

Moral judgments recruit domain-general valuation mechanisms to integrate representations of probability and magnitude.

Supplementary Information

¹Amitai Shenhav & ¹Joshua D. Greene

¹Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA, 02138, USA

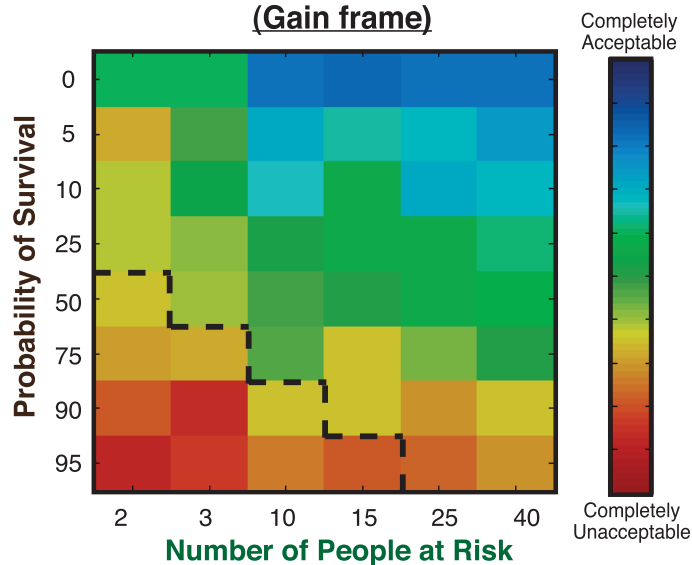
Correspondence should be addressed to A.S. (ashenhav@wjh.harvard.edu)

Supplementary Figures

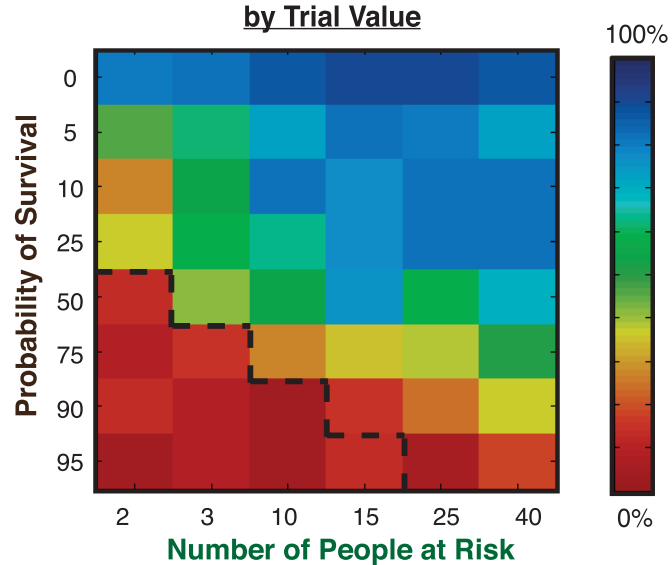
Supplementary Figure 1 (related to Figure 2). **a-c)** Average responses across trial value space for a replication of the fMRI experiment's behavioral task and two variations on that task, each with an independent group of subjects. Distribution and sampling of individual trial values was identical to the fMRI task in all versions (see Figure 2 and Methods). **(a)** Acceptability ratings (1-5 scale) replicating the behavioral results of the fMRI experiment, which employed a Gain frame, describing imperiled individuals as ones whose lives might be *saved*. **(b)** Behavioral results using binary choices instead of acceptability ratings (stay course = 0, change course = 1) and a Gain frame. **(c)** Behavioral results using acceptability ratings and a Loss frame, describing imperiled individuals as ones whose lives might be *lost*. **d)** Average log expected value for changing course at indifference point for responses shown in (a-c) as well as for a binarized version of (a). Average expected value is significantly higher than 1 (the expected value needed to break even; note that $\ln(1) = 0$) in all cases (all $p < 0.05$). Error bars indicate standard error of the mean. **e)** Distribution of subjects in each condition (a-c) classifying their options (one person vs group) as always consisting of two bad outcomes, two good outcomes, or mixed.

