The Neuropsychology of
Moral Luck

Seyed Payam Kamaneh

Institute for Research in Fundamental Sciences (IPM)

School of Analytic Philosophy

What we'll discuss:

1. Philosophical preliminaries

2. From moral dilemmas to moral psychology (I): the case of utilitarian judgments

3. From moral dilemmas to moral psychology (II): the case of moral luck

3.1. Cushman's dual-process account

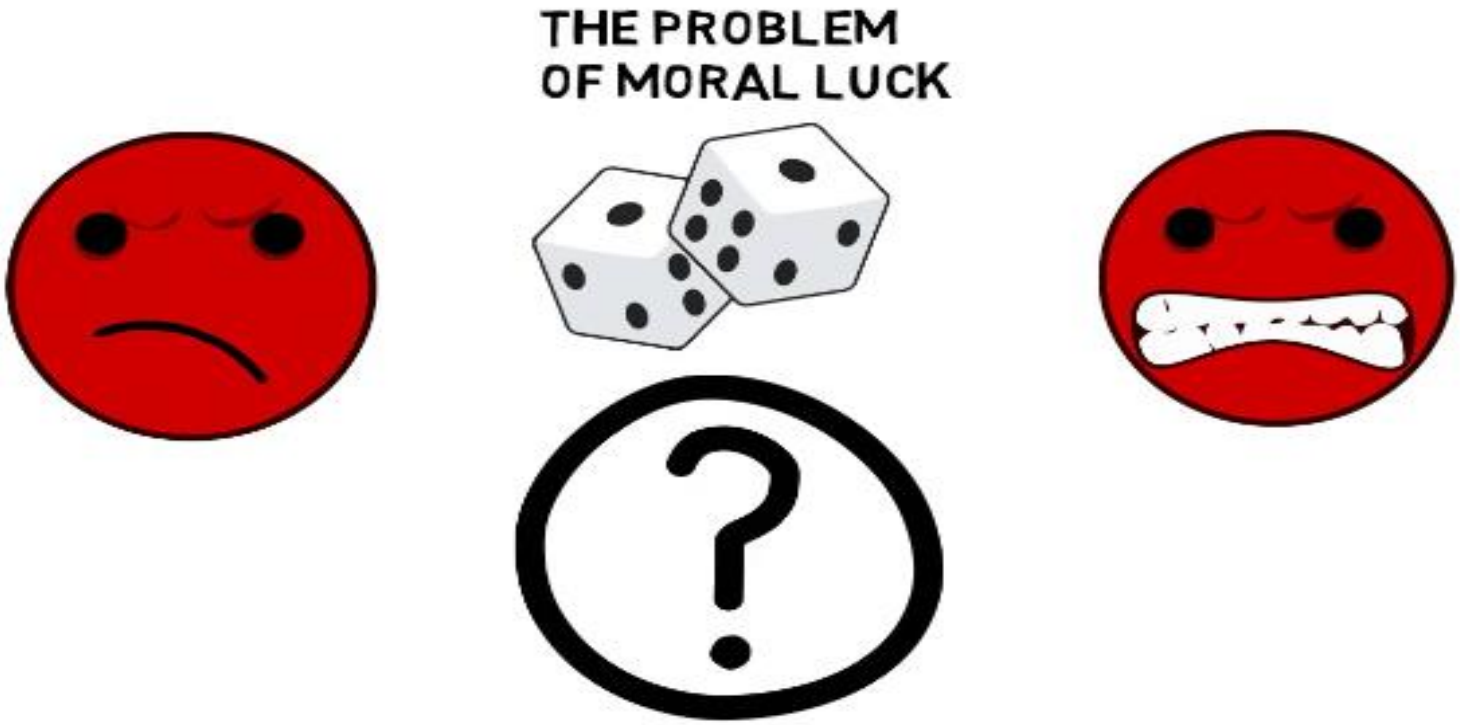
3.2. Nichols' epistemic assessment account

