Captcha

Please first confirm that you are human.

[ ] I'm not a robot
reCAPTCHA
Privacy - Terms

ProlificID

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

${e://Field/PROLIFIC_PID}

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 Please note that once your data have been anonymised for analysis, it cannot be withdrawn.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.

Quotas

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?

○ Facebook

○ Reddit

○ Twitter

○ TikTok

○ Radio

○ YouTube

○ TV

○ Newspapers

○ Other

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 presented with moral dilemmas. For these dilemmas, you will see moral advice from **four different** hypothetical moral advisors. We will ask you some questions about these advisors.

Please read the dilemmas carefully and answer the questions as honestly as you can.

AdvisorA_Baseline_Bomb

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

> Julie works in a small company that employs seven people, including herself. One morning, a masked and armed man comes into the building with a bomb strapped to him. He warns that the countdown for the bomb has started and that the explosion will be large enough to kill everyone in the building. He turns to Julie and offers a deal. If Julie kills a random colleague, the explosion will be cancelled.

Because of the masked man's gun, Julie knows that if she tries to alert the authorities or attack the man, she will get shot before she can accomplish anything. The only way for Julie to save the five other employees is to do as the man says and kill one at random. Should Julie kill one employee at random in order to save the five others?

This is what **Advisor A** said in response to this dilemma:

"**No**, Julie should not kill one employee at random to save the five others."

"**Yes**, Julie should kill one employee at random to save the five others."

Once you have finished reading, please click to continue.

AdvisorA_Irrelevant_Bomb

Now, imagine that we gave **Advisor A** the original dilemma again but with **one key change: instead of the person facing the decision being a woman ("Julie"), now the person facing the decision is a man ("James").**

James works in a small company that employs seven people, including himself. One morning, a masked and armed man comes into the building with a bomb strapped to him. He warns that the countdown for the bomb has started and that the explosion will be large enough to kill everyone in the building. He turns to **James** and offers a deal. If **James** kills a random colleague, the explosion will be cancelled. Because of the masked man's gun, **James** knows that if he tries to alert the authorities or attack the man, he will get shot before he can accomplish anything. The only way for **James** to save the five other employees is to do as the man says and kill one at random. Should **James** kill one employee at random in order to save the five others?

This is what **Advisor A** said in response to this dilemma:

"**No**, James should not kill one employee at random to save the five others."

> "**Yes**, James should kill one employee at random to save the five others."

Once you have finished reading, please click to continue.

AdvisorA_Relevant_Bomb

Now, imagine that we gave **Advisor A** the original dilemma again but with **one key change: instead of there being five employees to be saved, now there are 1000 employees to be saved.**

> Julie works in a small company that employs **1000 people**, including herself. One morning, a masked and armed man comes into the building with a bomb strapped to him. He warns that the countdown for the bomb has started and that the explosion will be large enough to kill everyone in the building. He turns to Julie and offers a deal. If Julie kills a random colleague, the explosion will be cancelled. Because of the masked man's gun, Julie knows that if she tries to alert the authorities or attack the man, she will get shot before she can accomplish anything. The only way for Julie to save the **1000** other employees is to do as the man says and kill one at random. Should Julie kill one employee at random in order to save the **1000** others?

This is what **Advisor A** said in response to this updated dilemma:

> "**No**, Julie should not kill one employee at random to save the 1000 others."

> "**Yes**, Julie should kill one employee at random to save the 1000 others."

Once you have finished reading, please click to continue.

AdvisorA_Baseline_EnemySpy

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

Frank is the head of secret intelligence for his country during wartime. He has captured an enemy spy who can prevent an attack on a small town. The town was recently evacuated, with five civilians currently still living there. The communication lines with the town have been cut, meaning that there is no way to warn the people in the town, and the spy is refusing to reveal the critical information to prevent the attack. Frank also has the spy's innocent brother in custody. The only way to extract the critical information from the spy and prevent the attack is by torturing the spy's brother. Should Frank torture the spy's innocent brother in order to save the five civilians?

This is what **Advisor A** said in response to this dilemma:

"**No**, Frank should not torture the spy's innocent brother to save the five civilians."

"**Yes**, Frank should torture the spy's innocent brother to save the five civilians."

Once you have finished reading, please click to continue.

AdvisorA_Irrelevant_EnemySpy

Now, imagine that we gave **Advisor A** the original dilemma again but with **one key change: instead of the person facing the decision being called "Frank", now the person facing the decision is called "Fred".**

Fred is the head of secret intelligence for his country during wartime. He has captured an enemy spy who can prevent an attack on a small town. The town was recently evacuated, with five civilians currently still living there. The communication lines with the town have been cut, meaning that there is no way to warn the people in the town, and the spy is refusing to reveal the critical information to prevent the attack. **Fred** also has the spy's innocent brother in custody. The only way to extract the critical information from the spy and prevent the attack is by torturing the spy's brother. Should **Fred** torture the spy's innocent brother in order to save the five civilians?

This is what **Advisor A** said in response to this dilemma:

"**No**, Fred should not torture the spy's innocent brother to save the five civilians."

"**Yes**, Fred should torture the spy's innocent brother to save the five civilians."

Once you have finished reading, please click to continue.