a

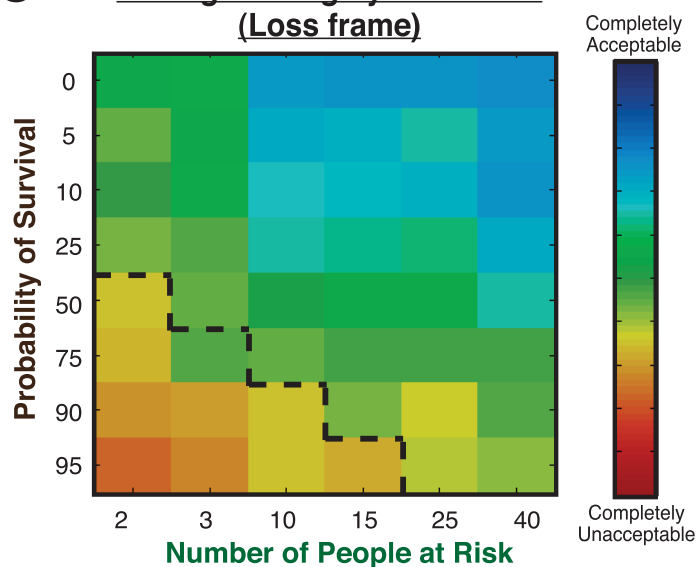
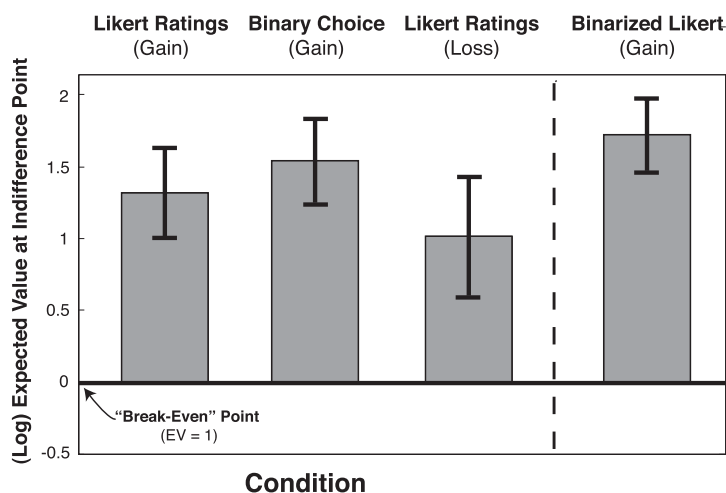
Average Rating by Trial Value (Gain frame)

**b**

Proportion Choosing to Switch by Trial Value

**c**

Average Rating by Trial Value (Loss frame)

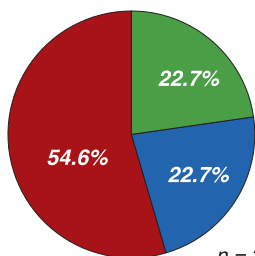
**d****e**

Proportion of subjects viewing their decisions as consistently between:

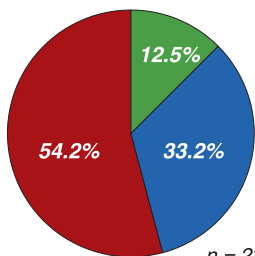
Two BAD outcomes

Two GOOD outcomes

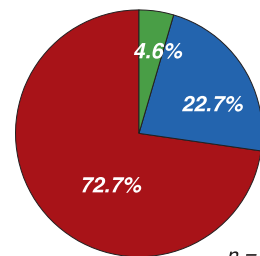
Mixed



Binary Choice
(Gain)



Likert Ratings
(Gain)



Likert Ratings
(Loss)

Supplementary Tables

Supplementary Table 1 (related to Figure 6). Regions exhibiting significant parametric increases in BOLD activity with linearly increasing tendency toward utilitarian judgment at the individual level.

