Two systems, etc.

Jonathan Baron    Burcu Gürçay    Antonios N. Cotaias
Razeen Jivani    Matthew H. Lazarus    Brett D. Neustadt
Don A. Moore

March 7, 2013
An act/omission dilemma

A guided missile was accidentally fired and is heading for a jet plane with 500 passengers. The only way to prevent this is for the air-traffic controller to instruct a smaller plane with 100 passengers to fly into the path of the missile (without telling the pilot why) and take the hit. Should the controller direct the smaller plane into

to path

yes no

What is the largest number of passenger deaths in the smaller plane that should be tolerated in order to prevent 500 passenger deaths in the larger plane?

500 400 300 200 100 0

This should not be tolerated no matter what harm is prevented by allowing it.
A utilitarian vs. rule dilemma with two acts

X is asked to testify for the prosecution at an insider trading trial. X knows that the defendant is innocent. But other witnesses have provided damaging testimony in which they distorted the truth. X also knows that, if he says what he knows, then the defendant will be wrongly convicted, because X’s testimony will be mis-interpreted as consistent with the distorted testimony of others. If X says he knows nothing, despite swearing to tell “the whole truth”, then the defendant will be acquitted, and nobody will find out that X lied about knowing nothing. What should X do?

Lie, saying that he knows nothing, in which case the court will correctly acquit the defendant.

Tell what he knows, as he swore under oath that he would do, in which case the defendant will be wrongly convicted.
Two systems theory (Greene et al.)

- long RTs for utilitarian responses to personal dilemmas
- cognitive interference affects utilitarian RT (but not choices)
- priming with CRT
- brain damage, psychopathy (but also anger)
- correlation with CRT
Two systems theory (Greene et al.)

- long RTs for utilitarian responses to personal dilemmas
- RT results may arise from response probability
- cognitive interference affects utilitarian RT (but not choices)
- priming with CRT

- brain damage, psychopathy (but also anger)
- correlation with CRT
Two systems theory (Greene et al.)

- long RTs for utilitarian responses to personal dilemmas
- RT results may arise from response probability
- cognitive interference affects utilitarian RT (but not choices)
- non-utilitarian responses may require less reading
- priming with CRT

- brain damage, psychopathy (but also anger)

- correlation with CRT
Two systems theory (Greene et al.)

- long RTs for utilitarian responses to personal dilemmas
- RT results may arise from response probability
- cognitive interference affects utilitarian RT (but not choices)
- non-utilitarian responses may require less reading
- priming with CRT
- utilitarian response may require more attention to number
- failed to replicate with expanded CRT
- brain damage, psychopathy (but also anger)
- correlation with CRT
Two systems theory (Greene et al.)

- long RTs for utilitarian responses to personal dilemmas
- RT results may arise from response probability
- cognitive interference affects utilitarian RT (but not choices)
- non-utilitarian responses may require less reading
- priming with CRT
- utilitarian response may require more attention to number
- failed to replicate with expanded CRT
- brain damage, psychopathy (but also anger)
- emotional response to action may affect it
- correlation with CRT
Two systems theory (Greene et al.)

- long RTs for utilitarian responses to personal dilemmas
- RT results may arise from response probability
- cognitive interference affects utilitarian RT (but not choices)
- non-utilitarian responses may require less reading
- priming with CRT
- utilitarian response may require more attention to number
- failed to replicate with expanded CRT
- brain damage, psychopathy (but also anger)
- emotional response to action may affect it
- correlation with CRT
- other explanations, discussed later
Mouse-pointer tracking study

A viral epidemic has spread across a town. You are a medical researcher and have developed two substances in your laboratory. You know that one of them is a vaccine, but you don't know which because both of the vials have been mislabeled as vaccine. You also know that the other substance is deadly.

Once you figure out which substance is the vaccine you can create more to save many lives.