4. Suggestions for future research

Philosophical preliminaries

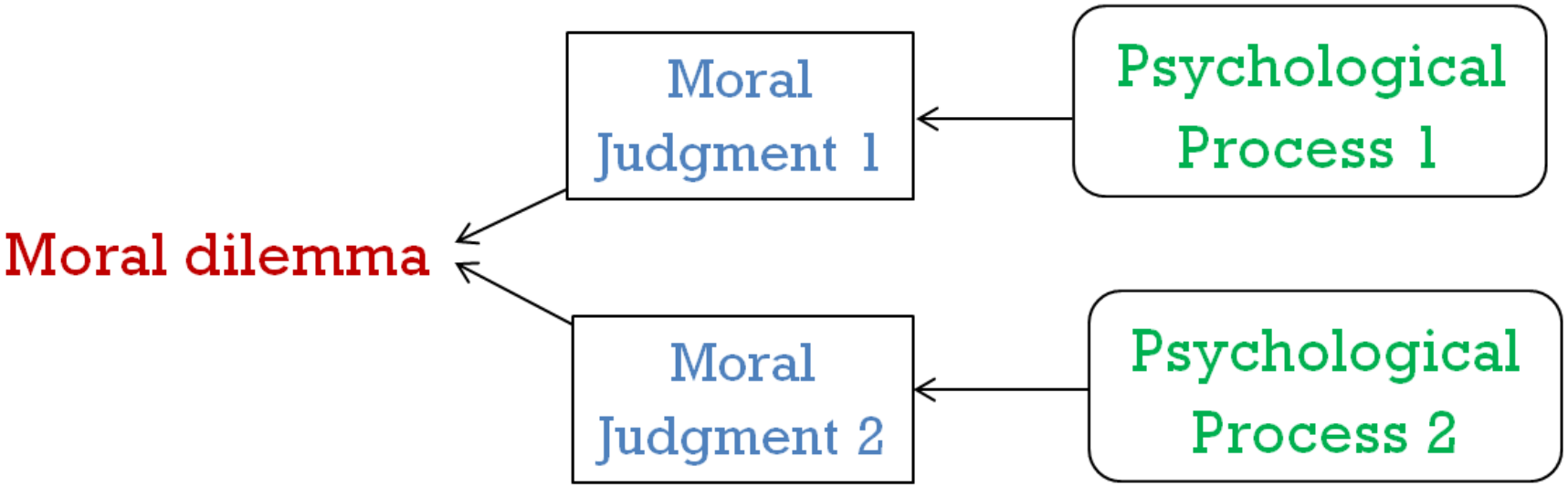
- If an action is not under the control of an agent, it depends on luck (e.g. our heartbeat).
- On many occasions, we are judged by having properties (values) their possession is not under our control (e.g. our ethnicity).
- *Moral* value is the supreme sort of value, and *luck* has (and should have) no role in making moral judgments.
- Control principle (CP): We are morally assessable only to the extent that what we are assessed for depends on factors under our control.
- CP-Corollary: Two people ought not to be morally assessed differently if the only other differences between them are due to factors beyond their control.
- Question: Does the CP-Corollary hold in the actual world?