Side	Region(s)	Cluster Extent (voxels)	Peak Z score	Peak MNI Coordinates		
				X	Y	Z
B	SFG (medial wall)	347	4.66	-8	44	50
L	Lateral OFC / frontal pole	468	4.54	-46	36	-16
R	Anterior MTG, temporal pole	265	4.18	60	0	-32
R	Lateral OFC / frontal pole	411	4.14	42	38	-18
L	SPL, anterior IPS	273	3.86	-36	-36	44

Significant clusters met a voxelwise threshold of $p < .005$ and a cluster-wise threshold of $p < .05$. L, left; R, right; B, bilateral. SFG, superior frontal gyrus; OFC, orbitofrontal cortex; MTG, middle temporal gyrus; SPL, superior parietal lobule; IPS, intraparietal sulcus.

Supplementary Table 2 (related to Figure 7). Regions exhibiting significant parametric increases in BOLD activity with increasing certainty of action outcome.

Side	Region(s)	Cluster Extent (voxels)	Peak Z score	Peak MNI Coordinates		
				X	Y	Z
L	MTG, IPL	1839	5.89	-54	-48	0
L	vlPFC/IFG, aIns	2279	5.76	-50	16	16
L	MFG/premotor	383	4.62	-40	0	50
L	SFG (medial wall)	292	4.24	-8	42	44
B	SMA	304	4.20	-2	4	64
B	Precuneus, PCC	659	4.10	-8	-70	34
R	vlPFC, aIns	349	3.93	42	30	-4
R	IPL	268	3.78	44	-56	32

Significant clusters met a voxelwise threshold of $p < .005$ and a cluster-wise threshold of $p < .05$. L, left; R, right; B, bilateral. MTG, middle temporal gyrus; vlPFC, ventrolateral prefrontal cortex; IPL, inferior parietal lobule; IFG, inferior frontal gyrus; aIns, anterior insula; MFG, middle frontal gyrus; SFG, superior frontal gyrus; SMA, supplementary motor area; PCC, posterior cingulate cortex.

Supplemental Experimental Procedures

Task Instructions

Thank you for participating in this study. As noted on your consent form, any data collected from you will be kept strictly confidential.

In this study you will be asked to evaluate a number of moral dilemmas under various conditions. You will be presented with 10 different scenario contexts¹ and will respond to 10 conditions for each one.

For each scenario, you will proceed through four screens. The first three screens will start to describe a situation that you are hypothetically faced with and an action that you could perform in response to that situation.

[screen transition]

When you are done reading each screen, you should press any key to move on to the next one. However, please try your best to get the fullest understanding of the scenario as described thus far before moving on to the next screen.

The scenario description will include all the information you need to make your decision EXCEPT that it will not explicitly state the values for two features of the dilemma: a) a number of people involved in part of the scenario and b) a likelihood that something will happen. You will be evaluating this scenario given a number of variations of these features.

The fourth and final screen will provide you with the prompt that you will be answering for each of these variations - namely, whether or not it is morally acceptable for you to perform the action in question.²

[screen transition]

After you have understood the scenario context and the question you will be answering, you will press any button to move on to the individual trials that will fill these gaps for you in the scenario. Before you do so, please try your best to hold in mind what the action is that you will be evaluating across trials, as you will not be reminded of this after this screen.

First you will see a "+" in the middle of the screen. Any time that this is up, all you need to do is fixate on the "+" and prepare to respond to the next trial. Next you will see text appear indicating the missing information. You should then evaluate the action in question in the context of these values given, and make a judgment of its moral acceptability.

¹ In addition to the scenarios used for this experiment, subjects responded to a set of five additional scenarios (each as additional functional runs) for a separate experiment. These runs had a similar trial structure but the content of those scenarios and the trial value space from which the trials were drawn was different than those described in the current experiment. These additional runs were randomly ordered and interspersed with the one described here.

² In the binary choice follow-up experiment, this line read "The fourth and final screen will provide you with the prompt that you will be answering for each of these variations - namely, which of the actions in question you would perform in this scenario. It is important that you pay special attention to this as you will not be reminded later what specific actions you are choosing between."