When you have finished reading, click the button below for the number and method, and then don't move the mouse/pointer until you are finished reading again.
After click

A viral epidemic has spread across a town. You are a medical researcher and have developed two substances in your laboratory. You know that one of them is a vaccine, but you don't know which because both of the vials have been mislabeled as vaccine. You also know that the other substance is deadly.

Once you figure out which substance is the vaccine you can create more to save many lives.

You have two lab assistants who work with you, and the only way to identify the vaccine with certainty is to inject the two substances into these people against their wishes.

One person will live, the other will die, and you will be able to start saving lives with your vaccine.

Is it morally wrong to kill one of these people with a deadly injection in order to identify a vaccine that will save 2,500 lives?

- no (not wrong)
- yes (wrong)
No switch
Bad fit (sometimes happens)
Experiments with standard dilemmas

All of these used personal and impersonal dilemmas, as modified by McGuire, Moore, Starcke, Gürçay, and their co-authors. Most contrasted personal/impersonal, but I ignore this here. Gürçay and Baron varied number of lives saved within each page, and I ignore that too. Most responses were independent of number, but for the rest I count 3/5 yes’s as “yes”.

We also varied instructions to respond carefully or intuitively. That had no effect on choice (although it did affect response times, as intended).

Although Suter and Hertwig (2011) claimed to find an effect of such instructions, a re-analysis of their data also found no effects of their manipulation and no interaction with dilemmas.
Best fitting Rasch curves for data from M2008.
Proportion yes responses as a function of mean Ability of each quartile of Ability (M2008). Each line is a dilemma.
Mean log normalized response time as a function of Ability–Difficulty (scaled and rounded) for data from Moore et al. (2011). Circle areas are proportional to the number of observations in each point; red indicates “no”; green, “yes”, black is both combined.
Response-time function for **moral** items (Moore et al., 2011)
Response-time function for CRT items

Normalized log response time median vs. Ability–Difficulty z score.
F1. A bat and a ball cost $1.10 in total. The bat costs a dollar more than the ball. How much does the ball cost?

F2. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?

F3. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? (Frederick, 2005)
N2. If it takes 2 nurses 2 minutes to measure the blood pressure of 2 patients, how long would it take 200 nurses to measure the blood pressure of 200 patients?

N1. Soup and salad cost $5.50 in total. The soup costs a dollar more than the salad. How much does the salad cost?

N3. Sally is making sun tea. Every hour, the concentration of the tea doubles. If it takes 6 hours for the tea to be ready, how long would it take for the tea to reach half of the final concentration? (Finucane & Gullion, 2010).
Some extended CRT items.

B1. All flowers have petals.
   Roses have petals.
   If these two statements are true, can we conclude from them that roses are flowers.

B2. All mammals walk.
    Whales are mammals.
    If these two statements are true, can we conclude from them that whales walk.

B3. All things that have a motor need oil.
    Automobiles need oil.
    If these two statements are true, can we conclude from them that automobiles have a motor.
B4. All living things need water. 
   Roses need water. 
   If these two statements are true, can we conclude from them that roses are living things. (Markovits and Nantel, 1989)

B5. All vehicles have wheels. 
   Boats are vehicles. 
   If these two statements are true, can we conclude from them that boats have wheels. (De Neys & Franssens, 2009)

L1. If animals need vitamin Q, can we conclude that oysters need vitamin Q?

L2. If oxygen in the air is poisonous to animals, can we conclude that oxygen in the air is poisonous to dogs?
O1. Jack is looking at Anne but Anne is looking at George. Jack is married but George is not. Is a married person looking at an unmarried person? (A) Yes (B) No (C) Cannot be determined. (Toplak & Stanovich, 2002; see also Böckenholt, 2012)

O2. Ann’s father has a total of five daughters: Lala, Lele, Lili, Lolo, and ____. What is the name of the fifth daughter? (Krizo, 2011, but apparently older.)