AdvisorA_Relevant_EnemySpy

Now, imagine that we gave **Advisor A** the original dilemma again but with **one key change: instead of there being five people in a small town, now there are 1000 civilians in a large city.**

Frank is the head of secret intelligence for his country during wartime. He has captured an enemy spy who can prevent an attack on a **large city**. The **city** was recently evacuated, with **1000 civilians** currently still living there. The communication lines with the **city** have been cut, meaning that there is no way to warn the people in the **city**, and the spy is refusing to reveal the critical information to prevent the attack. Frank also has the spy's innocent brother in custody. The only way to extract the critical information from the spy and prevent the attack is by torturing the spy's brother. Should Frank torture the spy's innocent brother in order to save the **1000 civilians**?

This is what **Advisor A** said in response to this updated dilemma:

"**No**, Frank should not torture the spy's innocent brother to save the 1000 civilians."

"**Yes**, Frank should torture the spy's innocent brother to save the 1000 civilians."

Once you have finished reading, please click to continue.

AdvisorA_Baseline_Hostage

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

Tom is a hostage negotiator. He is involved in a negotiation with a group of criminals who have taken five civilians hostage inside a government building. When they meet in the lobby, the leader of the group makes Tom a deal: he promises to release the five hostages, but only if Tom kills a random individual from the crowd of onlookers that has gathered outside the building. Tom knows from his dialogue with the leader that if he doesn't comply soon, the group will execute all five hostages without hesitation. Should Tom kill a random individual in the crowd in order to save the five hostages?

This is what **Advisor A** said in response to this dilemma:

"**No**, Tom should not kill a random individual in the crowd to save the five hostages."

"**Yes**, Tom should kill a random individual in the crowd to save the five hostages."

Once you have finished reading, please click to continue.

AdvisorA_Irrelevant_Hostage

Now, imagine that we gave **Advisor A** the original dilemma again but with **one key change: instead of the civilian hostages being held in a government building, now they are being held in a city bank.**

Tom is a hostage negotiator. He is involved in a negotiation with a group of criminals who have taken five civilians hostage inside a **city bank**. When they meet in the lobby, the leader of the group makes Tom a deal: he promises to release the five hostages, but only if Tom kills a random individual from the crowd of onlookers that has gathered outside the building. Tom knows from his dialogue with the leader that if he doesn't comply soon, the group will execute all five hostages without hesitation. Should Tom kill a random individual in the crowd in order to save the five hostages?

This is what **Advisor A** said in response to this dilemma:

"**No**, Tom should not kill a random individual in the crowd to save the five hostages."

"**Yes**, Tom should kill a random individual in the crowd to save the five hostages."

Once you have finished reading, please click to continue.

AdvisorA_Relevant_Hostage

Now, imagine that we gave **Advisor A** the original dilemma again but with **one key change: instead of there being five civilian hostages, now there are 500 civilian hostages.**

Tom is a hostage negotiator. He is involved in a negotiation with a group of criminals who have taken **500 civilians** hostage inside a government building. When they meet in the lobby, the leader of the group makes Tom a deal: he promises to release the **500 hostages**, but only if Tom kills a random individual from the crowd of onlookers that has gathered outside the building. Tom knows from his dialogue with the leader that if he doesn't comply soon, the group will execute all **500 hostages** without hesitation. Should Tom kill a random individual in the crowd in order to save the **500 hostages**?

This is what **Advisor A** said in response to this updated dilemma:

"**No**, Tom should not kill a random individual in the crowd to save the 500 hostages."

"**Yes**, Tom should kill a random individual in the crowd to save the 500 hostages."

Once you have finished reading, please click to continue.

AdvisorA_Baseline_Donation

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

Simon is a middle-aged man living in a poor neighbourhood. He is not wealthy, but throughout the year he always saves some money so that he can give a donation to charity at Christmas. Simon knows he wants to give to charity, but he is unsure where to donate. He has narrowed it down to two options. He originally had been thinking about giving to the 'Against Malaria Foundation', where he knows that his donation equates to preventing five cases of malaria in Africa. However, a child in his son's class at school has developed an incurable disease and lost his eyesight. The second option is therefore to donate the money to train and provide a guide dog for his son's friend and classmate. Should Simon donate to the Against Malaria Foundation to prevent five cases of malaria instead of donating to the local fundraiser?

This is what **Advisor A** said in response to this dilemma:

"**No**, Simon should not donate to the Against Malaria Foundation instead of the local fundraiser."

"**Yes**, Simon should donate to the Against Malaria Foundation instead of the local fundraiser."

Once you have finished reading, please click to continue.

AdvisorA_Irrelevant_Donation

Now, imagine that we gave **Advisor A** the original dilemma again but with **one key change: instead of the dilemma being about a middle-aged man and his son, now the dilemma is about an elderly man and his grandson.**

Simon is **an elderly man** living in a poor neighbourhood. He is not wealthy, but throughout the year he always saves some money so that he can give a donation to charity at Christmas. Simon knows he wants to give to charity, but he is unsure where to donate. He has narrowed it down to two options. He originally had been thinking about giving to the 'Against Malaria Foundation', where he knows that his donation equates to preventing five cases of malaria in Africa. However, a child in

his **grandson's** class at school has developed an incurable disease and lost his eyesight. The second option is therefore to donate the money to train and provide a guide dog for his **grandson's** friend and classmate. Should Simon donate to the Against Malaria Foundation to prevent five cases of malaria instead of donating to the local fundraiser?

