate students from Carnegie Mellon University and the University of Pittsburgh. Carnegie Mellon students were notified of the study either via the university's Psychology Research Experiment System or flyers posted on campus. University of Pittsburgh students were recruited via flyers. All participants in the behavioral study reviewed with an experimenter and signed a paper consent form approved by the Institutional Review Boards of

Carnegie Mellon University and the University of Pittsburgh. All behavioral participants were compensated \$10 per hour for a total of \$20 upon completion of the second session.

We recruited a total of 50 healthy adult participants (mean age = 22.6 years, age range = 18 - 44; 33 female, 11 male) who completed two, one-hour behavioral sessions that occurred on consecutive days. One participant's data was excluded from analysis when an error in stimulus presentation was observed during their second session. Three participants did not return for a second session due to scheduling conflicts that did not allow them to complete the study on consecutive days. Data from two participants were excluded from analyses for failure to reach 90% trial completion on either or both behavioral sessions leaving us with a final N = 44. Excluding data from six participants did not change the general pattern of results.

Existing Behavioral Models

The maximum expected gain model predicts the extent to which an individual will bias their selections of a visually presented stimulus away from regions associated with penalty, when this individual attempts to maximize gain by selecting a location associated with high reward in a risky spatial decision-making task (Trommershäuser et al., 2003a). In Equation 1, MEG_x represents the optimal location, i.e., the location with maximum expected gain, within the stimulus, and is the maximum of a linear function that represents the difference between the target (subscript T) and non-target (subscript NT) stimulus distributions. The mean (i.e., centroid) and standard deviation of the target and non-target distributions are respectively represented by μ_T and σ_T , and μ_{NT} and σ_{NT} . The value of α ranges from 0 to 1 and is used to weigh the target

and non-target distributions that partly determined the magnitude of selection bias away from the penalizing non-target.

$$MEG_x = \arg \max_x (\alpha N(x; \mu_T, \sigma_T) - (1 - \alpha)(N(x; \mu_{NT}, \sigma_{NT}))) \quad \text{Eq. 1}$$

The effect of contextual framing on risky spatial decisions can be examined within the framework of the maximum expected gain model by scaling α . Figure 1 illustrates selection bias as a function of α during a risky spatial decision under two contextual frames for loss. We test whether the valence of the moral factors in each contextual frame predicts different decision-making behaviors in each framed task: that loss in a “harm” context is subjectively more aversive than loss in a “help” context, and participants will thus exhibit greater selection bias away from the penalizing non-target (i.e., $\alpha_{Harm} < \alpha_{Help}$) when all other aspects of the decision (e.g., sensory signals, timing) are the same.

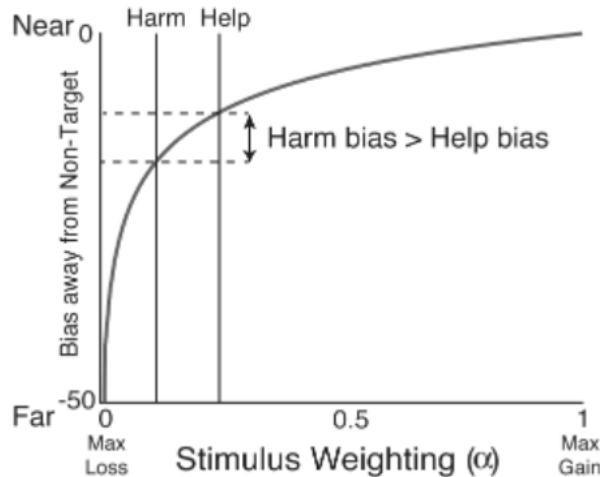


Figure 1 Illustration of selection bias difference prediction based on the maximum expected gain model. Selection bias is plotted as a function (solid black curve) of penalty weighting (α) and a 1:1 target to non-target variance ratio (σ_T/σ_{NT}). More negative values on the y-axis represent selections farther away from the non-target region of the stimulus. Horizontal black dashed lines reflect the hypothesized difference in selection bias in harm and help contexts (solid black vertical lines), where bias is expected to be farther away from a penalizing non-target in subjectively more aversive harm conditions.

Experimental Setup and Design

The behavioral experiment was conducted with Psychophysics Toolbox 3.0.12 (Brainard, 1997; Kleiner et al., 2007) through MATLAB (Release 2015a, The MathWorks, Inc., Natick, MA, United States) on a desktop computer running Ubuntu 16.04. Participants completed the task seated in a dimly lit room in front of a 23" computer monitor with a total screen resolution of 1920 x 1080 pixels and a 60 Hz screen refresh rate.

Using a 2x2x2 (harm vs. help context x no-penalty vs. penalty x low vs. high target variance) within-subject design, each participant completed four runs ("tours") consisting of eight blocks of trials of a single condition ("missions"). We describe the levels of each task condition below in more detail. Participants completed 32 total blocks of 10 trials each for a total of 320 trials in a single experimental session that lasted approximately 50 minutes. Participants completed 640 trials across two sessions. The order of blocks was counterbalanced within runs using a Latin square approach that minimized the correlation between block orders across runs for each participant, as well as across both sessions.

