Moral judgment and measures of cognitive style: cognitive reflection, reflection-impulsivity, and actively open-minded thinking

Jonathan Baron (with contributions from others, especially Sydney Scott and Katrina Fincher)

July 9, 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 the 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.
CRT items

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).
Utilitarianism scale used in obd3

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 **do this**.

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></td>
<td></td>
<td>.29</td>
<td>.30</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).
Study with elaborate CRT (OBD4)

- Do lures matter for correlation with U-scale? (Omitted some CRT items to make room for new ones)
- Does belief bias matter directly?
- Also included U-rule items, before CRT
- Improved U-scale (after CRT)
CRT types in addition to original arithmetic items (n1–n3 & f1–f3)

- **Arithmetic no-lure (a1–a6)** If it takes 1 nurse 5 minutes to measure the blood pressure of 6 patients, how many minutes would it take 100 nurses to measure the blood pressure of 300 patients?

- **Belief consistent (1c–4c)**, All aunts are sisters. Some women are aunts. If these two statements are true, can we conclude from them that some women are sisters?

- **Belief neutral (1n–4n)** All laloobays are rich. Sandy is a laloobay. If these two statements are true, can we conclude from them that Sandy is rich?

- **Belief inconsistent (1i–4i & b1–b5)** All bears are ferocious. Some stuffed animals are bears. If these two statements are true, can we conclude from them that some stuffed animals are ferocious?
Revised U-scale

- When a moral rule leads to outcomes that are worse than those from breaking the rule, we should **follow** the rule.
- 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 one option has better effects on some people and worse effects on nobody than any other option, then we should choose this option.
- For decision making that affects other people, all that matters is doing good and preventing harm.
- 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 (obd4) (\(\alpha\) in diagonals)

<table>
<thead>
<tr>
<th>CRT-score</th>
<th>lure time</th>
<th>no-lure time</th>
<th>U-scale</th>
<th>U-rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRT-score</td>
<td>.92</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lure-time</td>
<td>.15</td>
<td>.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>no-lure-time</td>
<td>.34</td>
<td>.79</td>
<td>.91</td>
<td></td>
</tr>
<tr>
<td>U-scale</td>
<td>-.04</td>
<td>.31</td>
<td>.25</td>
<td>.72</td>
</tr>
<tr>
<td>U-rules</td>
<td>.11</td>
<td>.07</td>
<td>.11</td>
<td>.48</td>
</tr>
</tbody>
</table>

.30 is \(p = .001\) (1 tail), .23 is \(p = .01\), .19 is \(p = .025\), .16 is \(p = .05\).
IRT parameters (obd4)
Correlations with all other items and arithmetic-lure items (obd4)
Conclusions of obd4

- Response times (RT) correlate positively with CRT.
- Items without lures behave like other items. (But consistent belief-bias items are useful only to keep Ss honest.)
- RT to CRT correlate with utilitarian beliefs
- Lures do not matter for this correlation (difference n.s.).
- Utilitarian beliefs correlate with utilitarian responses to rule dilemmas.
- But correlation of CRT score with beliefs, or with dilemmas, does not replicate.
- Still, beliefs correlate with dilemmas, and RT with beliefs.
CRT, AOT, religion and utilitarian judgment (obd5)

- Best ActNum (7 items, $\alpha = .73$) and ActRule (8 items, $\alpha = .69$) from previous studies
- Short CRT scale ($\alpha = .78$) with times ($\alpha = .91$): 1i, 1n, 1c, n2, n1, n3, 2i, 2n, a1, a2, 3n, 3i
- Expanded utilitarian belief scale (Uscale, 2 parts, “choices” and “morality”) ($\alpha = .60$)
- Religion items from Piazza and Landy (2013) ($\alpha = .83$)
- AOT scale with one added item ($\alpha = .67$)
X is the inspector of a nuclear power plant that X suspect has not met its safety requirements. The plant foreman and X are touring the facility when one of the nuclear fuel rods overheats. The emergency coolant system fails to activate, and a chain reaction is about to begin, which will result in a nuclear meltdown. This will release lethal radiation into the nearby town, killing many people.

X realizes that the only way to stop the meltdown is to manually release liquid nitrogen into the fuel rod chamber.

This will remove just enough heat energy from the rod assembly to prevent the nuclear chain reaction. However, it will also instantly kill an employee trapped nearby.