Philosophical preliminaries



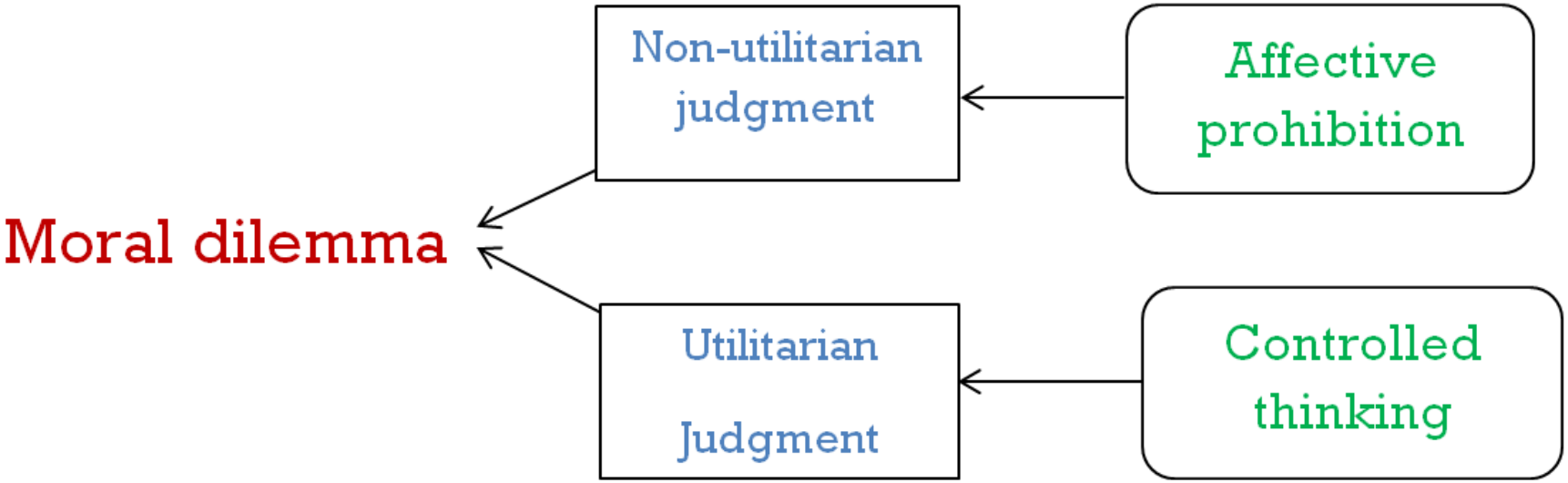
- Intuition 1: They should <u>not</u> be morally assessed differently.
- Intuition 2: They should be morally assessed differently.

From moral dilemmas to moral psychology



From moral dilemmas to moral psychology (I):

The case of utilitarian judgments



The case of utilitarian judgments

Evidence for Greene's dual-process account:

1) Moral judgments made in response to dilemmas like the "crying baby" case are associated with increased neural activity in regions associated with *emotion*, including subregions of medial prefrontal cortex (mPFC) and the amygdala.

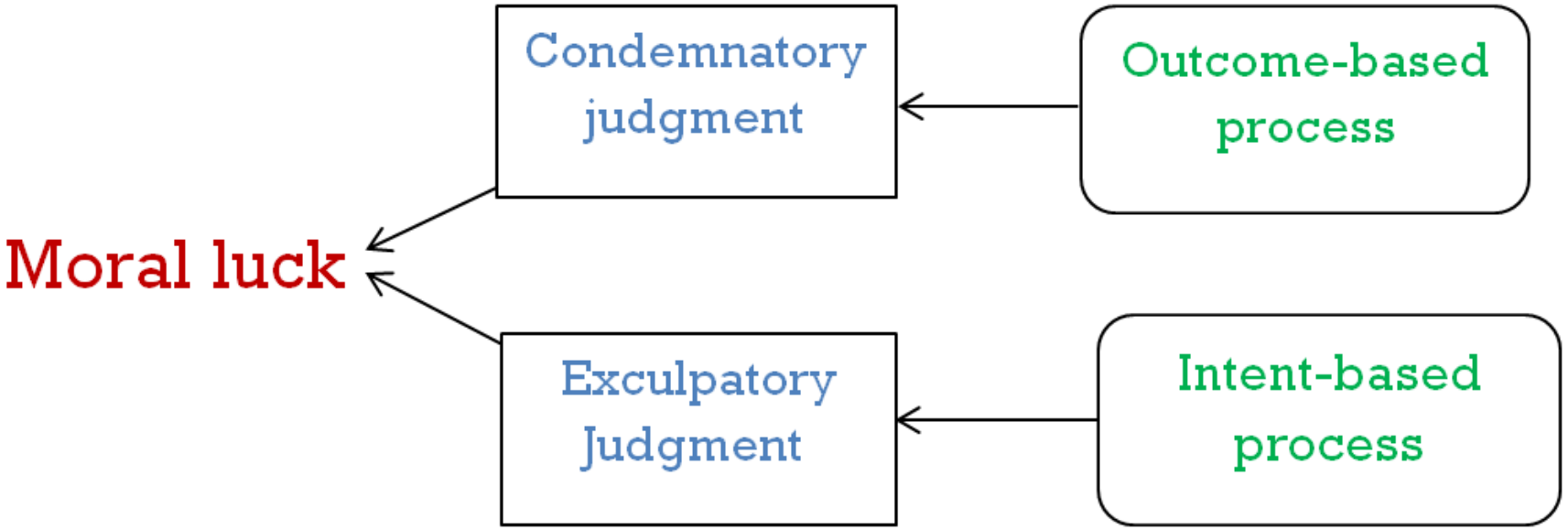
2) Individuals with damage to vmPFC are far more likely than healthy individuals to endorse harmful behavior in order to promote a greater good.