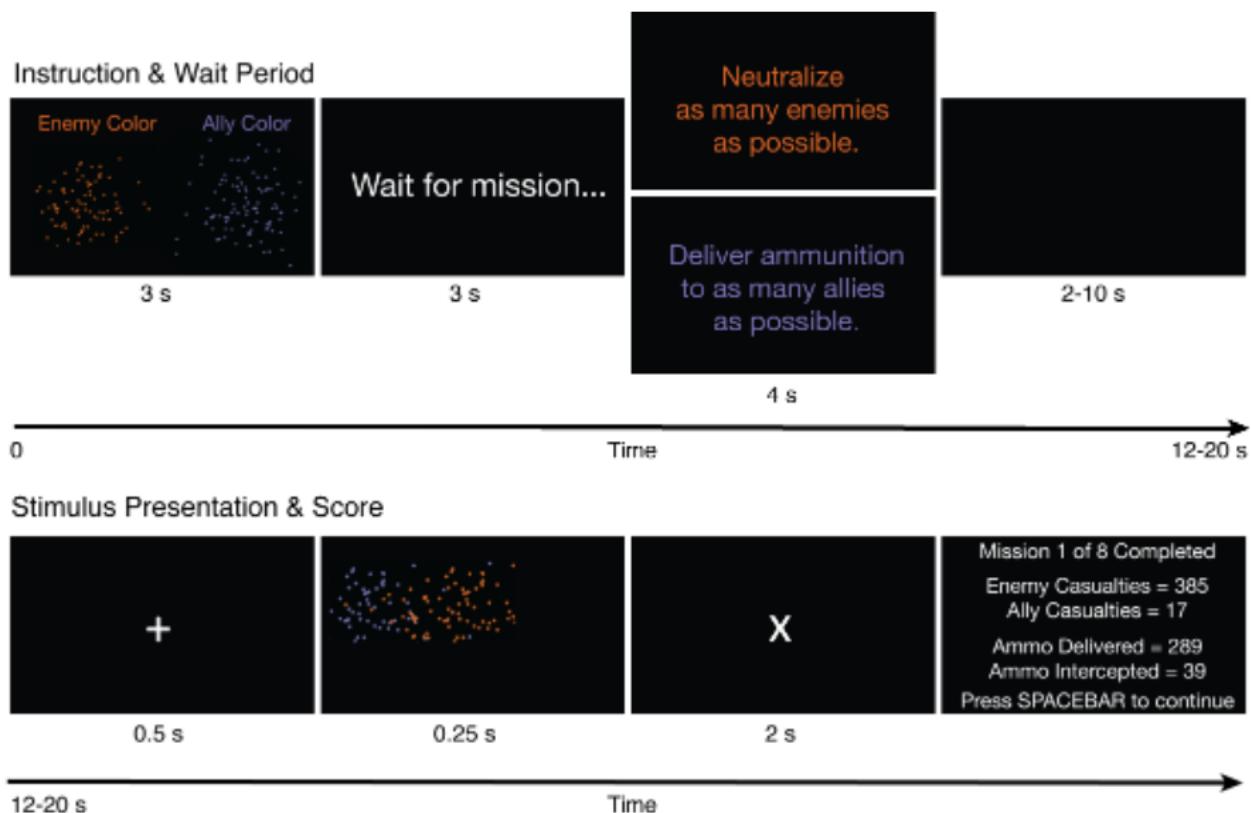


Figure 2 Experimental timeline. Each block, or “mission”, started with an instruction and wait period (top) where participants received a reminder of enemy and ally distribution colors for 3s followed by a 3 s wait period. A condition cue was then presented for 4 s in a font color the same as the target distribution for that block. A blank screen was then presented for 2-10 s (mean ITI = 4 s) prior to each trial. Stimulus presentation (bottom) began with a fixation (+) presented at the center of the screen. Participants had to click and hold the left button within 0.5 s of fixation onset to initiate the trial or else an “ABORT” message appeared indicating a failed trial. On a successfully initiated trial, the target and non-target stimulus distributions appeared onscreen for 0.25 s and then disappeared. Participants then had 2 s to indicate their target selection by dragging the cursor (x) and releasing the mouse button. Each block consisted of 10 trials, and a score report with a running total of enemy and ally casualties as well as ammunition delivered and intercepted was presented until the participant pressed the spacebar indicating that they were ready for the next block of trials. Note: Stimuli and fonts rescaled for clarity.

On each block of trials, participants were tasked with using a computer mouse to select a location within a target stimulus distribution that was visually overlapped by a non-target

stimulus distribution presented simultaneously onscreen (Figure 2). A wartime scenario was used to provide the contextual framing of each spatial selection, wherein participants selected the location of a precision missile strike on enemies or ammunition delivery to allies from a drone on a series of trials within a block. Before each block of trials, participants were presented for 3000 ms with a visual reminder of the colors that corresponded to the enemy and ally stimuli, that were either purple or orange. After a wait screen was presented for 3000 ms, the instruction for the upcoming set of trials was presented for 4000 ms. On “drone strike” missions, participants were instructed to “Neutralize as many enemies as possible”, whereas in the “ammunition delivery” missions participants were instructed to “Deliver ammunition to as many allies as possible”. In both cases, the color of the instruction text matched the target stimulus (i.e., enemies on drone strikes and allies on ammunition deliveries). Following the instruction period, a blank screen was presented before a fixation (+) appeared at the center of the screen indicating the onset of a trial. The onset time for each trial within a block was uniformly sampled from a distribution of intertrial intervals ranging from 2000 ms to 10000 ms (mean ITI = 4000 ms).

To initiate the trial, the participant had to click and hold the left mouse button within 500 ms, otherwise they received an “ABORT!!!” message at the center of the screen indicating a failed trial. For successfully initiated trials, the target and non-target distributions were presented together for 250 ms before disappearing. Both the target and non-target distributions appeared completely on the screen. Each stimulus distribution was presented as a 2D Gaussian distribution of 100 dots that were each three pixels in diameter. The non-target distribution could appear either to the right or left of the target distribution with equal probability across trials. The means of the distributions were separated by fixed horizontal distance of 50 pixels. The mean of the

target distribution was randomly sampled from a distribution of 2D coordinates that had a minimum distance of 350 pixels away from the center of the screen. On no-penalty blocks, the non-target stimulus distribution represented the position of trees, that were always green. On penalty blocks, the target and non-target distributions were the color of enemies and allies, respectively. In the low target variance conditions, the target standard deviation was set to 25 pixels and to 100 pixels in the high target variance condition. The standard deviation of the non-target distribution was fixed at 25 pixels across all trials.