You will rate each trial on a 1-5 scale, with 1 indicating that the action would be "Completely Unacceptable" and 5 indicating that it would be "Completely Acceptable."³

[screen transition]

After you answer you will again see a "+" in the middle of the screen, followed by the next trial. This will occur for 10 different variations of each scenario. It is important that you try your best to judge each trial in isolation, and avoid consideration of past responses for the current scenario or past scenarios.

You will only have 10 seconds to respond to each variation. If the "+" appears before you have responded, that means you are out of time. If this happens, do not attempt to respond. Simply look at the "+" and wait for the next trial.

Once you have pressed a button there is no way to go back to the previous screen. If you press the wrong button or if you press a button too soon, don't worry.

[screen transition]

Moral judgments can be difficult to make, and we understand that people sometimes change their minds about moral questions or feel conflicted about the answers they've given. Don't think of your answers as "written in stone." All we want from you is a thoughtful first response.

While we want your answers to be thoughtful, you may find that in some cases the right answer seems immediately obvious. If that happens, it's okay to answer quickly. There are no trick questions, and in every case we have done our best to make the relevant information as clear as possible.

Note, however, that no two scenarios are the same, although many are similar to each other. To answer a question properly you will have to read it carefully because it will always be different in some way from the questions you have already answered.

[screen transition]

In some cases, you might feel that the situation we've described is not realistic. For example, it might say that if you do X, then Y will happen, and you might think that this is not realistic, that Y might not necessarily happen if you do X. If you find yourself having these sorts of doubts, you should "suspend disbelief" and assume that the situation really is the way it's described, even if it doesn't seem realistic to you.

Likewise, you may feel that you need more information than is provided about the situation before you can give your answer. If this happens, you should make your best guess about what you think the situation is like without making any unnecessary assumptions. For example, if it doesn't say that the other person in the situation is related to you, then you should assume that you and the other person are unrelated.

[screen transition]

³ In the binary choice experiment, these two lines read "You should then evaluate the actions in question in the context of these values given, and make your choice. You will press the 1 key to continue with your current action (STAY) or the 2 key to choose the alternate action (SWITCH)."

When you are done reading this screen you will begin two practice scenarios. Once again, you should use any key to advance through the first four screens. Then you will encounter a number of trials which you will rate on a 1-5 scale.

While you will see 10 such trials for each scenario in the actual experiment, you will only respond to 4 variations of each practice scenario.

Furthermore, when the actual experiment begins there will be an additional 10 seconds of fixation before the scenario text comes up, while the scanner warms up. Please begin reading as usual when the text comes up.

If you have any questions at this time, please wait until the current scan completes and then squeeze the ball to let the experimenter know. Otherwise, you may press a key to move onto the practice.

Dilemma text

Below is the text for the moral dilemmas used in the present study (including the wording used to describe the variable Magnitude and Probability parameters). See Fig 1 and Methods in the main text for details concerning presentation. Dilemmas have been adapted from a variety of sources (Boorse and Sorensen, 1988; Cushman et al., 2006; Foot, 1978; Greene et al., 2001; Royzman and Baron, 2002; Thomson, 1985; Waldmann and Dieterich, 2007).

Practice dilemmas

Skydivers

You are the head of a safety crew that has been alerted about a group of skydivers that is wearing defective parachutes and is currently in a freefall. You've arrived in time to see that a group of divers is holding onto one another and will fall to their death a few feet to your right unless your crew's trampoline is moved beneath them.

Your crew sees one diver apart from the rest who is much closer to reaching the ground, and instinctively moves a few feet to the left to catch him. You know that there won't be enough time to reach the group if you wait to catch the one. If nothing is done, the safety crew will catch the diver to the left, avoiding his death, but the separate group of divers will land on the ground, causing their deaths. You can re-direct your crew to the right, preventing the death of the group, but the one diver won't be caught.

You also know that there is a chance that one of the parachutes of the group will still open in time to prevent all of their deaths, but there is no chance that this will happen for the one given how close he is to the ground.

Consider each of the following scenarios and, for each one, determine how morally acceptable you think it would be to direct your crew to the right, re-positioning the trampoline underneath the group.