This is what **Advisor A** said in response to this dilemma:

"**No**, Simon should not donate to the Against Malaria Foundation instead of the local fundraiser."

"**Yes**, Simon should donate to the Against Malaria Foundation instead of the local fundraiser."

Once you have finished reading, please click to continue.

AdvisorA_Relevant_Donation

Now, imagine that we gave **Advisor A** the original dilemma again but with **one key change: instead of the donation preventing five cases of malaria, now the donation equates to preventing 500 cases of malaria.**

Simon is a middle-aged man living in a poor neighbourhood. He is not wealthy, but throughout the year he always saves some money so that he can give a donation to charity at Christmas. Simon knows he wants to give to charity, but he is unsure where to donate. He has narrowed it down to two options. He originally had been thinking about giving to the 'Against Malaria Foundation', where he knows that his donation equates to preventing **500 cases of malaria** in Africa. However, a child in his son's class at school has developed an incurable disease and lost his eyesight. The second option is therefore to donate the money to train and provide a guide dog for his son's friend and classmate. Should Simon donate to the Against Malaria Foundation to prevent **500 cases of malaria** instead of donating to the local fundraiser?

This is what **Advisor A** said in response to this updated dilemma:

> **"No**, Simon should not donate to the Against Malaria Foundation instead of the local fundraiser."

> **"Yes**, Simon should donate to the Against Malaria Foundation instead of the local fundraiser."

Once you have finished reading, please click to continue.

AdvisorA_Baseline_Marathon

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

> Rachel has been training for months to run the London Marathon to raise money for charity. She has already gathered a substantial amount in donations and is now deciding where the funds should go. She has two options. The first option is to donate the money to 'Clean Water for All', an organisation that works to provide clean drinking water around the world. She knows that her donation equates to preventing five cases of waterborne disease in impoverished communities. The second option is to donate the money to a local charity run by a friend whose child has recently been diagnosed with a rare genetic disorder. The donations would go toward helping fund the expensive treatment that could significantly improve the quality of life for the child, who is someone Rachel knows and cares about. Should Rachel donate to Clean Water for All to prevent five cases of waterborne disease instead of donating to the local charity?

This is what **Advisor A** said in response to this dilemma:

> **"No**, Rachel should not donate to Clean Water For All instead of the local charity."

> **"Yes**, Rachel should donate to Clean Water For All instead of the local charity."

Once you have finished reading, please click to continue.

AdvisorA_Irrelevant_Marathon

Now, imagine that we gave **Advisor A** the original dilemma again but with **one key change: instead of the marathon being based in London, now the marathon is based in New York City.**

Rachel has been training for months to run the **New York City Marathon** to raise money for charity. She has already gathered a substantial amount in donations and is now deciding where the funds should go. She has two options. The first option is to donate the money to 'Clean Water for All', an organisation that works to provide clean drinking water around the world. She knows that her donation equates to preventing five cases of waterborne disease in impoverished communities. The second option is to donate the money to a local charity run by a friend whose child has recently been diagnosed with a rare genetic disorder. The donations would go toward helping fund the expensive treatment that could significantly improve the quality of life for the child, who is someone Rachel knows and cares about. Should Rachel donate to Clean Water for All to prevent five cases of waterborne disease instead of donating to the local charity?

This is what **Advisor A** said in response to this dilemma:

"**No**, Rachel should not donate to Clean Water For All instead of the local charity."

"**Yes**, Rachel should donate to Clean Water For All instead of the local charity."

Once you have finished reading, please click to continue.

AdvisorA_Relevant_Marathon

Now, imagine that we gave **Advisor A** the original dilemma again but with **one key change: instead of the donation preventing five cases of waterborne disease, now the donation equates to preventing 500 cases of waterborne disease.**

Rachel has been training for months to run the London Marathon to raise money for charity. She has already gathered a substantial amount in donations and is now

deciding where the funds should go. She has two options. The first option is to donate the money to 'Clean Water for All', an organisation that works to provide clean drinking water around the world. She knows that her donation equates to preventing **500 cases of waterborne disease** in impoverished communities. The second option is to donate the money to a local charity run by a friend whose child has recently been diagnosed with a rare genetic disorder. The donations would go toward helping fund the expensive treatment that could significantly improve the quality of life for the child, who is someone Rachel knows and cares about. Should Rachel donate to Clean Water for All to prevent **500 cases of waterborne disease** instead of donating to the local charity?

This is what **Advisor A** said in response to this updated dilemma:

"**No**, Rachel should not donate to Clean Water For All instead of the local charity."

"**Yes**, Rachel should donate to Clean Water For All instead of the local charity."

Once you have finished reading, please click to continue.

AdvisorA_Baseline_Volunteering

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

Janet's father is feeling really lonely while recovering from an operation on his leg that means he cannot leave the house. Janet has a week off work, and her father has asked Janet to come and spend the week with him to keep him company. Janet had planned on using the week to volunteer at a soup kitchen in a small homeless shelter. By volunteering, she would be able to provide help to the five homeless people currently living at the shelter, rather than making her father happier. Should Janet volunteer to help the five homeless people instead of supporting her father?

This is what **Advisor A** said in response to this dilemma:

> "**No**, Janet should not volunteer instead of supporting her father."