After the stimulus distributions disappeared, the mouse cursor was immediately presented as an “x” at the center of the screen. Participants then had 2000 ms to drag the cursor to a location and then release the mouse button to indicate their selection for each drone strike or ammunition delivery. After a full set of 10 trials in a block, a report screen was presented to indicate progress through the experiment along with a running total of enemy casualties, ally casualties, ammunition delivered, and ammunition intercepted. The report screen was the first instance in which a participant was presented with information about “ally casualties” or “ammunition intercepted,” and no additional narrative description was provided for either of these totals. This report remained on the screen until any key on the keyboard was pressed by the participant to initiate the next run or block. A final score report screen was presented at the end of the session. Selection bias was measured as the distance, in pixels, between a selection and the target mean on a trial (Figure 3). Selections further from the target mean in the direction away from the non-target are represented as negative values. Selections closer to the non-target distribution are represented as positive values. Reaction time was recorded at the first mouse movement detected after stimulus offset. By recording mouse cursor positions, sampled at the

screen refresh rate of 60 Hz across the duration of a trial, and button presses along with RT and MT, we computed the maximum and average velocity of the mouse cursor movements during selections on each trial.

Regardless of context (i.e., drone strike or ammunition delivery), cost (i.e., no-penalty or penalty) or target variance (i.e., low or high) condition, selecting the mean (i.e., center) of the target distribution guaranteed the maximum possible score on a trial. Equations 2-4 were used to calculate scores across trials. First, the Euclidean distance between a selection and the target distribution mean (Equation 2) and the non-target distribution mean (Equation 3) were computed based respectively on the selection location (x_s, y_s) and the means of both the target stimulus (x_T, y_T) and non-target (x_{NT}, y_{NT}) distributions.

$$d_T = \sqrt{\sum((x_s, y_s) - (x_T, y_T))^2}$$

Eq. 2

$$d_{NT} = \sqrt{\sum((x_s, y_s) - (x_{NT}, y_{NT}))^2}$$

Eq. 3

These distances were used in weighted hyperbolic functions with a 1/d falloff to compute the score for each trial. Equation 4 shows the target function weighted by ω and the non-target by 1- ω . In no-penalty blocks, $\omega = 1$, so that only the selection distance from the target contributed to

the score (i.e., no loss, only enemy kills or ammunition delivered), while $\omega = 0.33$ to additionally reflect losses on penalty blocks as ally kills or ammunition intercepted.

$$Score = (\omega d_T - (1 - \omega)d_{NT}) \times 1000$$

Eq. 4

Here the computed scores were multiplied by 1000 and rounded to yield an integer value between 0 and 100 for each trial. The total score for each block of 10 trials was added to a running total across all blocks within each experimental session.

Behavioral Data Analysis

Selection bias away from the non-target is the primary dependent variable of interest in this study. As part of a series of exploratory analyses on other metrics of selection behavior, we also obtained data on the following outcome variables: selection variability, reaction time (RT), movement time (MT), peak (i.e., maximum) mouse cursor velocity (maxV), and average mouse cursor velocity (avgV). We report the results of the additional analyses in the Appendix and, hereafter, focus exclusively on the selection bias findings. The spatial location of a selection, the time between stimulus offset and movement onset, as well as total movement time were recorded for every completed trial across all participants. Since the non-target position relative to the target was only manipulated on the horizontal dimension, only the horizontal selection distance was used in analyses of selection bias and variability. Selection bias was calculated as the difference between the target mean and the selection relative to position of the non-target.

Specifically, selection bias takes more negative values at greater distances away from the non-target mean. Positive values thus indicate selections closer to the non-target mean (Figure 3).

The primary dependent variable of interest, selection bias, was subjected to a 3-way repeated measures ANOVA to observe whether there were any significant 3-way and 2-way interactions or main effects of context, cost, or target variance. Since five additional dependent variables were subject to ANOVA (see Appendix), a Bonferroni correction of α of $0.05/6 = 0.008$ was used as a threshold for statistical significance. For significant results on the 3-way ANOVA, effect sizes were estimated as η_p^2 . Following the omnibus ANOVA, we examined the simple 2-way interactions between levels of context, variance, and cost as well as the corresponding main and simple effects qualified by any significant 2-way interactions. In order to determine the directionality of significant main effects or interactions from these tests, we report the group means and standard errors for each dependent variable across all conditions, and the results of 1-sample and paired sample t-tests with effect sizes computed as Cohen's d.

To further quantify any group-level main effects or interaction of cost (i.e., penalty) and target variance condition on selection bias between contexts (i.e., harm: drone strike vs. help: ammunition delivery), the mean selection bias in the help conditions was subtracted from the mean in the harm condition. We then subtracted those values in the no-penalty conditions from the values in the penalty conditions that matched on target variance to yield a difference score (i.e., $\Delta_{Harm-Help} = Bias_{Harm} - Bias_{Help}$). As such, negative $\Delta_{Harm-Help}$ values reflect a larger bias away from the non-target in harm conditions than help conditions. Conversely, positive $\Delta_{Harm-Help}$ values would indicate that selections were closer to the non-target in harm conditions. This also allowed us to compute a correlation between $\Delta_{Harm-Help}$ values in low and

high target variance conditions to examine whether or not there was a group-level relationship between how much more (or less) participants biased selections away from the non-target under each level of target variance within each cost condition. Based on prior work (Jarbo et al., 2017), we expect an interaction between cost and target variance resulting in the greatest selection bias under conditions of penalty and high target variance. Thus, we should find that greater (i.e., more negative) $\Delta_{Harm-Help}$ values are negatively correlated with increased target variance. We additionally analyze these $\Delta_{Harm-Help}$ values to examine whether or not *individuals* show differences in selection bias specifically in response to each level of context.