Should X kill the employee in order to save the people in the nearby town?
Some people suppose that there are other options, or that the consequences might be different from what the story says. If you did this and it affected your answer, please change your answer now, so that it is the answer you would give if there were no other options and no additional consequences.
Utilitarian beliefs (Uscale): “Choices” \( (\alpha = .60) \)

Response scale: Always \ldots Never (4 points)

- When a moral rule leads to outcomes that are worse than those from breaking the rule, we should **follow** the rule. (-)
- 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 one option has better effects on some people and worse effects on nobody than any other option, then we should choose this option.
- When we can help some people a lot by harming fewer people a little, we should do this.
Utilitarian beliefs: “Choices”

Response scale: Agree … Disagree (4 points)

► We should not harm some people in order to help other people.
► For decision making that affects other people, all that matters is doing good and preventing harm.
► 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. (-)
Utilitarian beliefs: “Morality questions”

Response scale: Agree . . . Disagree (4 points)

- Killing someone can be morally right if it is for the greater good.
- It is always morally wrong to assist people in ending their lives. (-)
- Torture can sometimes be morally right, if it is for the greater good.
- It is always morally wrong to have sexual relations with a family member. (-)
- It is always morally wrong to betray your country. (-)
Religion scale (Relig): “Morality questions” ($\alpha = .83$)

Response scale: Agree ... Disagree (4 points)

- The truth about morality is revealed only by God.
- It is possible to live a righteous life without knowledge of God's laws. (-)
- Acts that are immoral are immoral because God forbids them.
- We don't need to try to figure out what is right and wrong, the answers have already been given to us by God.
- An atheist can still understand what is morally right and wrong. (-)
- Without God, humans still have a way to distinguish right from wrong. (-)
AOT scale: “Questions about thinking” ($\alpha = .67$)

Response scale: Strongly agree . . . Strongly disagree (5 points)

- Allowing oneself to be convinced by an opposing argument is a sign of good character.
- People should take into consideration evidence that goes against their beliefs.
- People should revise their beliefs in response to new information or evidence.
- Changing your mind is a sign of weakness. (-)
- Intuition is the best guide in making decisions. (-)
- It is important to persevere in your beliefs even when evidence is brought to bear against them. (-)
- One should disregard evidence that conflicts with one’s established beliefs. (-)
- People should search actively for reasons why their beliefs might be wrong.
Item discrimination and “validity”
Correlations, disattenuated correlations on top (obd5)

<table>
<thead>
<tr>
<th></th>
<th>Relig</th>
<th>AOT</th>
<th>CRT</th>
<th>CRTrt</th>
<th>Uscale</th>
<th>ActRule</th>
<th>ActOmit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relig</td>
<td>0.83</td>
<td>−0.817</td>
<td>−0.392</td>
<td>−0.272</td>
<td>−0.808</td>
<td>−0.264</td>
<td>−0.346</td>
</tr>
<tr>
<td>AOT</td>
<td>−0.609</td>
<td>0.67</td>
<td>0.530</td>
<td>0.339</td>
<td>0.683</td>
<td>0.417</td>
<td>0.285</td>
</tr>
<tr>
<td>CRT</td>
<td>−0.315</td>
<td>0.383</td>
<td>0.78</td>
<td>0.469</td>
<td>0.577</td>
<td>0.323</td>
<td>0.212</td>
</tr>
<tr>
<td>CRTrt</td>
<td>−0.237</td>
<td>0.265</td>
<td>0.395</td>
<td>0.91</td>
<td>0.250</td>
<td>0.389</td>
<td>0.104</td>
</tr>
<tr>
<td>Uscale</td>
<td>−0.570</td>
<td>0.433</td>
<td>0.395</td>
<td>0.185</td>
<td>0.60</td>
<td>0.318</td>
<td>0.611</td>
</tr>
<tr>
<td>ActRule</td>
<td>−0.200</td>
<td>0.284</td>
<td>0.237</td>
<td>0.308</td>
<td>0.205</td>
<td>0.69</td>
<td>0.436</td>
</tr>
<tr>
<td>ActNum</td>
<td>−0.270</td>
<td>0.200</td>
<td>0.160</td>
<td>0.085</td>
<td>0.404</td>
<td>0.310</td>
<td>0.73</td>
</tr>
</tbody>
</table>

Reliabilities in bold. Raw $r = .200$ is $p = .05$ 2-tailed.
Oblimin factor analysis, 2 factors

Relig
Uscale
AOT
ActNum
CRT-rt
CRT
ActRule
MR1
-0.88
0.65
0.61
0.31
MR2
0.64
0.53
0.46
0.54
Summary of stepwise (down) regressions

Significant predictors in bold

- **Uscale**  \( AOT + \textbf{Relig} + \textbf{CRT} + \text{CRTrt} \)
- **CRT**  \( \textbf{AOT} + \text{Relig} \)
- **CRTrt**  \( \textbf{AOT} + \text{Relig} \)
- **AOT**  \( \textbf{Relig} + \textbf{CRT} + \text{CRTrt} \)
- **Relig**  \( \textbf{AOT} + \text{CRT} + \text{CRTrt} \)
- **ActNum**  \( \textbf{Uscale} + \text{CRT} + \text{CRTrt} + \text{AOT} + \text{Relig} \)
- **ActRule**  \( \textbf{Uscale} + \textbf{CRT} + \text{CRTrt} + \textbf{AOT} + \text{Relig} \)
Overall conclusions

- Utilitarian responses to dilemmas seem to relate to general moral beliefs that are explicitly endorsed.
- Utilitarian beliefs, religion, and AOT are all closely related.
- Utilitarian beliefs (and sometimes choices) seem related to reflection-impulsivity, and the expanded CRT.
- The CRT is also related to AOT and religion.
- The predictive power of the CRT seems more related to R/I than to sequential use of two systems.