3) According to fMRI studies, *willingness to harm* was associated with greater activation in regions associated with *cognitive control* and *thinking guided by explicit rules* (including dorsolateral prefrontal cortex (DLPFC)).

4) Subjects under *cognitive load* take significantly longer to make utilitarian judgments.

From moral dilemmas to moral psychology (II):

Cushman's dual-process account of moral luck



A useful distinction

Kinds of harm	Intention	Negative outcome
Intentional	+	+
Accidental	-	+
Attempted	+	-

Evidence for Cushman's dual-process account of moral luck

1) Children exhibit consistency across development in the judgment of *attempted* harms, but exhibit developmental change in the judgment of *accidental* harms.

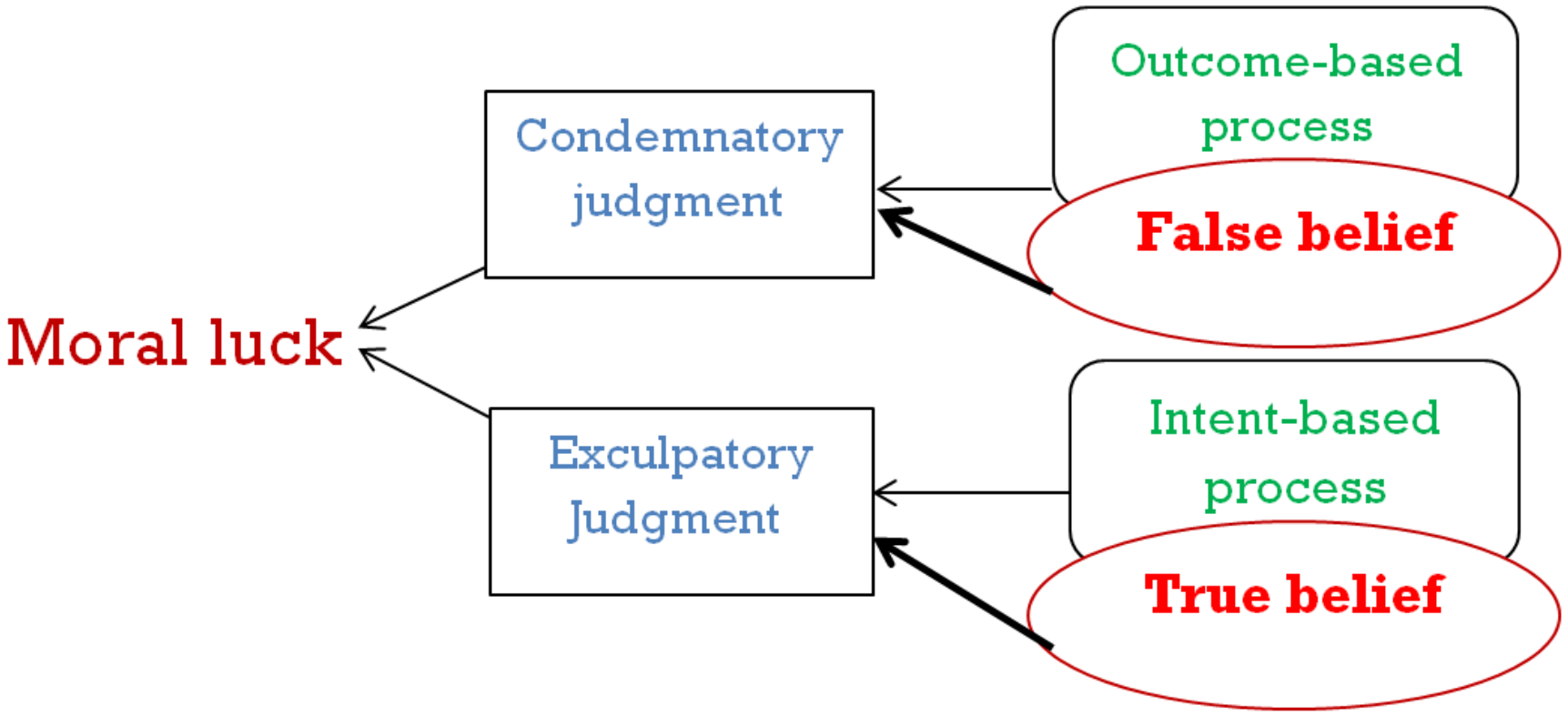
2) Functional neuroimaging of adults engaging in moral judgment reveals activation of brain regions implicated in *cognitive conflict* and control for the judgment of *accidental harms*, but not attempted harms.