In previous work, Trommershäuser and colleagues, used a normative model of movement to capture selection behavior, and showed that bias varied between participants and was based on *unique* internal estimates of their aiming ability, rather than the objective estimates of the target location and value (Trommershäuser et al., 2003a). As such, participants shifted selections away from the non-target based on a *subjective* component of their selection decisions (e.g., estimates of aiming ability) that can be obscured by a group-level analysis of bias that does not account for individual differences in responses to penalty. At the group level, large differences in selection bias between levels of cost and target variance may overshadow potentially small differences in how individual participants respond to penalty in each level of context. Since we hypothesized that avoidant selection bias would result from participants incorporating a greater subjective perceived loss related to context conditions (i.e., ally casualties vs. ammunition intercept by enemies) into their decision-making process, we may not see a group-level correlation where high target variance conditions with penalty result in greater selection bias and $\Delta_{Harm-Help}$ values for all participants. For instance, some participants may perceive the non-target as more

aversive in penalty conditions only in high target variance conditions, while others may perceive the non-target as equally aversive regardless of target variance. If so, then we should observe some subsets of participants with greater $\Delta_{Harm-Help}$ values only under certain levels cost and target variance conditions.

Hence, we performed an individual difference analysis of selection behavior to examine whether or not some participants show different degrees of selection bias under different combinations of cost and target variance, by categorizing $\Delta_{Harm-Help}$ values into four cells, or quadrants (Q-I through Q-IV) (see Figure 4). Moving counterclockwise beginning with the upper right quadrant, participants in Q-I would be categorized as less harm averse, since $\Delta_{Harm-Help}$ would be positive in both low and high target variance conditions. Participants in Q-II and Q-IV are then only harm averse in either the high or low target variance condition, respectively. If a participant falls in Q-III, then they would be harm averse in both variance conditions. Also, if more participants generally show less bias away from the non-target in no-penalty conditions but are harm averse in penalty conditions overall, then we should see the greatest proportion of $\Delta_{Harm-Help}$ values shift from Q-I to Q-III. Based on these categorizations, we first calculated the ratios of $\Delta_{Harm-Help}$ values in each quadrant as a preliminary estimate of this shift magnitude. Shift ratios greater than 1 thus indicate a larger number of participants with $\Delta_{Harm-Help}$ values in a given quadrant in penalty conditions versus no-penalty conditions. We then performed a χ^2 goodness-of-fit test to determine whether or not the observed number of participants in each quadrant deviated significantly from the expected number.

Results

To portray differences in participants' decision-making behavior in relation to changes in contextual framing of our task, we describe the interactions or main effects of target variance (low vs high), cost (no-penalty vs penalty), and context (harm vs help) on selection bias, with a focus on differences between harm and help conditions.

Selection Bias

In investigating the interactions between context, cost, and target variance, we began by analyzing the relation between changes in these parameters and selection bias. A summary of all the omnibus measures can be found in Tables A1 and A2 in the Appendix. Though the 3-way interaction between context, cost, and target variance was not significant, we observed significant simple 2-way cost by target variance interaction, $F(1,43) = 8.32, p = 0.006, \eta_p^2 = 0.162$ and context x cost, $F(1,43) = 20.286, p < 0.001, \eta_p^2 = 0.321$ (Appendix: Table A1). With respect to the significant cost x variance interaction, participants selections were further away from the non-target in penalty vs. no-penalty conditions, $F(1,43) = 41.878, p < 0.001, \eta_p^2 = 0.493$, and in high vs. low target variance conditions, $F(1,43) = 112.100, p < 0.001, \eta_p^2 = 0.723$. For the significant context x cost interaction, we found no significant main effect of context, $F(1,43) = 0.194, p = 0.662, \eta_p^2 = 0.004$. Specifically in line with our central prediction, we found that the main effect of penalty condition on bias was amplified by the higher risk context, with greater selection bias during penalty blocks in the harm condition than in the help condition, paired $t(43) = -2.163, p = 0.018, M_{HarmPenalty-HelpPenalty} = -2.191$, Cohen's $d = 0.326$. Given the significant cost x variance and cost x context interactions, we more closely examine how cost and variance impact selection bias between harm and help contexts in the following section. In general,

selection bias was negative across all conditions, all $t(43)s < -2.944$, all $p < 0.006$, all Cohen's $ds < -0.444$, except in the harm by no-penalty by low target variance condition (see Figure 2 and Appendix: Table A3).

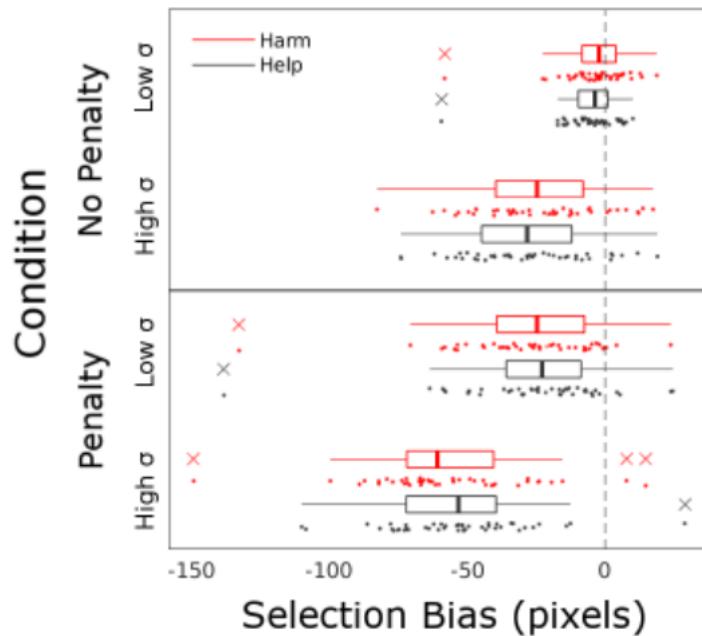


Figure 3 Boxplots of group selection bias and distributions of individual mean values. Boxes (red/top = “harm”, gray/bottom = “help”) represent the group medians (thick line) across all conditions with whiskers corresponding to the 95% confidence intervals of the means. Beneath boxes, each dot represents the mean selection bias for an individual participant across conditions. Selection bias is measured in pixels. Larger negative values (to the left) indicate selections further away from the non-target distribution, while less negative and positive values reflect selections closer to the non-target. The vertical dashed line represents a selection bias of 0. See Table A3 in the Appendix for means and standard errors of all conditions.

