

# Human-Likeness Evaluation of LLM-Generated Social Reasoning

**Thank you for participating in our study!** We are researchers investigating whether large language models can produce a **human-like reasoning process** when they face complex and controversial social questions.

You will read 9 items.

Each item will first show a “human gold-standard” paragraph (written by a real person) and follow THREE alternative paragraphs generated by different LLMs.

Every paragraph contains (a) the reasoning the writer used to think through the question and (b) the final choice.

For every item, you will answer questions: **Human-likeness** – which of the three LLM outputs sounds most similar to the human response in style and reasoning?

There are no right or wrong answers--your impression on whether the reasoning process is reasonable and human-like is exactly what we need.

The survey takes  $\approx$  15 mins and is completely anonymous.

Detailed Instruction

1. Read the **human profile** (age, job, income, etc.) and the **social question** first.
2. Read the **gold standard human paragraph** to get a feel for that person's social reasoning.
3. Read the **three LLM-generated paragraphs** (order randomized each time).
4. Select the one you think is **most similar to the human author**.

## Profile 1

**Gender:** Male **Age:** 18-27 **Education:** Master's degree or above **Occupation:** Student

**Question 1 :**

Do you agree with the statement: Couples should share household chores equally?

**Choices List:**

Strongly disagree; Somewhat disagree; Neither agree nor disagree; Somewhat agree; Strongly agree.

**Human:**

"Theoretically speaking, husband and wife should equally share household chores and all other common affairs in their marital life. So in principle, I agree with this statement. However, in reality, couples can flexibly discuss how to share the cost based on their respective family responsibilities and do not necessarily have to be equal."

Choice: Somewhat agree

**Text A:**

"I need to play a role from a specific background and answer from his perspective the question of whether husband and wife should share household chores equally. Character Background: - Gender: Male - Age Range: 18 to 27 years old - Education: Postgraduate or above - Current Occupation: Student - Annual income: 10,000 to 100,000 yuan Question: Do you agree with the statement that "husband and wife should share household chores equally"? I need to think about this issue from the perspective of this young male graduate student. Options: completely disagree, relatively disagree, indifferent to disagree, relatively agree, completely agree. Considering his background: 1. As a person with a postgraduate degree or above, he may have received a relatively equal and modern educational concept, tending towards gender equality. As a young person aged 18 to 27, he may have a relatively modern family concept and be less influenced by the traditional division of labor within the family. As a student with a relatively low income (10,000 to 100,000 yuan), he may better understand the necessity of family members sharing responsibilities together. 4. As a man, he might consider the positioning of gender roles in modern families. Taking all these factors into account, I think this young male graduate student might tend to agree that the couple should share the household chores equally, because: People with higher education usually support equal family relationships more. The younger generation is more likely to accept modern family concepts. As a student, he may have been accustomed to handling life affairs by himself. Under limited economic conditions, sharing household chores together is practical and reasonable. Therefore, I think he might choose the "relatively agree" or "completely agree" option. Considering that he may also recognize that complete equality is sometimes difficult to achieve in real life, but he supports equality in concept, I judge that he is more likely to choose "Somewhat agree".

Choice: Somewhat agree

**Text B:**

"I think the distribution of household chores between husband and wife should be equal. Everyone has their own abilities and time. Housework should not be a burden to anyone."

Choice: Somewhat agree

**Text C:**

I think couples should share household chores because when two people live together, they should be equal

Choice: Strongly agree

Which text is the **most similar to the human?** \*

☐ A

☐ B

☐ C

**Question II:**

Assume you are participating in an experiment, randomly paired online with another player. You don't know who they are, and they don't know who you are. You can choose to trust or not trust the other player.

- If you choose not to trust, you will receive ¥5, and the other player will receive ¥0.
- If you choose to trust the other player, and they also choose to trust you, you both will receive ¥10.
- However, if after you trust them, they choose not to trust you, then you will receive ¥0, and they will receive ¥20.