3) *The Blame blocking effect*: The occurrence of a *coincidental harm* blocks people from assigning punishment to an attempted murder.

4) Transcranial magnetic stimulation (TMS) to the right temporoparietal junction (rTPJ) causes participants to judge attempted harms as less morally forbidden and more morally permissible.

From moral dilemmas to moral psychology (II):

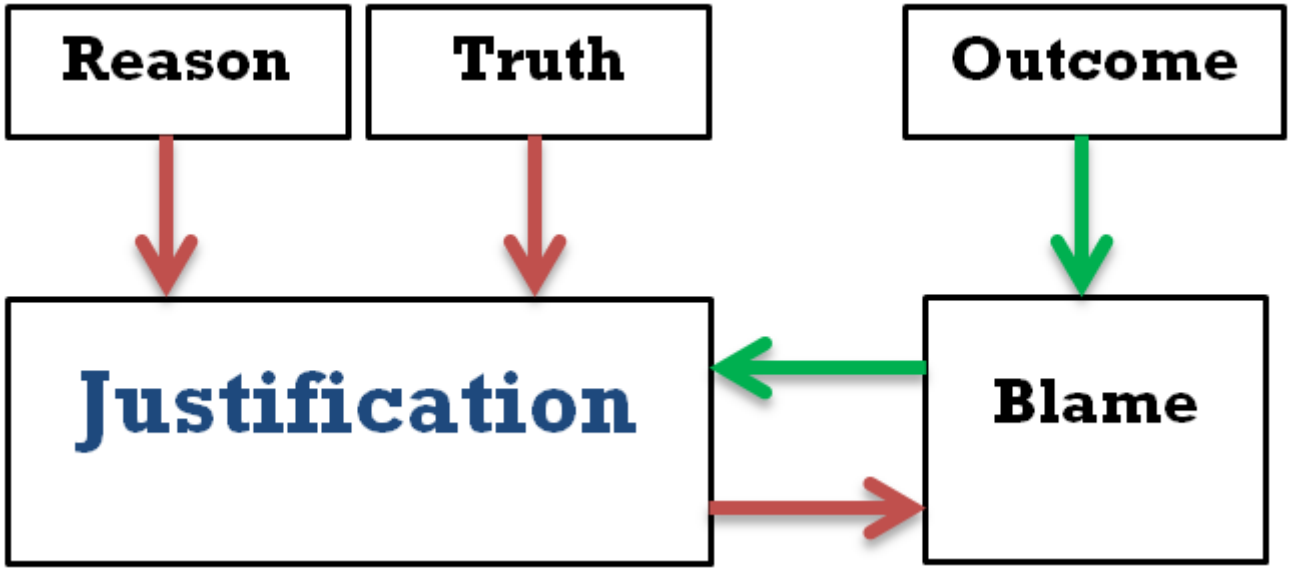
Nichols' Epistemic assessment account of moral luck