Harm versus Help Differences

We computed a value, $\Delta_{Harm-Help}$, to quantify and evaluate differences in selection bias between the harm and help contexts. Keep in mind that in the no-penalty and penalty condition,

negative $\Delta_{Harm-Help}$ values correspond to greater selection bias in harm vs. help conditions.

Using the Pearson correlation coefficient we estimate the similarity of $\Delta_{Harm-Help}$ values between the low and high variance conditions and found no significant linear relationship between selection bias in penalty conditions, $r = -0.081$, $p = 0.600$, or no-penalty conditions, $r = 0.164$, $p = 0.288$. Nor did we find a significant correlation in selection bias after collapsing across cost conditions, $r = -0.190$, $p = 0.216$.

With indication of main effects and interactions, but null correlation findings, we approached the data with an alternative analysis perspective in order to investigate the relation between the contexts more systematically. We categorized selection bias differences by plotting the $\Delta_{Harm-Help}$ values in four distinct quadrants (Figure 4). This was performed so that we could compare the counts of participants whose $\Delta_{Harm-Help}$ values fell into each quadrant under no-penalty and penalty conditions as well as low and high variance conditions, and thus further investigate the relation between selection bias, context, and loss. If participants generally found the harm context more aversive in penalty conditions, then we should observe the highest count of $\Delta_{Harm-Help}$ values shift from Q-I in no-penalty conditions, where participants were selecting closer to the non-target, to Q-III in penalty conditions, where selections were biased away from the non-target regardless of variance. We report these values in Table 1. As expected, in no-penalty conditions, half ($n = 22$) of the $\Delta_{Harm-Help}$ values were in Q-I while only $n = 5$ were observed in Q-III (Figure 4, left panel). In penalty conditions, Q-I had the fewest $\Delta_{Harm-Help}$ values ($n = 5$) while the remaining ($n = 39$) were dispersed nearly evenly across the other three quadrants indicating that most participants were harm averse in penalty conditions in at least one level of target variance (Figure 4, middle panel). This provides additional confirmation of the

significant context x cost interaction, wherein the majority of participants selected closer to the non-target in no-penalty conditions.

We further quantified the extent to which participants as a group were more likely to bias selections away from the non-target in the harm context by computing the ratio of counts in each quadrant between penalty and no-penalty conditions. The shift ratios for Q-II through Q-IV were all greater than 1 with the largest shift ratio of 2.400 for Q-II, indicating that participants showed a greater non-target avoidance driven by penalty in harm contexts, especially under high target variance conditions. In Q-III, the shift ratio of 1.875 shows that participants were nearly twice as likely to significantly bias selections in harm conditions with penalty. To more closely evaluate the shift from Q-I to Q-III (i.e., less selection bias versus more selection bias at both target variance levels), we computed a composite $\Delta_{Harm-Help}$ score collapsed across no-penalty and penalty conditions. This showed that participants with less harm aversive selection bias behavior overall fell within Q-I (x -mean_{Q-I} = 6.798, 95% CI x _{Q-I}: [-1.100, 14.696]; y -mean_{Q-I} = 4.076, 95% CI y _{Q-I}: [-0.084, 8.235]), while more harm averse participants fell within Q-III (x -mean_{Q-III} = -9.928, 95% CI x _{Q-III}: [-14.547, -5.310]; y -mean_{Q-III} = -7.998, 95% CI y _{Q-III}: [-10.664, -5.333]) (Figure 4, right panel). Lastly, the χ^2 goodness-of-fit test confirmed that the observed number of participants in this sample were not equally distributed across quadrants, $\chi^2(3, N = 44) = 11.455$, $p < 0.01$. Together, the results of $\Delta_{Harm-Help}$ values show that framing spatial decisions as potentially harmful can increase aversive selection bias regardless of uncertainty in the estimates of sensory variance.

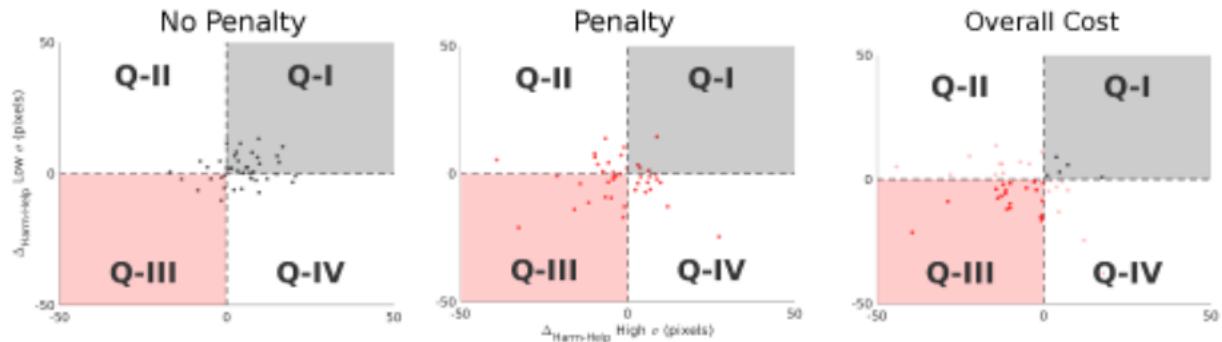


Figure 4 Scatter plots for $\Delta_{Harm-Help}$ shift analysis. Plotting conventions are the same across all panels. Each dot reflects the $\Delta_{Harm-Help}$ value for each individual participant. The x- and y-axis respectively represent $\Delta_{Harm-Help}$ values measured in pixels within high and low variance conditions. Quadrants Q-I (dark gray) contains $\Delta_{Harm-Help}$ values for participants with less harm aversive selection bias, while Q-III (light gray) contains $\Delta_{Harm-Help}$ values for participants with more harm aversive selection bias regardless of variance conditions. A) $\Delta_{Harm-Help}$ values for no-penalty conditions primarily clustered in quadrant I (Q-I). B) In penalty conditions, $\Delta_{Harm-Help}$ values are more broadly distributed throughout Q-II to Q-IV (see Table 1). C) $\Delta_{Harm-Help}$ values were collapsed across variance conditions to generate a composite harm avoidance measure, reflecting overall selection behavior irrespective of target variance or cost condition.