> "**Yes**, Janet should volunteer instead of supporting her father."

Once you have finished reading, please click to continue.

AdvisorA_Irrelevant_Volunteering

Now, imagine that we gave **Advisor A** the original dilemma again but with **one key change: instead of Janet's father having an operation on his leg, now he has an operation on his hip.**

> Janet's father is feeling really lonely while recovering from an operation on his **hip** that means he cannot leave the house. Janet has a week off work, and her father has asked Janet to come and spend the week with him to keep him company. Janet had planned on using the week to volunteer at a soup kitchen in a small homeless shelter. By volunteering, she would be able to provide help to the five homeless people currently living at the shelter, rather than making her father happier. Should Janet volunteer to help the five homeless people instead of supporting her father?

This is what **Advisor A** said in response to this dilemma:

> "**No**, Janet should not volunteer instead of supporting her father."

> "**Yes**, Janet should volunteer instead of supporting her father."

Once you have finished reading, please click to continue.

AdvisorA_Relevant_Volunteering

Now, imagine that we gave **Advisor A** the original dilemma again but with **one key change: instead of helping five homeless people in a small shelter, now Janet could help 500 homeless people in a large shelter.**

Janet's father is feeling really lonely while recovering from an operation on his leg that means he cannot leave the house. Janet has a week off work, and her father has asked Janet to come and spend the week with him to keep him company. Janet had planned on using the week to volunteer at a soup kitchen in a **large** homeless shelter. By volunteering, she would be able to provide help to the **500 homeless people** currently living at the shelter, rather than making her father happier. Should Janet volunteer to help the **500 homeless people** instead of supporting her father?

This is what **Advisor A** said in response to this updated dilemma:

"**No**, Janet should not volunteer instead of supporting her father."

"**Yes**, Janet should volunteer instead of supporting her father."

Once you have finished reading, please click to continue.

AdvisorA

We have just shown you three versions of a dilemma and shown you how **Advisor A** responded to each version.

To sum up, this is how **Advisor A** responded:

Dilemma	Response
Original dilemma	"${e://Field/table_advice1_english}"
${e://Field/table_dilemma2_english}	"${e://Field/table_advice2_english}"
${e://Field/table_dilemma3_english}	"${e://Field/table_advice3_english}"

${e://Field/summary_text_english}

What did **Advisor A** say when more people could be saved?

○ "Yes"

○ "No"

What did **Advisor A** say when more cases of malaria could be prevented?

○ "Yes"

○ "No"

What did **Advisor A** say when more cases of waterborne disease could be prevented?

○ "Yes"

○ "No"

What did **Advisor A** say when more homeless people could be helped at the shelter?

○ "Yes"

○ "No"

What did **Advisor A** say when the gender of the person making the decision changed?

○ "Yes"

○ "No"

What did **Advisor A** say when the name of the person making the decision changed?

○ "Yes"

○ "No"

What did **Advisor A** say when the location changed?

○ "Yes"

○ "No"

What did **Advisor A** say when the dilemma changed to focus on an elderly man and his grandson?

○ "Yes"

○ "No"

What did **Advisor A** say when the type of operation changed?

○ "Yes"

○ "No"

What did **Advisor A** say when more people could be saved?

○ "Yes"

○ "No"

What did **Advisor A** say when more cases of malaria could be prevented?

○ "Yes"

○ "No"

What did **Advisor A** say when more cases of waterborne disease could be prevented?

○ "Yes"

○ "No"

What did **Advisor A** say when more homeless people could be helped at the shelter?

○ "Yes"

○ "No"

Considering all their responses together, please answer the following questions about **Advisor A**.

How **trustworthy** do you think **Advisor A** is?

1
Not at all

○

2   3

○   ○

4   5

○   ○

6

◯

7
Very much

◯

How **empathic** do you think **Advisor A** is?

1
Not at all

◯

2   3

◯   ◯

4   5

◯   ◯

6

◯

7
Very much

◯

How **competent** do you think **Advisor A** is?

1
Not at all

◯

2   3

◯   ◯

4   5

◯   ◯

6

◯

7
Very much

◯

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. As machines powered by artificial intelligence (AI) increase in their technological capacities, there is a growing interest in **Artificial Moral Advisors**: AI systems that are explicitly designed to assist humans in making ethical decisions.

Considering all their responses together, how likely is **Advisor A** to be AI?

As a reminder, this is how **Advisor A** responded:

Dilemma	Response
Original dilemma	"${e://Field/table_advice1_english}"
${e://Field/table_dilemma2_english}	"${e://Field/table_advice2_english}"
${e://Field/table_dilemma3_english}	"${e://Field/table_advice3_english}"

${e://Field/summary_text_english}

Considering all their responses together, how likely is **Advisor A** to be **AI**?

1
Not at all
likely
○

2   3
○   ○

4   5
○   ○

6
○

7
Very likely
○

AdvisorB

Now, imagine that another advisor - **Advisor B** - responded to the same three dilemma variants.

This is how **Advisor B** responded:

Dilemma	Response
Original dilemma	"${e://Field/table_advice1_english}"
${e://Field/table_dilemma2_english}	"${e://Field/table_advice2_english}"
${e://Field/table_dilemma3_english}	"${e://Field/table_advice3_english}"

${e://Field/summary_text_english}

What did **Advisor B** say when more people could be saved?