A novel concept

Kinds of agent	Negative outcome	Belief
Lucky	-	T
Unlucky	+	F
Extra lucky	-	F

Nichols' Epistemic assessment account: more details

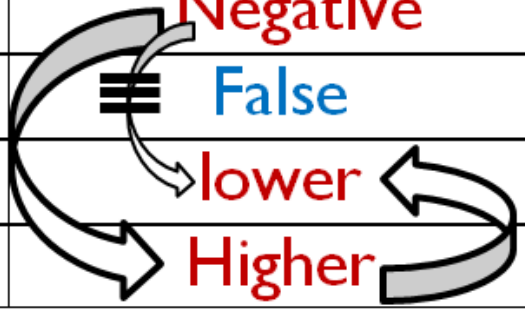


Comparison 1 (lucky vs. extra lucky agents)

Variables	Lucky agent	Extra lucky agent
Outcome	Neutral	Neutral
Truth value	True	False
Justification	Higher	lower
Blame	Lower	Higher

Comparison 2 (unlucky vs. extra lucky agents)

Variables	unlucky agent	Extra lucky agent
Outcome	Negative	Neutral
Truth value	False	False
Justification	lower	Higher
Blame	Higher	lower



Neural evidence for the *direct* influence of bad outcome on moral judgment

- If bad outcomes lead to harsher moral judgments and then to greater consideration of agents' beliefs, we should expect bad outcomes to be associated with enhanced activation of brain regions for belief reasoning, *late in the trial*.

- In two regions consistently associated with belief reasoning in moral and non-moral contexts, the RTPJ and the LTPJ, the response was significantly higher following bad outcomes versus neutral outcomes. The differential neural response appeared quite late in the stimulus, around the time of the moral judgment, rather than at the time that the belief was presented, prior to the judgment.

Philosophical implications of epistemic assessment account

Moral judgments do appear to be dominated by factors we reflectively endorse as morally relevant: whether agents have good or bad reasons for their beliefs, whether these beliefs are true or false.

When assigning moral blame, we care mostly about whether agents are justified in thinking that *they won't cause harm*. To the extent that moral luck asymmetries are driven by such mental state assessments, we may be able to defend a rational approach to morality.

Question:

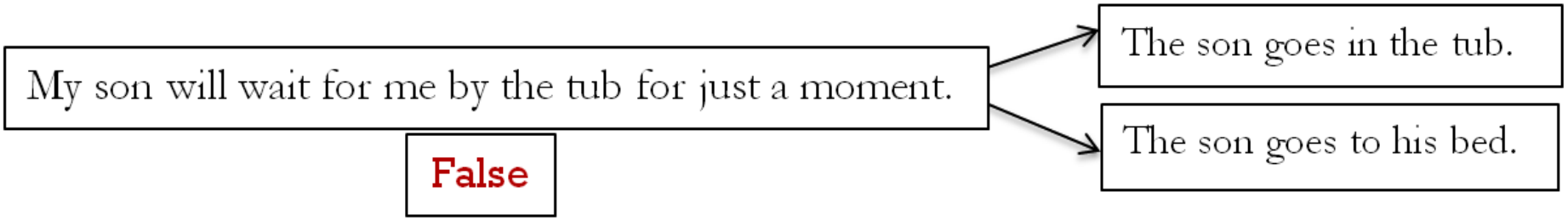
- Should the truth or falsity of a belief play a substantial role in *rationally assessing* its degree of justification?

Suggestions for future research

- Determining the exact content of ascribed belief in real-world cases:

Content of father's belief	T-value in The Lucky	T-value in The Extra Lucky
My son will wait for me by the tub for just a moment.	True	False
I won't cause harm to my son by leaving him alone.	True	True

- Examining different ways in which the ascribed belief can be false:



- Examining the impact of true belief in moral assessment of attempted harm.