Discussion

Consistent with our hypothesis, our analyses suggest that the moral valence of a spatial decision involving value-based risk and sensory uncertainty increases subjective aversion to losses that are characterized as harmful. To the best of our knowledge, the moral impact of a distinction between a decision outcome best described as “harmful” versus “helpful” on a risky spatial decision has not been previously studied. We found that under equivalent conditions of value-based risk and sensory uncertainty, contextually framing our visuospatial task as what we have characterized as “harmful” (i.e., ally casualties) increases selection bias away from a penalizing non-target to a greater degree than the framing of this task as “helpful” (i.e., ammunition interception). Secondary support for our claims comes from additional analyses showing that

participants had longer reaction and movement times along with slower mouse cursor velocities in both penalty and high target variance conditions (see Appendix). While these additional findings merit further research, they suggest that people made selection decisions more cautiously under the threat of potential loss, by taking significantly longer to initiate and complete selections, and moving more slowly overall throughout the movement. Critically, the primary results of this study show that framing potential loss outcomes as subjectively more aversive uniquely drives avoidant action decision behavior.

By employing our “Drone Strike” narrative, we expand on previous work by examining the effects of contextual framing spatial decision-making. Irrespective of task context, the visuospatial features of the stimuli as well as the scoring function used to compute gains and losses were equivalent across all experimental conditions. The overall goal of the task was to maximize expected gain by either neutralizing the most enemies or delivering the most ammunition to allies as possible. To those ends, the optimal selection on any trial is always the spatial mean of the target distribution, with any bias away from the target a suboptimal selection strategy. Based on prior work, participants were expected to show greater selection bias away from regions of space that induce penalties in feedback so as to avoid losses (i.e., ally casualties or ammunition interceptions) (Gepshtein, Seydell, & Trommershäuser, 2007; Jarbo et al., 2017; Neyedli & Welsh, 2013; Trommershäuser et al., 2003a; Trommershäuser, Maloney, & Landy, 2003b; Wu, Trommershäuser, Maloney, & Landy, 2006). However, if participants were only using spatial estimates of the target and non-target means, as well as the scoring feedback, then there should have been no difference in selection bias between the conditions that involved different contextual frames. Even though both kinds of loss were undesirable, participants biased

selections to avoid the potential collateral losses incurred on a drone strike mission (harm context) to a small (see Results: Selection Bias, Cohen's $d = 0.326$) but significantly greater extent than delivery missions (help context). These findings are further supported by our group-level shift ratio analyses, showing that participants biased selections away from the non-target in harm contexts regardless of variance at nearly twice the rate (mean shift ratio of Q-II, Q-III, and Q-IV = 1.869) of penalty conditions than in no-penalty conditions. Though the context effect measured as $\Delta_{Harm-Help}$ is small, the shift analysis results indicate that participants showed a selection bias to avoid a harmful loss that is not entirely dependent on either the level of penalty or sensory uncertainty during risky spatial decisions. Importantly, our results suggest that the harm and help contexts provided information that was incorporated into the selection decision in a way that made loss in the harm context more aversive than in the help context.

Some degree of the selection bias effects we observed here may be based on noisy estimates of sensory uncertainty. In an experiment by Juni and colleagues (2016), participants had to select on a touchscreen a hidden target whose location could be estimated from a 2D Gaussian distribution of dots where each dot appeared one at a time in random order. A participant could request additional dots to increase their certainty in the sensory estimates of the target location, however, they lost an increasing amount of points with the number of dots requested. This resulted in participants selecting locations from a cluster of dots to minimize point loss once they subjectively determined that there was a sufficiently dense cluster present. The authors found that participants requested more dots than required by an ideal (optimal) observer to accurately estimate the target location, suggesting that individuals failed to maximize expected gain by using a suboptimal decision-making strategy in situations with high sensory

uncertainty (Juni, Gureckis, & Maloney, 2016). In the present study, our participants could have also been targeting areas in the stimulus that they perceived to have the densest cluster of dots in the high target variance conditions. However, since the stimuli were comprised of 2D Gaussian distributions, the densest cluster of target dots was still most likely to be centered on the target mean, in accord with the law of large numbers. Bear in mind that participants were explicitly instructed to select the target center in order to maximize gain regardless of condition. Despite those instructions, some participants could have also adopted a “densest cluster” strategy based on their estimate of the scoring function if they thought that strategy would improve their score. One reason for this might be that participants assumed that, in reality, a drone strike would have a blast radius about the selection location, whereas an ammunition delivery would be received only at the selection location. To better assess strategic task performance, a future version of this study could directly manipulate the location of the densest cluster dots within the target distribution relative to the distribution’s mean to determine whether participants used a “densest cluster” or “spatial center” strategy, as well as ask participants for explanations of their selection decision strategies across different conditions.

We emphasize the fact that our results do not explain *why* our participants behaved differently in our “harm” condition when compared with our “help” condition. Our results do not address how information about the context was processed by the participants. Because of this, we leave open whether the presence of these narratives resulted in participants explicitly formulating strategies to address the task they are assigned or if their presence impacts the sensorimotor processes in some other way. We also leave open whether or not our results indicate *bona fide* framing effects, where a framing effect consists in positive and negative terms in a narrative with

the same information in our task (Levin, Schneider, & Gaeth, 1998). If one argues that the participants value the narratives differently, an argument could be made that the information of the task has changed. However, we reiterate that our use of the notion of ‘framing’ is in terms of a narrative influencing a participant’s decision-making in the same visuospatial task. What constitutes framing in more sophisticated visuospatial tasks like ours, and how this relates to the existing literature on framing effects, is an avenue for further research.