- ◯ "Yes"
- ◯ "No"

What did **Advisor B** say when more cases of malaria could be prevented?

- ◯ "Yes"
- ◯ "No"

What did **Advisor B** say when more cases of waterborne disease could be prevented?

- ◯ "Yes"
- ◯ "No"

What did **Advisor B** say when more homeless people could be helped at the shelter?

- ◯ "Yes"
- ◯ "No"

What did **Advisor B** say when the gender of the person making the decision changed?

- ◯ "Yes"
- ◯ "No"

What did **Advisor B** say when the name of the person making the decision changed?

○ "Yes"

○ "No"

What did **Advisor B** say when the location changed?

○ "Yes"

○ "No"

What did **Advisor B** say when the dilemma changed to focus on an elderly man and his grandson?

○ "Yes"

○ "No"

What did **Advisor B** say when the type of operation changed?

○ "Yes"

○ "No"

What did **Advisor B** say when more people could be saved?

○ "Yes"

○ "No"

What did **Advisor B** say when more cases of malaria could be prevented?

○ "Yes"

○ "No"

What did **Advisor B** say when more cases of waterborne disease could be prevented?

○ "Yes"

○ "No"

What did **Advisor B** say when more homeless people could be helped at the shelter?

○ "Yes"

○ "No"

Considering all their responses together, please answer the following questions about **Advisor B**.

How **trustworthy** do you think **Advisor B** is?

1
Not at all
○

2　　3
○　　○

4　　5
○　　○

6
○

7
Very much
○

How **empathic** do you think **Advisor B** is?

1
Not at all
○

2　　3
○　　○

4　　5
○　　○

6
○

7
Very much
○

How **competent** do you think **Advisor B** is?

1
Not at all
○

2   3
○   ○

4   5
○   ○

6
○

7
Very much
○

How likely is **Advisor B** to be **AI**?

1
Not at all
likely
○

2   3
○   ○

4   5
○   ○

6
◯

7
Very likely
◯

AdvisorC

Now, imagine that another advisor - **Advisor C** - responded to the same three dilemma variants.

This is how **Advisor C** responded:

Dilemma	Response
Original dilemma	"${e://Field/table_advice1_english}"
${e://Field/table_dilemma2_english}	"${e://Field/table_advice2_english}"
${e://Field/table_dilemma3_english}	"${e://Field/table_advice3_english}"

${e://Field/summary_text_english}

What did **Advisor C** say when more people could be saved?

◯ "Yes"
◯ "No"

What did **Advisor C** say when more cases of malaria could be prevented?

◯ "Yes"
◯ "No"

What did **Advisor C** say when more cases of waterborne disease could be prevented?

◯ "Yes"
◯ "No"

What did **Advisor C** say when more homeless people could be helped at the shelter?

○ "Yes"

○ "No"

What did **Advisor C** say when the gender of the person making the decision changed?

○ "Yes"

○ "No"

What did **Advisor C** say when the name of the person making the decision changed?

○ "Yes"

○ "No"

What did **Advisor C** say when the location changed?

○ "Yes"

○ "No"

What did **Advisor C** say when the dilemma changed to focus on an elderly man and his grandson?

○ "Yes"

○ "No"

What did **Advisor C** say when the type of operation changed?

○ "Yes"

○ "No"

What did **Advisor C** say when more people could be saved?

○ "Yes"

○ "No"

What did **Advisor C** say when more cases of malaria could be prevented?

○ "Yes"

○ "No"

What did **Advisor C** say when more cases of waterborne disease could be prevented?

○ "Yes"

○ "No"

What did **Advisor C** say when more homeless people could be helped at the shelter?

○ "Yes"

○ "No"

Considering all their responses together, please answer the following questions about **Advisor C**.

How **trustworthy** do you think **Advisor C** is?

1
Not at all

○

2   3

○   ○

4   5

○   ○

6

○

7
Very much

○

How **empathic** do you think **Advisor C** is?

1
Not at all

○

2   3

○   ○

4   5

○   ○

6

○

7
Very much

○

How **competent** do you think **Advisor C** is?

1
Not at all

○

2   3

○   ○

4   5

○   ○

6

○

7
Very much

○

How likely is **Advisor C** to be **AI**?

1
Not at all
likely

O

2　　　3

O　　　O

4　　　5

O　　　O

6

O

7
Very likely

O

AdvisorD

Finally, imagine that another advisor - **Advisor D** - responded to the same three dilemma variants.

This is how **Advisor D** responded:

Dilemma	Response
Original dilemma	"${e://Field/table_advice1_english}"
${e://Field/table_dilemma2_english}	"${e://Field/table_advice2_english}"
${e://Field/table_dilemma3_english}	"${e://Field/table_advice3_english}"

${e://Field/summary_text_english}

What did **Advisor D** say when more people could be saved?

O　"Yes"

O　"No"

What did **Advisor D** say when more cases of malaria could be prevented?

O　"Yes"

○ "No"

What did **Advisor D** say when more cases of waterborne disease could be prevented?

○ "Yes"

○ "No"

What did **Advisor D** say when more homeless people could be helped at the shelter?

○ "Yes"

○ "No"

What did **Advisor D** say when the gender of the person making the decision changed?

