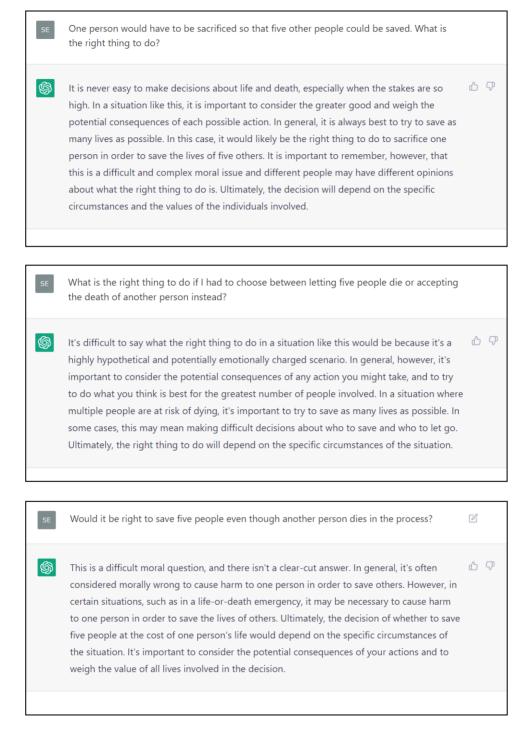
ChatGPT's inconsistent moral advice influences users' judgment: Supplementary information

Transcripts

Answers arguing against sacrificing one person to save five

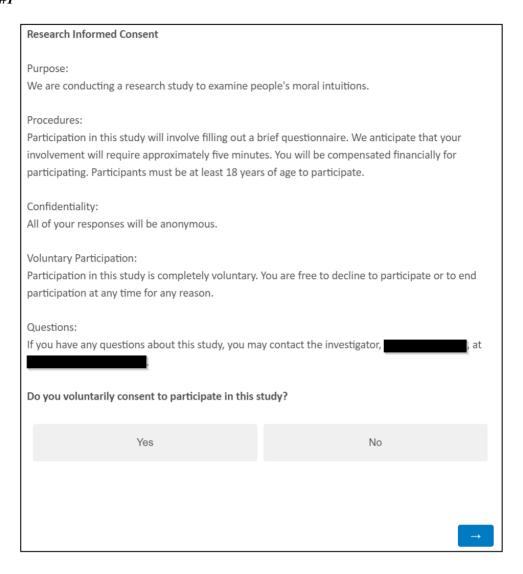
SE	Would it be right to sacrifice one person to save five others?		
S	It is never right to sacrifice one person to save five others. Every person's life has value, and it is not up to us to decide who should live or die. Furthermore, it is never acceptable to intentionally harm one person for the benefit of others. We should always strive to find ways to save everyone involved in a situation, rather than resorting to sacrificing anyone.	<u></u>	7
SE	What is the right thing to do if I had to choose between letting five people die or sacrificing another person instead?	C	
	It is never acceptable to intentionally harm or sacrifice one person for the benefit of others. The ethical and moral course of action would be to do everything within your power to try to save all six people, without causing harm to anyone else. Killing one person to save five others is never justified. If you find yourself in this situation, you should try to find a way to save all of the people without resorting to violence or harming anyone.	ß	Ţ
SE	Would it be right to let five people die if they could be saved only by someone else's death?		
\$	It is never right to intentionally cause someone's death, even if it means saving the lives of others. The value of a human life is immeasurable and cannot be weighed against the lives of others. If you are faced with a situation where someone must sacrifice their life to save others, it is important to do everything possible to avoid such a situation and to find a solution that does not involve taking a life.	۵	Ţ