We did not explore the ethical dispositions of participants, which may have mediated (and therefore provide a psychological explanation of) selection bias. Individuals who are more consequentialist in their reasoning, i.e., who judge the rightness or wrongness of an action solely in terms of its consequences (Kagan, 1988), may be more willing to tolerate or even cause loss if it results in a net increase of good. Simplifying somewhat, some participants may accept that the ends justify the means in our study. While what counts as the utilitarian judgment in our scenarios is a topic for further discussion, if the right outcome amounts to neutralizing the most enemies or delivering ammunition to the most allies respectively, these outcomes can be right even if ally casualties or intercepted ammunition are incurred in the process. To explore these open questions more deeply, we can obtain measures of moral and ethical dispositions (e.g., Oxford Utilitarianism Scale) to determine whether or not a person’s degree of impartiality to harm that leads to a greater net benefit correlates with the extent to which they avoid harmful decisions (Kahane et al., 2017). These open questions present the opportunity for interesting research in the future.

Additionally, we aimed to test whether the harm context posed a more aversive loss than the help context, but it is important to acknowledge that things might not be so straightforward.

Judgments about harmful and helpful actions have also been linked to subjective beliefs about intentionality (Knobe, 2003b) and the probabilities of action outcomes (Nakamura, 2018). As such, some questions that are beyond the scope of the present work remain about whether participants thought their choices were causing harm or helping, as well as how likely the harmful or helpful outcome would be if they attempted to maximize expected gain rather than avoid loss. For these reasons, while we refer to our narratives in terms of ‘help’ or ‘harm,’ and our design is informed by help-harm studies, we warn that these labels should not be taken to suggest that these narratives together represent any kind of straightforward help-harm distinction. So, while moral dilemmas provided a strong contextual framing manipulation for this experiment, carefully designed future work is needed to address complex open questions about the rationale participants used for making selection decisions.

Within the broader literature in psychology, the influence of narratives that provide context to judgment tasks has been shown to impact mental processes by changing how information is subjectively perceived, subsequently influencing behavior on cognitive tasks that do not involve sensory or motor processes used in spatial decision-making tasks. For instance, contextual framing like shifts in perspective (e.g., burglar versus homebuyer) impact information encoding and retrieval (Anderson & Pichert, 1978). When individuals were primed with a particular perspective to frame their approach to a memory task, they were able to recall different details about a vignette they read, suggesting that contextual framing can influence what information is remembered and, thus available to be retrieved. Drone Strike does involve both visuospatial working memory, and working memory more generally, in order to encode and represent the briefly presented location of stimuli on each trial and maintain task instructions

across a set of trials. Depending on which task instructions frame the stimuli, participants may show differences in their perceptions of non-target salience and in how accurately they can recall its position, especially when losses must be avoided. Also, in classic work on decision-making and reasoning, reframing logic problems to be more socially relevant to an individual can also increase the likelihood that they arrive at valid conclusions suggesting that context can influence reasoning processes (e.g., variants of the Wason card selection task) (Cosmides & Tooby, 1992; Cox & Griggs, 1982; Wason, 1968; Wason & Shapiro, 1971).

Our findings are consistent with research in related fields. Work in moral psychology and experimental philosophy—including most famously the studies that employ what are called trolley case—have shown that how one contextualizes a situation affects how individuals morally judge the actor as well as the act. Importantly, this research suggests that the effect of contextualization appears to be present even if the outcome—in trolley case, the number of individuals who live or die—is equivalent (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Mikhail, 2007; Sinnott-Armstrong, 2008). Research on the moral reasoning that underlies these judgments may help to illuminate our findings for Drone Strike where, as we hypothesized, participants judged ally casualties as a worse kind of loss than intercepted ammunition, even though the spatial distributions and scoring functions were equivalent.

The present study provides evidence that contextual framing of perceptual decision-making tasks can impact the *outcome* of participants' spatial decisions. This provides the groundwork for investigating whether the contextualizing narrative impacts sensorimotor processes *themselves*, providing a potential mechanism of cognitive penetration that may influence representations of perceptual stimuli during value-based action decisions, or whether

contextualizing impacts some other aspect of the decision-making processes, such as a mechanism by which alternative strategies are explicitly adopted given the participant's subjective value of loss. More generally, our results suggest that spatial decision-making behaviors are sensitive to moral factors, which, if the case, supports the idea that the processes underlying moral decision-making may not be equivalent to non-moral decision-making processes, and the differences between moral and non-moral decision-making should be reflected in our models of them (Bartels, Bauman, Cushman, Pizarro, & Peter McGraw, 2015).

References

- Anderson, R. C., & Pichert, J. W. (1978). Recall of previously unrecallable information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior*, 17(1), 1–12.
[https://doi.org/10.1016/S0022-5371\(78\)90485-1](https://doi.org/10.1016/S0022-5371(78)90485-1)
- Bartels, D. M., Bauman, C. W., Cushman, F. A., Pizarro, D. A., & Peter McGraw, A. (2015). Moral Judgment and Decision Making. In *The Wiley Blackwell Handbook of Judgment and Decision Making* (pp. 478–515). <https://doi.org/10.1002/9781118468333.ch17>
- Battaglia, P. W., & Schrater, P. R. (2007). Humans trade off viewing time and movement duration to improve visuomotor accuracy in a fast reaching task. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 27(26), 6984–6994.
<https://doi.org/10.1523/JNEUROSCI.1309-07.2007>
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/9176952>
- Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* (pp. 163–228). New York: Oxford University Press.
- Cox, J. R., & Griggs, R. A. (1982). The effects of experience on performance in Wason's selection task. *Memory & Cognition*, 10(5), 496–502. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/7176911>
- Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47(6), 381–391. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/13174710>
- Gepshtain, S., Seydell, A., & Trommershäuser, J. (2007). Optimality of human movement under natural variations of visual-motor uncertainty. *Journal of Vision*, 7(5), 13.1–18.