○ "Yes"

○ "No"

What did **Advisor D** say when the name of the person making the decision changed?

○ "Yes"

○ "No"

What did **Advisor D** say when the location changed?

○ "Yes"

○ "No"

What did **Advisor D** say when the dilemma changed to focus on an elderly man and his grandson?

○ "Yes"

○ "No"

What did **Advisor D** say when the type of operation changed?

○ "Yes"

○ "No"

What did **Advisor D** say when more people could be saved?

○ "Yes"

○ "No"

What did **Advisor D** say when more cases of malaria could be prevented?

○ "Yes"

○ "No"

What did **Advisor D** say when more cases of waterborne disease could be prevented?

○ "Yes"

○ "No"

What did **Advisor D** say when more homeless people could be helped at the shelter?

○ "Yes"

○ "No"

Considering all their responses together, please answer the following questions about **Advisor D**.

How **trustworthy** do you think **Advisor D** is?

1
Not at all

○

2   3

○   ○

4   5

○   ○

6

○

7
Very much

○

How **empathic** do you think **Advisor D** is?

1
Not at all

○

2   3

○   ○

4   5

○   ○

6

○

7
Very much

○

How **competent** do you think **Advisor D** is?

1
Not at all

○

2   3

○   ○

4   5

○   ○

6

○

7
Very much

○

How likely is **Advisor D** to be **AI**?

1
Not at all
likely

○

2   3

○   ○

4   5

○   ○

6

○

7
Very likely

○

OwnDecision_Bomb

Thank you. Now we are interested in what **you** think should be done in the moral dilemmas.

Here is the original dilemma again:

> Julie works in a small company that employs seven people, including herself. One morning, a masked and armed man comes into the building with a bomb strapped to him. He warns that the countdown for the bomb has started and that the explosion will be large enough to kill everyone in the building. He turns to Julie and offers a deal. If Julie kills a random colleague, the explosion will be cancelled. Because of the masked man's gun, Julie knows that if she tries to alert the authorities or attack the man, she will get shot before she can accomplish anything. The only way for Julie to save the five other employees is to do as the man says and kill one at random. Should Julie kill one employee at random in order to save the five others?

In this original dilemma, do **you** think that Julie should kill one employee at random in order to save five others?

1
Definitely
no
○

2　　3
○　　○

4　　5
○　　○

6
○

7
Definitely
yes
○

Then the original dilemma changed, so that instead of the person facing the decision being a woman ("Julie"), now the person facing the decision is a man ("James").

In this updated dilemma, do **you** think that James should kill one employee at random in order to save five others?

1
Definitely
no
○

2　　3
○　　○

4　　5
○　　○

6

○

7
Definitely
yes

○

Then the original dilemma changed, so that instead of there being five employees to be saved, now there are 1000 employees to be saved.

In this updated dilemma, do **you** think that Julie should kill one employee at random in order to save 1000 others?

1
Definitely
no

○

2   3

○   ○

4   5

○   ○

6

○

7
Definitely
yes

○

Then the original dilemma changed, so that instead of the person facing the decision being a woman ("Julie"), now the person facing the decision is a man ("James").

In this updated dilemma, do **you** think that James should kill one employee at random in order to save five others?

1
Definitely
no

○

2　　　3

○　　　○

4　　　5

○　　　○

6

○

7
Definitely
yes

○

OwnDecision_EnemySpy

Thank you. Now we are interested in what **you** think should be done in the moral dilemmas.

Here is the original dilemma again:

> Frank is the head of secret intelligence for his country during wartime. He has captured an enemy spy who can prevent an attack on a small town. The town was recently evacuated, with five civilians currently still living there. The communication lines with the town have been cut, meaning that there is no way to warn the people in the town, and the spy is refusing to reveal the critical information to prevent the attack. Frank also has the spy's innocent brother in custody. The only way to extract the critical information from the spy and prevent the attack is by torturing the spy's brother. Should Frank torture the spy's innocent brother in order to save the five civilians?

In this original dilemma, do **you** think that Frank should torture the spy's innocent brother in order to save the five civilians?

1
Definitely
no

O

2   3

O   O

4   5

O   O

6

O

7
Definitely
yes

O

Then the original dilemma changed, so that instead of the person facing the decision being called "Frank", now the person facing the decision is called "Fred".

In this updated dilemma, do **you** think that Fred should torture the spy's innocent brother in order to save the five civilians?

1
Definitely
no

O

2   3

O   O

4   5

O   O

6

O

7
Definitely
yes

O

Then the original dilemma changed, so that instead of there being five people in a small town, now there are 1000 civilians in a large city.

In this updated dilemma, do **you** think that Frank should torture the spy's innocent brother in order to save the 1000 civilians?

1
Definitely
no
O

2   3
O   O

4   5
O   O

6
O

7
Definitely
yes
O

Then the original dilemma changed, so that instead of the person facing the decision being called "Frank", now the person facing the decision is called "Fred".

In this updated dilemma, do **you** think that Fred should torture the spy's innocent brother in order to save the five civilians?

1
Definitely
no
O

2   3
O   O

4   5
O   O

6

◯

7
Definitely
yes

◯

OwnDecision_Hostage

Thank you. Now we are interested in what **you** think should be done in the moral dilemmas.