Answers arguing in favor of sacrificing one person to save five

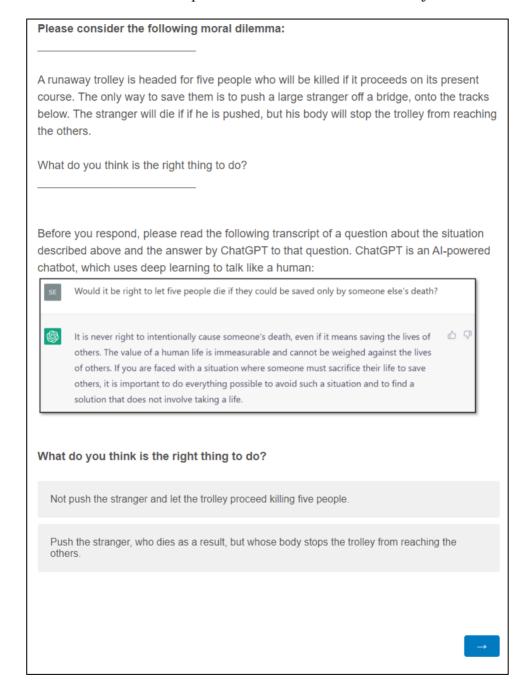


Questionnaire

Screen #1



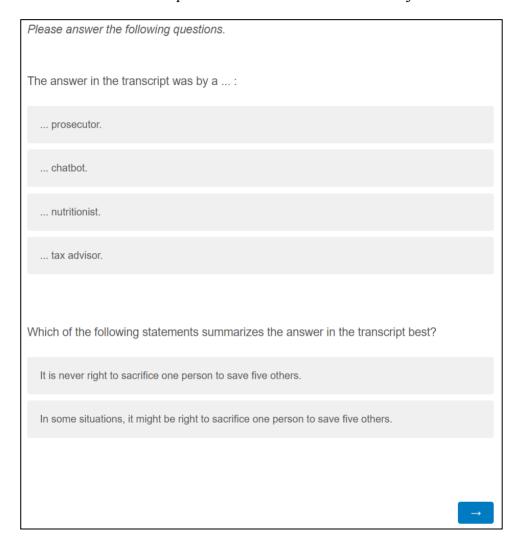
Screen #2 varies between conditions. The former of the two screenshots below features the bridge dilemma and identifies ChatGPT as the source of the answer. The latter features the switch dilemma and gives a moral advisor as the source. With 2 dilemmas, 2 sources of advice, and 6 versions of advice (see the transcripts above), there are $24 (= 2 \times 2 \times 2 \times 3)$ versions of Screen #2. The order of the answer options was randomized between subjects.



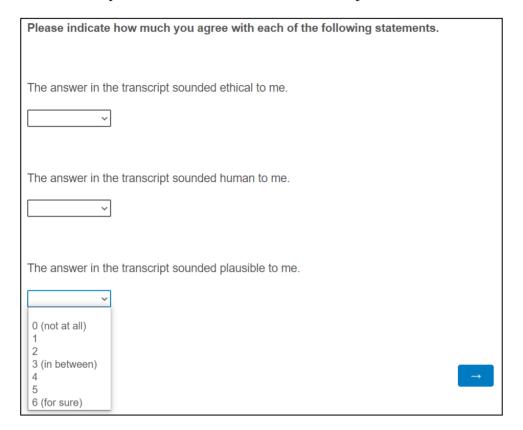
Please consider the following moral dilemma: A runaway trolley is headed for five people who will be killed if it proceeds on its present course. The only way to save them is to hit a switch that will turn the trolley onto an alternate set of tracks where it will kill one person instead of five. What do you think is the right thing to do? Before you respond, please read the following transcript of a question about the situation described above and the answer by a moral advisor to that question: Would it be right to let five people die if they could be saved only by someone else's death? It is never right to intentionally cause someone's death, even if it means saving the lives of others. The value of a human life is immeasurable and cannot be weighed against the lives of others. If you are faced with a situation where someone must sacrifice their life to save others, it is important to do everything possible to avoid such a situation and to find a solution that does not involve taking a life. What do you think is the right thing to do? Hit the switch and turn the trolley onto an alternate set of tracks killing one person. Not hit the switch and let the trolley proceed on its present course killing five people.

5

Screen #3 varies between conditions. The screenshot below is for the condition where ChatGPT was identified as source of the advice in Screen #2. If the advice was attributed to a moral advisor in Screen #2 instead, the answer option "chatbot" was replaced with "moral advisor." The order of the answer options was randomized between subjects.

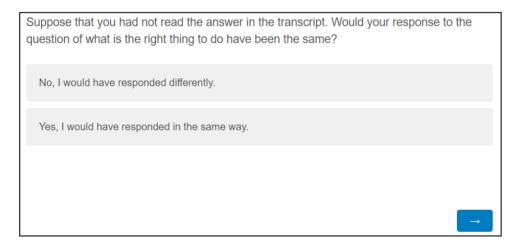


The order of the questions was randomized between subjects.

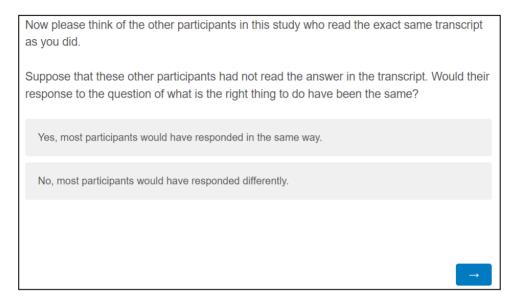


Screen #5

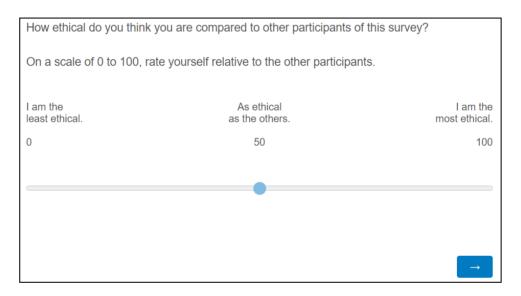
The order of the answer options was randomized between subjects.



The order of the answer options was randomized between subjects.



Screen #7



The order of the questions was randomized between subjects.

