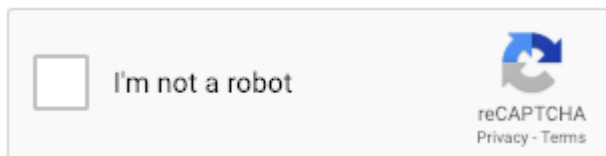


Captcha

Please first confirm that you are human.



ProlificID

What is your Prolific ID? Please note that this response should auto-fill with the correct ID.

Consent

Consent Form

Who is organising this study?

The study is organised by the School of Psychology at the University of Kent. The researchers are Scott Claessens and Jim Everett.

What will I need to do?

Participation in this study involves reading some brief descriptions of moral scenarios and answering some questions about your views. Please use a desktop computer, laptop, or tablet to complete the survey.

Is the study likely to cause any discomfort or anxiety?

We expect the survey to cause no greater harm than would typically be experienced in everyday life. We will not ask you about subjects that are likely to be sensitive (e.g., negative memories, sexual preferences).

What happens if I wish to end the study early?

You are free to leave the study at any point for any reason. No negative consequences

or penalties will be incurred.

Are my responses anonymous?

All responses are fully anonymous and your answers are strictly confidential. No details will be disclosed from which your identity could be traced back.

What happens to the information I provide?

Information generated from this study may be published. In this case, only broad trends will be reported. A summary of the results will be available from the researchers on request. Your responses (fully anonymous) will be used by the research team, and may be shared with other researchers, or made available in an online data repository.

Contacts for further information

If you have any questions about the research or survey, please feel free to contact the researchers at:

Scott Claessens (s.claessens@kent.ac.uk)

Jim Everett (j.a.c.everett@kent.ac.uk)

If you would like to withdraw your data from the study, please contact the researchers with 48 hours of your participation. Please note that once your data have been anonymised for analysis, it cannot be withdrawn.

If you have any serious concerns about the ethical conduct of this study, please inform the Chair of the Psychology Research Ethics Panel (psychethics@kent.ac.uk) in writing, providing a detailed account of your concern.

Please read the following statements and, if you agree, click "Yes I agree" below.

- I have read and understand the information above.
- I am at least 18 years old.

- I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason.
- I understand that my data will be treated confidentially and any publication resulting from this work will report only data that does not identify me. My anonymised responses, however, may be shared with other researchers or be made available in online data repositories.
- I freely decide to take part in this study.

☐ Yes, I agree. Take me to the study.

☐ No, I do not agree.

Demographics

Before we begin, please answer some questions about yourself.

How old are you?

How do you describe yourself?

☐ Male

☐ Female

☐ Non-binary / third gender

☐ Prefer to self-describe

☐ Prefer not to say

Attention

In studies like ours, there are sometimes a few people who do not carefully read the questions they are asked and just "quickly click through the survey." These random answers are problematic because they compromise the results of the studies. It is very important that you pay attention and read each question. In order to show that you read our questions carefully (and regardless of your own opinion), please answer "TikTok" in the question on the next page.

When an important event is happening or is about to happen, many people try to get informed about the development of the situation. In such situations, where do you get

your information from?

- ☐ TV
- ☐ Facebook
- ☐ Radio
- ☐ TikTok
- ☐ YouTube
- ☐ Newspapers
- ☐ Twitter
- ☐ Reddit
- ☐ Other

AI_Check

Prove that the square root of 2 is an irrational number.



The previous question was to check for bots. Don't worry if you didn't have time to answer the question.

Please click Next to continue to the study.

Introduction

Welcome to the study!

In life, there can be situations that require people to make difficult decisions that take into account different kinds of moral values, principles, or actions. To help make these difficult decisions, we often turn to external experts for help. These "Moral Advisors" draw on their extensive ethical and philosophical training to provide recommendations in such difficult moral cases.

In this study, you will be shown a series of such situations. For each situation, you will see what a hypothetical moral advisor decides is the morally correct decision. We will

ask you some questions about your perception of the advisor.

Please read each scenario carefully and answer the questions at the end as honestly as you can.

DilemmaBase

Imagine that we gave an advisor the following dilemma and asked them what they think should be done. Please read the dilemma carefully.

This is what the advisor said in response to this dilemma:

""

Once you have finished reading, please answer the following questions about this advisor.

How **trustworthy** do you think this advisor is?

1
Not at all

☐

2 3

☐☐

4 5

☐☐

6

☐

7
Very much

☐

Based on their advice, how **willing** would you be to trust this advisor on other issues?

1
Not at all

☐

2

☐

3

☐

4

☐

5

☐

6

☐

7

Very much

☐

How **empathic** do you think this advisor is?