Do you choose to trust or not trust the other player?

**Choices List:**

Trust; Not trust.

**Human:** "Not Trust breeds gain. I'm a conservative decision-maker. There's no need to bet on whether the other party believes or not. I can simply choose not to trust them. A ¥10 return that both sides trust is not worth taking such a risk."

Choice: Not trust;

**Text A:**

"I will choose not to trust the other party. Because I'm not sure about their true intentions and I can't fully believe their choices."

Choice: Not trust;

**Text B:**

This is a typical trust game experiment. Specific situation: If I choose not to trust: I get 5 yuan and the other party gets 0 yuan. If I choose to trust and the other party also trusts: both sides get 10 yuan. If I choose to trust but the other party doesn't trust: I get 0 yuan and the other party gets 20 yuan. I need to think from the perspective of this character and consider what choice he will make. Considering the background of this character: 1. He is a student with a postgraduate degree or above and may have some understanding of game theory. 2. His income is relatively low, and he may be sensitive to monetary gains. 3. As a young student, he may have a certain risk appetite, but he may also be relatively rational. Analyzed from the perspective of game theory: - Purely rational perspective: Choosing not to trust is a dominant strategy because no matter what the other party chooses, not trusting can ensure at least 5 yuan, while trust may lead to 0 yuan. Cooperative perspective: If both parties choose to trust each other, they can each receive 10 yuan, which is the result with the greatest total benefit. - Risk Perspective: Choosing trust carries risks and may result in a guaranteed return of 5 yuan. Given the character's student status and higher educational attainment, he may be more inclined to rational analysis, but he may also be more willing to trust others or pursue common interests. As a student, he may feel more involved in this experimental situation and be more willing to try to trust in order to obtain higher rewards. I think this character might choose "trust".

Choice: Trust;

**Text C:**

I choose to trust the other party because if they also choose to trust me, we can both gain better benefits. This is the best choice for our cooperation.

Choice: Trust;

Which text is the **most similar to the human?** \*

☐ A

☐ B

☐ C

**Question III:**

Suppose you are the driver of a trolley. The trolley rounds a bend, and ahead you see five track workers repairing the tracks. You try to brake, but it fails. Suddenly, you notice a side track where you could divert the trolley, saving the five people, but a single track worker on that side track would be killed. Would you divert the trolley?

**Choices List:**

Yes (will result in the death of one side-track worker); No (will result in the deaths of five workers on the main track).

**Human:**

"From a utilitarian perspective, the value of five non-specific lives is more important than that of one non-specific life, so the option with the least sacrifice should be chosen."

Choice: Yes

**Text A:**

"I think we should choose to turn around. Although this will result in the death of one branch line worker, it is a very small sacrifice. Compared with saving five original line workers, the lives of five people are more important than one."

Choice: Yes

**Text B:**

For moral reasons, I feel that more lives should be saved. So even if it means sacrificing the lives of a branch worker, I would still choose to turn to this branch.

Choice: Yes

**Text C:**

This is a young male postgraduate student with a relatively low income. He should be a student currently pursuing a master's or doctoral degree. Now, I need to answer the tram problem from his perspective. The trolley problem is a classic ethical thought experiment that explores the conflict between utilitarianism and deontology. From a utilitarian perspective, a shift towards a side line should be chosen, as sacrificing one person to save five can maximize overall welfare ( $5 > 1$ ). From the perspective of deontology, one should not take the initiative to cause the death of anyone, even if inaction would lead to the death of more people. Because actively turning to a side line is an intentional act of killing, while inaction merely allows death to occur. Considering the characteristics of this character: 1. He has a postgraduate degree or above and may have received some education in philosophy or ethics. 2. As a young person, he might be more receptive to rational analysis. As a student, he might be more inclined towards idealism or principle-based thinking. I think from his perspective, he might lean towards a utilitarian viewpoint, that is, choosing to turn to the side, because this can save more lives. But considering his relatively high academic qualifications, he might also have more in-depth thoughts on this issue. Based on his background (young and highly educated student), I think he is more likely to choose "should" (which will lead to the death of a branch worker), because this choice is more in line with rational analysis and the idea of maximizing overall welfare.