Divers on right: [variable Magnitude value appears here]

Probability of parachute opening: [variable Probability value appears here]

Loss Frame: *Probability of continuing freefall: [variable 100-Probability value appears here]*

Scuba Divers (only included in fMRI task)

You are working on a ship when you see a large group of sharks in the distance heading towards a team of scuba divers. You know that when the sharks reach this group of people they will encircle them and kill them all.

You also know that there is one scuba diver separate from this group who is closer to you and currently not in the path of the sharks. If you shoot this one diver with a harpoon, he will certainly die and the sharks will sense his presence and switch trajectory away from the team of divers. If you do this, all of the divers will be spared.

You also know that the divers are carrying an electromagnetic (EM) pulse that will deter the sharks, but given the team's current depth there is only a small chance that the pulse will function. In order to ensure that the group will be spared, you must shoot the one diver now.

Consider each of the following scenarios and, for each one, determine how morally acceptable it would be for you to shoot the one diver.

Scuba divers in team: [variable Magnitude value appears here]

Probability EM pulse will function: [variable Probability value appears here]

Main experiment dilemmas

Gas Vent

There has been an accident in the basement of an office building. A pipe has broken and is releasing a lethal gas into the building's ventilation system. If nothing is done, the gas will travel upward to Room A and cause the death of all workers there.

You are the building manager. You know that the only way to avoid the deaths of the workers in Room A is to block the main vent immediately by hitting a button that will close one of the vent doors. You also know that if you close the vent door the gas will be diverted into a different room, Room B, and cause the death of the one worker there.

You are also aware that there is a chance that the workers in Room A will escape before the gas reaches them. This would be impossible for the one in Room B.

Consider each of the following scenarios and, for each one, determine how morally acceptable you think it would be to hit the button that closes the vent door and diverts the lethal gas to Room B.

Workers in Room A: [variable Magnitude value appears here]

Probability of escape: [variable Probability value appears here]

Loss Frame: *Probability of remaining trapped: [variable 100-Probability value appears here]*

Bridge

You are operating a crane in front of a tall bridge when an earthquake hits. As a result of the earthquake, a gap has formed in the bridge, and a bus carrying a number of people is now partially hanging off of this gap in the bridge. If nothing is done, the bus will soon plummet and all of these people will be killed.

You know that you can move your crane in a way to push the bus back onto the bridge, avoiding the deaths of all the people on board. However, you are also aware that in its current position your crane is supporting one small part of the bridge on which there is currently a single car. If you move the crane, this part will certainly collapse and the one person in that car will be killed.

You also know that there is a chance the bus will be able to drive back onto the bridge, preventing the death of everyone on board. There is no chance that the car will be able to avoid falling if you move your crane.

Consider each of the following scenarios and, for each one, determine how morally acceptable you think it would be for you to move your crane to push the bus, causing part of the bridge with a single car to collapse.

People on bus: [variable Magnitude value appears here]

Probability bus will drive to safety: [variable Probability value appears here]

Loss Frame: *Probability bus will fall off bridge: [variable 100-Probability value appears here]*

Rescue Boat

You are driving a rescue boat in the ocean, heading east towards one drowning man. You receive a distress signal informing you that a small boat has capsized in the opposite direction, and all the people aboard are now drowning.

You know that if you immediately change course and go full speed, bearing west, you will reach these people in time to save them. However, if you do this, the one man to the east will certainly die. If you do nothing and hold your course, the one man will be saved, but you will not reach the people to the west in time to save them.

You also know that the only other rescue boat in the area is much further to the west, so would be unable to reach the one drowning man. But there is a chance the rescue boat will reach the group drowning to the west.

Consider each of the following scenarios and, for each one, determine how morally acceptable you think it would be to change your course to head toward the group to the west.

Individuals drowning to the west: [variable Magnitude value appears here]

Probability of alternate rescue boat reaching them: [variable Probability value appears here]

Loss Frame: *Probability of alternate rescue boat failing to reach them: [variable 100-Probability value appears here]*

Boxcar

You are operating the switch at a railroad station when you see an empty, out of control boxcar coming down the main track. It is moving so fast that anyone it hits will die immediately. The boxcar is headed towards a tunneled section in which a group of repairmen are working.