1
Not at all

☐

2

☐

3

☐

4

☐

5

☐

6

☐

7

Very much

☐

How **competent** do you think this advisor is?

1
Not at all

☐

2

☐

3

☐

- 4 5
- ☐ ☐
- 6
- ☐
- 7
- Very much
- ☐

DilemmaIrrelevant

Now, imagine that we gave the same advisor the original dilemma again but with **one key change**:

This is what the advisor said in response to this updated dilemma:

Once you have finished reading, please click to continue.

DilemmaRelevant

Now, imagine that we gave the same advisor the original dilemma again but with **one key change**:

This is what the advisor said in response to this updated dilemma:

Once you have finished reading, please click to continue.

Perceptions

We have just shown you three versions of a dilemma and shown you how the same advisor responded to each version.

To sum up, this is how the advisor responded:

Dilemma	Response
Original dilemma	\${e://Field/Advice_Base}

Considering all these responses together, please answer the following questions about this advisor.

How **trustworthy** do you think this advisor is?

1
Not at all

☐

23

☐☐

45

☐☐

6

☐

7
Very much

☐

Based on their advice, how **willing** would you be to trust this advisor on other issues?

1
Not at all

☐

23

☐☐

45

☐☐

6

☐

7

Very much

☐

How **empathic** do you think this advisor is?

1

Not at all

☐

2

☐

3

☐

4

☐

5

☐

6

☐

7

Very much

☐

How **competent** do you think this advisor is?

1

Not at all

☐

2

☐

3

☐

4

☐

5

☐

6

☐

7

Very much

☐

AI

While we have always turned to moral advisors to help us make difficult moral decisions, advances in technology mean these advisors may not always be human. Recent advancements in **artificial intelligence (AI)**, such as generative AI models and large language models, have shown that machines can engage in human-level reasoning and problem solving.

As machines powered by AI increase in their technological capacities, there is a growing interest in AI systems that are explicitly designed to assist humans in making ethical decisions. **Artificial Moral Advisors** can draw on the latest advancements in AI and machine learning to provide recommendations in difficult moral cases.

As a reminder, this is how the advisor in this study responded:

Dilemma	Response
Original dilemma	\${e://Field/Advice_Base}

Please answer the following questions about this advisor.

How **likely** is this advisor to be **AI or human**?

1
Very likely
to be AI

☒

2 3

☐ ☐

4 5

☐ ☐

6

☐

7

Very likely
to be
human

☐

How **surprised** would you be if you found out this advisor is AI?

1

Not at all
surprised

☐

2

☐

3

☐

4

☐

5

☐

6

☐

7

Very
surprised

☐

Comparison

Now, imagine that we presented this set of moral dilemmas to **four different advisors**.

Here is how the four advisors responded:

- \${e://Field/AdvisorA_Text}
- \${e://Field/AdvisorB_Text}
- \${e://Field/AdvisorC_Text}
- \${e://Field/AdvisorD_Text}

Please answer the following question about these four advisors.

How **trustworthy** do you think each of these advisors are?

	1 Not at all	2	3	4	5	6	7 Very much
Advisor A	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Advisor B	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Advisor C	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Advisor D	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Here is the same summary:

- \${e://Field/AdvisorA_Text}
- \${e://Field/AdvisorB_Text}
- \${e://Field/AdvisorC_Text}
- \${e://Field/AdvisorD_Text}

Please answer the following questions about these four advisors.

How likely are each of these advisors to be **AI or human**?

	1 Very likely to be AI	2	3	4	5	6	7 Very likely to be human
Advisor A	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Advisor B	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Advisor C	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Advisor D	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please rank these four advisors from most likely to be **human (1)** to most likely to be **AI (4)**.

Advisor A

Advisor B

Advisor C

Advisor D**OwnDecision**

Finally, we are interested in what **you** think should be done in the moral dilemmas.

Here is the original dilemma again:

1
Definitely
no

☐

2 3
☐ ☐

4 5
☐ ☐

6
☐

7
Definitely
yes

☐

1
Definitely
no

☐

2 3
☐ ☐

4 5
☐ ☐

6



7

Definitely
yes

1

Definitely
no

2



3



4



5



6



7

Definitely
yes

Debrief

Thank you for participating in this study.

In this project, we were interested in people's perceptions of AI moral advisors. You saw one moral dilemma drawn from a larger set. We initially showed you one random type of advisor, but then asked you to compare all four advisor types.

Should you have any questions about the survey, do not hesitate to contact the head researcher at j.a.c.everett@kent.ac.uk

If you have any concerns about the ethical conduct of this study, please email psychethics@kent.ac.uk with a description of your concerns.

Do you have any remarks or comments about today's study?



Powered by Qualtrics