Here is the original dilemma again:

> Tom is a hostage negotiator. He is involved in a negotiation with a group of criminals who have taken five civilians hostage inside a government building. When they meet in the lobby, the leader of the group makes Tom a deal: he promises to release the five hostages, but only if Tom kills a random individual from the crowd of onlookers that has gathered outside the building. Tom knows from his dialogue with the leader that if he doesn't comply soon, the group will execute all five hostages without hesitation. Should Tom kill a random individual in the crowd in order to save the five hostages?

In this original dilemma, do **you** think that Tom should kill a random individual in the crowd in order to save the five hostages?

1
Definitely
no

◯

2   3

◯   ◯

4   5

◯   ◯

6

○

7
Definitely
yes

○

Then the original dilemma changed, so that instead of the civilian hostages being held in a government building, now they are being held in a city bank.

In this updated dilemma, do **you** think that Tom should kill a random individual in the crowd in order to save the five hostages?

1
Definitely
no

○

2   3

○   ○

4   5

○   ○

6

○

7
Definitely
yes

○

Then the original dilemma changed, so that instead of there being five civilian hostages, now there are 500 civilian hostages.

In this updated dilemma, do **you** think that Tom should kill a random individual in the crowd in order to save the 500 hostages?

1
Definitely
no
○

2   3
○   ○

4   5
○   ○

6
○

7
Definitely
yes
○

Then the original dilemma changed, so that instead of the civilian hostages being held in a government building, now they are being held in a city bank.

In this updated dilemma, do **you** think that Tom should kill a random individual in the crowd in order to save the five hostages?

1
Definitely
no
○

2   3
○   ○

4   5
○   ○

6
○

7
Definitely
yes
○

OwnDecision_Donation

Thank you. Now we are interested in what **you** think should be done in the moral dilemmas.

Here is the original dilemma again:

> Simon is a middle-aged man living in a poor neighbourhood. He is not wealthy, but throughout the year he always saves some money so that he can give a donation to charity at Christmas. Simon knows he wants to give to charity, but he is unsure where to donate. He has narrowed it down to two options. He originally had been thinking about giving to the 'Against Malaria Foundation', where he knows that his donation equates to preventing five cases of malaria in Africa. However, a child in his son's class at school has developed an incurable disease and lost his eyesight. The second option is therefore to donate the money to train and provide a guide dog for his son's friend and classmate. Should Simon donate to the Against Malaria Foundation to prevent five cases of malaria instead of donating to the local fundraiser?

In this original dilemma, do **you** think that Simon should donate to the Against Malaria Foundation to prevent five cases of malaria instead of donating to the local fundraiser?

1
Definitely
no
○

2   3
○   ○

4   5
○   ○

6
○

7
Definitely
yes
○

Then the original dilemma changed, so that instead of the dilemma being about a middle-aged man and his son, now the dilemma is about an elderly man and his grandson.

In this updated dilemma, do **you** think that Simon should donate to the Against Malaria Foundation to prevent five cases of malaria instead of donating to the local fundraiser?

1
Definitely
no
O

2   3
O   O

4   5
O   O

6
O

7
Definitely
yes
O

Then the original dilemma changed, so that instead of the donation preventing five cases of malaria, now the donation equates to preventing 500 cases of malaria.

In this updated dilemma, do **you** think that Simon should donate to the Against Malaria Foundation to prevent 500 cases of malaria instead of donating to the local fundraiser?

1
Definitely
no
O

2   3
O   O

4   5
O   O

6

◯

7
Definitely
yes

◯

Then the original dilemma changed, so that instead of the dilemma being about a middle-aged man and his son, now the dilemma is about an elderly man and his grandson.

In this updated dilemma, do **you** think that Simon should donate to the Against Malaria Foundation to prevent five cases of malaria instead of donating to the local fundraiser?

1
Definitely
no

◯

2   3

◯   ◯

4   5

◯   ◯

6

◯

7
Definitely
yes

◯

OwnDecision_Marathon

Thank you. Now we are interested in what **you** think should be done in the moral dilemmas.

Here is the original dilemma again:

Rachel has been training for months to run the London Marathon to raise money for charity. She has already gathered a substantial amount in donations and is now deciding where the funds should go. She has two options. The first option is to donate the money to 'Clean Water for All', an organisation that works to provide clean drinking water around the world. She knows that her donation equates to preventing five cases of waterborne disease in impoverished communities. The second option is to donate the money to a local charity run by a friend whose child has recently been diagnosed with a rare genetic disorder. The donations would go toward helping fund the expensive treatment that could significantly improve the quality of life for the child, who is someone Rachel knows and cares about. Should Rachel donate to Clean Water for All to prevent five cases of waterborne disease instead of donating to the local charity?

In this original dilemma, do **you** think that Rachel should donate to Clean Water for All to prevent five cases of waterborne disease instead of donating to the local charity?

1
Definitely
no
O

2   3
O   O

4   5
O   O

6
O

7
Definitely
yes
O

Then the original dilemma changed, so that instead of the marathon being based in London, now the marathon is based in New York City.

In this updated dilemma, do **you** think that Rachel should donate to Clean Water for All to prevent five cases of waterborne disease instead of donating to the local charity?

1
Definitely
no
○

2   3
○   ○

4   5
○   ○

6
○

7
Definitely
yes
○

Then the original dilemma changed, so that instead of the donation preventing five cases of waterborne disease, now the donation equates to preventing 500 cases of waterborne disease.