You can flip the switch, redirecting the boxcar to a sidetrack on which there is one repairman working. If you do nothing, the boxcar will continue toward the repairmen in the tunnel on the main track and kill them all. If you hit the switch, the repairmen on the main track will be spared but the one repairman on the sidetrack will be hit by the boxcar and die.

You know that there is a chance an alarm on the main track will be triggered in time to alert the repairmen to evacuate before the boxcar arrives. There is no such alarm on the sidetrack, and therefore no chance the one workman would evacuate in time.

Consider each of the following scenarios and, for each one, determine how morally acceptable you think it would be to hit the switch, redirecting the boxcar onto the sidetrack.

Repairmen on main track: [variable Magnitude value appears here]

Probability of evacuation: [variable Probability value appears here]

Loss Frame: *Probability of remaining in tunnel: [variable 100-Probability value appears here]*

Cafe Grenade

You are working in the kitchen of a café, and see a terrorist throw a grenade next to the main dining room, in which a number of customers are eating. If nothing is done the grenade will explode and the walls of the main dining room will collapse and kill these customers.

There is only one other location the grenade could be thrown before it explodes, and that is the patio outside of the main dining room. There is one customer sitting on the patio, and he would certainly be killed if the grenade is thrown there.

You have time to run out and throw the grenade to the patio, without risking any harm to yourself. If you do this, the customers in the dining room will live but the one on the patio will die. You also know that there is a chance that the dining room walls will withstand the blast of the grenade explosion, and the customers would be spared.

Consider each of the following scenarios and, for each one, determine how morally acceptable you think it would be to throw the grenade onto the patio.

Customers in main dining room: [variable Magnitude value appears here]

Probability customers will be spared: [variable Probability value appears here]

Loss Frame: *Probability walls will collapse on customers: [variable 100-Probability value appears here]*

Post-Task Survey (only used in behavioral experiments subsequent to fMRI task; only final question was included in analysis, wherein the scale was split into 3 even parts such that: 1-3 was categorized as “consistently two bad outcomes”; 4-6 as “mixed”; and 7-9 as “consistently two good outcomes”))

Please circle a number to indicate how much you agree with each of the following statements about what motivated your earlier judgments:

1. My judgments were driven overall by a positive feeling towards saving the one person.

Strongly Disagree

1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---

 Strongly Agree

2. My judgments were driven overall by a positive feeling towards saving the group.

Strongly Disagree

1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---

 Strongly Agree

3. My judgments were driven overall by a negative feeling towards letting the one person die.

Strongly Disagree

1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---

 Strongly Agree

4. My judgments were driven overall by a negative feeling towards letting the members of the group die.

Strongly Disagree

1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---

 Strongly Agree

When judging the scenarios you read earlier, people can have two different kinds of reactions to their options. Some people feel as though the choice they are endorsing is primarily one of **two bad outcomes** (the “lesser of two evils”) because of the deaths that may occur either way. Others feel like they are primarily choosing between **two good outcomes** because either choice offers the opportunity to save lives.

Which of these better describes how you felt about your choices?

Always felt like two BAD outcomes

1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---

 Always felt like two GOOD outcomes

Supplemental References

Boorse, C., and Sorensen, R. (1988). Ducking harm. *Journal of Philosophy* 85, 115-134.

Cushman, F., Young, L., and Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment: testing three principles of harm. *Psychol Sci* 17, 1082-1089.

Foot, P. (1978). The problem of abortion and the doctrine of double effect. In *Virtues and Vices* (Oxford: Blackwell).

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

Royzman, E.B., and Baron, J. (2002). The preference for indirect harm. *Social Justice Research* 15, 165-184.

Thomson, J. (1985). The Trolley Problem. *Yale Law Journal* 94, 1395-1415.

Waldmann, M.R., and Dieterich, J.H. (2007). Throwing a bomb on a person versus throwing a person on a bomb: intervention myopia in moral intuitions. *Psychol Sci* 18, 247-253.