O3. On the side of a boat hangs a ladder with six rungs. Each rung is one foot from the next one, and the bottom rung is resting on the surface of the water. The tide rises at a rate of one foot an hour. How long will take the water to reach the top rung? 5 hours, 6 hours, never (Edward Royzman, personal communication)
Results of obd1 (103 Ss, 91 with consistent moral responses)

- No effect of order of CRT and moral judgment.
- CRT correlates .24 with threshold (eliminating inconsistent Ss), p=.020.
- −.22 with PVs, p=.036.
- .25 with composite of both, p=.017.
- Seems to depend both on PVs and threshold.
- Belief-bias and numerical items best.
- Syllogisms poor (but also not well correlated with other items).
Correlation of CRT with moral judgment responses

Response (0 is PV, 6 is harm caused = harm prevented)
Correlation of Belief-bias and Numerical CRT items

CRT correlations of moral response

Response (0 is PV, 6 is harm caused = harm prevented)
Possible explanations of CRT results

- Ss view task as math problem. (Doesn’t account for belief-bias correlations.)
- Ss are unreflective about PVs (Baron and Leshner).
- Thoughtful people become utilitarians before the experiment.
- High CRT correlates with reflective education, and so does utilitarianism.
- CRT measures reflection-impulsivity ($r = 0.21$ for RT and accuracy).
Experiment (obd3) with rule dilemmas and utilitarianism scale

When a moral rule leads to outcomes that are worse than those from breaking the rule, we should **follow** the rule.

*Always*  

Sometimes but *not always*  

*Never*

When a moral rule leads to outcomes that are worse than those from breaking the rule, we should **break** the rule.

When two options harm other people in the same ways, we should choose the option that harms fewer people.

When we can help some people a lot by harming other people a little, we should **not** harm the second group of people.

When we can help some people a lot by harming other people a little, we should **not** harm the second group of people.

When one option has better effects on some people and worse effects on nobody than any other option, than this option should be chosen.
When one option has better effects on some people and worse effects on nobody than any other option, this option is not always the one that should be chosen.

Agree    Mostly agree    Mostly disagree    Disagree

For decision making that affects other people, all that matters is doing good and preventing harm.

It is worse to intentionally cause some harm through action than to cause the same harm intentionally by doing nothing to prevent it (through some easy action).

Sometimes we should follow rules that require us to do things that are harmful on the whole.

Sometimes we should follow rules that prevent us from doing what is best on the whole.

Some things should not be done even if they lead to very good outcomes.
## Correlations (obd3) ($\alpha$ in diagonals)

<table>
<thead>
<tr>
<th></th>
<th>CRT</th>
<th>CRT time</th>
<th>U-scale</th>
<th>U-rules</th>
<th>U-act/omit</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRT</td>
<td>.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRT time</td>
<td>.28</td>
<td>.93</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U-scale</td>
<td>.21</td>
<td>.26</td>
<td>.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U-rules</td>
<td>.23</td>
<td>.26</td>
<td>.35</td>
<td>.53</td>
<td></td>
</tr>
<tr>
<td>U-act/omit</td>
<td>.05</td>
<td>.03</td>
<td>.21</td>
<td>.20</td>
<td>.77</td>
</tr>
<tr>
<td>R/I</td>
<td></td>
<td>.29</td>
<td>.30</td>
<td>.05</td>
<td></td>
</tr>
</tbody>
</table>

P-levels: .29 is $p < .01$, .22 is $p < .05$. 
Conclusions of obd3

- Failed to replicate CRT correlation with act-omit items (and also failed with threshold and PV measures.)
- But did find correlation with U-rules items
- The correlation is found both for time and accuracy
- And it is highest when both are used together to form a measure of reflection/impulsivity.
- CRT measures also correlate with endorsement of utilitarian principles stated abstractly.
- And this scale of endorsement correlates with utilitarian responses to dilemmas (but n.q.s. for act/omit cases).
Overall conclusions

- Moral judgment in experiments does not seem to involve sequential processes.
- It can be explained by conflict between competing and simultaneous processes.
- Utilitarian responses to dilemmas seem to relate to general moral beliefs that are explicitly endorsed.
- Utilitarian beliefs and choices seem related to reflection-impulsivity.
- Why?