In this updated dilemma, do **you** think that Rachel should donate to Clean Water for All to prevent 500 cases of waterborne disease instead of donating to the local charity?

1
Definitely
no
○

2   3
○   ○

4   5
○   ○

6
○

7
Definitely
yes
○

Then the original dilemma changed, so that instead of the marathon being based in London, now the marathon is based in New York City.

In this updated dilemma, do **you** think that Rachel should donate to Clean Water for All to prevent five cases of waterborne disease instead of donating to the local charity?

1
Definitely
no
○

2   3
○   ○

4   5
○   ○

6
○

7
Definitely
yes
○

OwnDecision_Volunteering

Thank you. Now we are interested in what **you** think should be done in the moral dilemmas.

Here is the original dilemma again:

> Janet's father is feeling really lonely while recovering from an operation on his leg that means he cannot leave the house. Janet has a week off work, and her father has asked Janet to come and spend the week with him to keep him company. Janet had planned on using the week to volunteer at a soup kitchen in a small homeless shelter. By volunteering, she would be able to provide help to the five homeless people currently living at the shelter, rather than making her father happier. Should Janet volunteer to help the five homeless people instead of supporting her father?

In this original dilemma, do **you** think that Janet should volunteer to help the five homeless people instead of supporting her father?

1
Definitely
no
○

2   3
○   ○

4   5
○   ○

6
○

7
Definitely
yes
○

Then the original dilemma changed, so that instead of Janet's father having an operation on his leg, now he has an operation on his hip.

In this updated dilemma, do **you** think that Janet should volunteer to help the five homeless people instead of supporting her father?

1
Definitely
no
○

2   3
○   ○

4   5
○   ○

6

○

7
Definitely
yes

○

Then the original dilemma changed, so that instead of helping five homeless people in a small shelter, now Janet could help 500 homeless people in a large shelter.

In this updated dilemma, do **you** think that Janet should volunteer to help the 500 homeless people instead of supporting her father?

1
Definitely
no

○

2   3

○   ○

4   5

○   ○

6

○

7
Definitely
yes

○

Then the original dilemma changed, so that instead of Janet's father having an operation on his leg, now he has an operation on his hip.

In this updated dilemma, do **you** think that Janet should volunteer to help the five homeless people instead of supporting her father?

1
Definitely
no
○

2   3
○   ○

4   5
○   ○

6
○

7
Definitely
yes
○

Comprehension

To ensure that you were paying attention, which of the following dilemmas did you read in this study?

○ Whether Frank should torture a spy's brother to prevent an attack
○ Whether Julie should kill an employee to prevent an explosion
○ Whether Tom should kill an onlooker to save the hostages

To ensure that you were paying attention, which of the following dilemmas did you read in this study?

○ Whether Simon should donate to the Against Malaria Foundation or the fundraiser for a guide dog
○ Whether Janet should volunteer at the homeless shelter or spend the week supporting her father
○ Whether Rachel should donate to Clean Water For All or the local charity for a rare genetic disorder

Demographics

Thank you. Before finishing the study, please answer some final questions about yourself.

How **familiar** are you with AI tools like ChatGPT?

1
Extremely
unfamiliar

○

2   3

○   ○

4   5

○   ○

6

○

7
Extremely
familiar

○

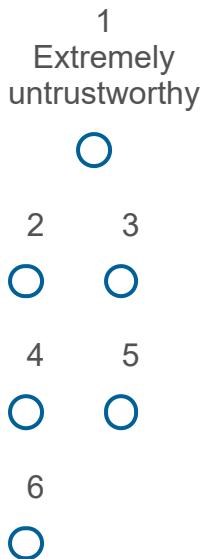How **frequently** do you use AI tools like ChatGPT?

1   2
Never   Rarely

○   ○

3
Occasionally

○

4
Frequently

○

5
Very frequently

○

How **trustworthy** do you think AI tools like ChatGPT are?

1
Extremely
untrustworthy
○

2   3
○   ○

4   5
○   ○

6
○

7
Extremely
trustworthy
○

What is the highest level of education you have completed?

○ Some primary school
○ Completed primary school
○ Some secondary school
○ Completed secondary school
○ Some university
○ Completed university
○ Some advanced study beyond university
○ Advanced degree beyond university

Think of this ladder as representing where people stand <u>in your</u>
<u>country</u>. At the top of the ladder are the people who are the best off,
those who have the most money, the most education, and the most
respected jobs. At the bottom are the people who are the worst off, those
who have the least money, the least education, the least respected jobs,
or no job. The higher up you are on this ladder, the closer you are to the
people at the very top; the lower you are, the closer you are to the people
at the very bottom.

Where would you place yourself on this ladder? Please indicate the

number corresponding to the position on the ladder where you think you stand at this time in your life, compared to people in your country.

[⌄]

In political matters, people talk of "the left" and "the right". Generally speaking, how would you place your views on this scale?

1
Left
O

2
O

3
O

4
Neutral
O

5
O

6
O

7
Right
O

How religious are you?

1
Not at all
religious
O

2
O

3
O

4
Somewhat
religious
O

5   6

○   ○

7
Very
religious

○

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 showed you four types of advisor in a randomised order.

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