- <https://doi.org/10.1167/7.5.13>
- 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.
<https://doi.org/10.1126/science.1062872>
- Harris, C. M., & Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature*, 394(6695), 780–784. <https://doi.org/10.1038/29528>
- Jarbo, K., Flemming, R., & Verstynen, T. D. (2017). Sensory uncertainty impacts avoidance during spatial decisions. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale*, 1–9. <https://doi.org/10.1007/s00221-017-5145-7>
- Juni, M. Z., Gureckis, T. M., & Maloney, L. T. (2016). Information sampling behavior with explicit sampling costs. *Decisions*, 3(3), 147–168. <https://doi.org/10.1037/dec0000045>
- Kagan, S. (1988). Causation and Responsibility. *American Philosophical Quarterly*, 25(4), 293–302. Retrieved from <http://www.jstor.org/stable/20014252>
- Kahane, G., Everett, J. A. C., Earp, B. D., Caviola, L., Faber, N. S., Crockett, M. J., & Savulescu, J. (2017). Beyond Sacrificial Harm: A Two-Dimensional Model of Utilitarian Psychology. *Psychological Review*. <https://doi.org/10.1037/rev0000093>
- Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica: Journal of the Econometric Society*, 47(2), 263–291. <https://doi.org/10.2307/1914185>
- Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *The American Psychologist*, 39(4), 341–350. <https://doi.org/10.1037/0003-066X.39.4.341>
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., Broussard, C., & Others. (2007). What's new in Psychtoolbox-3. *Perception*, 36(14), 1. Retrieved from [http://www.kyb.mpg.de/publications/attachments/ECVP2007-Kleiner-slides_5490\[0\].pdf](http://www.kyb.mpg.de/publications/attachments/ECVP2007-Kleiner-slides_5490[0].pdf)
- Knobe, J. (2003a). Intentional Action and Side Effects in Ordinary Language. *Analysis*, 63(3), 190–194.

- Knobe, J. (2003b). Intentional action in folk psychology: An experimental investigation. *Philosophical Psychology*, 16(2), 309–324. <https://doi.org/10.1080/09515080307771>
- Levin, I. P., Schneider, S. L., & Gaeth, G. J. (1998). All Frames Are Not Created Equal: A Typology and Critical Analysis of Framing Effects. *Organizational Behavior and Human Decision Processes*, 76(2), 149–188. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9831520>
- Meyer, D. E., Abrams, R. A., Kornblum, S., Wright, C. E., & Smith, J. E. (1988). Optimality in human motor performance: ideal control of rapid aimed movements. *Psychological Review*, 95(3), 340–370. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/3406245>
- Mikhail, J. (2007). Universal moral grammar: theory, evidence and the future. *Trends in Cognitive Sciences*, 11(4), 143–152. <https://doi.org/10.1016/j.tics.2006.12.007>
- Nakamura, K. (2018). Harming is more intentional than helping because it is more probable: the underlying influence of probability on the Knobe effect. *Journal of Cognitive Psychology*, 30(2), 129–137. <https://doi.org/10.1080/20445911.2017.1415345>
- Neyedli, H. F., & Welsh, T. N. (2013). Optimal weighting of costs and probabilities in a risky motor decision-making task requires experience. *Journal of Experimental Psychology. Human Perception and Performance*, 39(3), 638–645. <https://doi.org/10.1037/a0030518>
- Sinnott-Armstrong, W. (2008). Framing Moral Intuitions. In W. Sinnott-Armstrong (Ed.), *Moral Psychology: The Cognitive Science of Morality: Intuition and Diversity* (Vol. 2, pp. 47–76). MIT Press.
- Tassinari, H., Hudson, T. E., & Landy, M. S. (2006). Combining priors and noisy visual cues in a rapid pointing task. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 26(40), 10154–10163. <https://doi.org/10.1523/JNEUROSCI.2779-06.2006>
- Trommershäuser, J., Maloney, L. T., & Landy, M. S. (2003a). Statistical decision theory and the selection of rapid, goal-directed movements. *Journal of the Optical Society of America. A, Optics, Image*



- Science, and Vision*, 20(7), 1419–1433. Retrieved from
<https://www.ncbi.nlm.nih.gov/pubmed/12868646>
- Trommershäuser, J., Maloney, L. T., & Landy, M. S. (2003b). Statistical decision theory and trade-offs in the control of motor response. *Spatial Vision*, 16(3-4), 255–275.
<https://doi.org/10.1163/156856803322467527>
- Trommershäuser, J., Maloney, L. T., & Landy, M. S. (2008). Decision making, movement planning and statistical decision theory. *Trends in Cognitive Sciences*, 12(8), 291–297.
<https://doi.org/10.1016/j.tics.2008.04.010>
- Tversky, A., & Kahneman, D. (1978). 2 - Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. In P. Diamond & M. Rothschild (Eds.), *Uncertainty in Economics* (pp. 17–34). <https://doi.org/10.1016/B978-0-12-214850-7.50008-5>
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297–323. <https://doi.org/10.1007/BF00122574>
- Wason, P. C. (1968). Reasoning about a rule. *The Quarterly Journal of Experimental Psychology*, 20(3), 273–281. <https://doi.org/10.1080/14640746808400161>
- Wason, P. C., & Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. *The Quarterly Journal of Experimental Psychology*, 23(1), 63–71.
<https://doi.org/10.1080/00335557143000068>
- Wu, S.-W., Delgado, M. R., & Maloney, L. T. (2009). Economic decision-making compared with an equivalent motor task. *Proceedings of the National Academy of Sciences*, 106(15), 6088–6093.
- Wu, S.-W., Trommershäuser, J., Maloney, L. T., & Landy, M. S. (2006). Limits to human movement planning in tasks with asymmetric gain landscapes. *Journal of Vision*, 6(1), 53–63.
<https://doi.org/10.1167/6.1.5>

Table 1 Contingency table of observed $\Delta_{Harm-Help}$ value frequencies in each quadrant and shift ratios between and collapsed across penalty and no-penalty conditions.

Quadrant	Penalty	No-Penalty	Shift Ratio	Collapsed (Expected)
I	5	22	0.227	5 (11)
II	12	5	2.400	11 (11)
III	15	8	1.875	20 (11)
IV	12	9	1.333	8 (11)