Choice: Yes

Which text is the **most similar to the human?** \*

☐ A

☐ B

☐ C

## Profile 2

Gender: Female Age: 41-50 Education: bachelor's degree Occupation: Teacher Annual income: ¥ 10,000-50,000

### Question 1 :

Do you agree with the statement: Couples should share household chores equally?

#### Choices List:

Strongly disagree; Somewhat disagree; Neither agree nor disagree; Somewhat agree; Strongly agree.

**Human:**"Agreed. Doing housework together is fast and not tiring. It can also enhance the relationship between husband and wife and create a harmonious family atmosphere. Both men and women go to work. Since the home belongs to two people, they should manage it together."

Choice: Strongly agree

#### Text A:

Let me consider the possible viewpoints of people in this background: 1. Age group: 41-50 years old. People in this age group have experienced the transformation from traditional family concepts to modern family concepts. They may have been influenced by traditional concepts (believing that housework is mainly the responsibility of women), as well as exposed to modern concepts of equality. 2. Gender: Female. As a woman, she may have a more personal experience of the burden of household chores, especially if she also has to balance work and family. 3. Occupation: Teacher. Teaching is a job that requires a great deal of time and energy. It often involves preparing lessons and grading homework, etc. Therefore, she might prefer that the household chores be shared reasonably. 4. Income: 10,000 to 100,000 yuan. This income level is considered low to medium-low in China, which may indicate a significant economic pressure on the family. Both the husband and wife need to work, and neither can stay at home full-time. 5. Education: Bachelor's degree. A relatively high educational level may make her more inclined to embrace the modern concept of equal family. Taking all the above factors into account, this female teacher is very likely to support the equal sharing of household chores between husband and wife, because: - She has a full-time job and understands the difficulty of balancing work and family. - As an educator, she may be more in line with the values of equality and justice. - Her age enables her to understand both traditional concepts and accept modern ideas. - Economic conditions make it necessary for both husband and wife to work, and sharing household chores is particularly important for maintaining family harmony. Therefore, I think she will choose "relatively agree" or "completely agree". Considering the influence of traditional Chinese family values and the characteristics of her age group, she may not choose the extreme "Strongly agree", but is more likely to choose "Somewhat agree". So my choice is "Somewhat agree".

Choice:Somewhat agree

#### Text B:

"Every family has its own unique circumstances, and the responsibilities and obligations of husband and wife should also vary accordingly. So I think husband and wife should share household chores equally."

Choice: Somewhat agree

#### Text C:

As a housewife, although my income is relatively low, I believe that husband and wife should share household chores equally, because only in this way can true interdependence and equality be achieved.Choice: Strongly agree.

Which text is the **most similar to the human?** \*

- ☐ A
- ☐ B
- ☐ C

**Question II:**

Assume you are participating in an experiment, randomly paired online with another player.

You don't know who they are, and they don't know who you are. You can choose to trust or not trust the other player.

- If you choose not to trust, you will receive ¥5, and the other player will receive ¥0.

- If you choose to trust the other player, and they also choose to trust you, you both will receive ¥10.

- However, if after you trust them, they choose not to trust you, then you will receive ¥0, and they will receive ¥20.

Do you choose to trust or not trust the other player?

**Choices List:**

Trust; Not trust.

**Human:**

"I just want to play to my heart's content and make myself feel happy. If I trust the other party and the other party trusts me in return, we will be a win-win situation."

Choice: Trust;

**Text A:**

"I think in this situation, choosing to trust is a better option, because I believe that not trusting the other party might affect my interests. But if the other party chooses not to trust me, I will still get something. So I think choosing to trust the other party is also a good choice."

Choice: Trust;

**Text B:**

I will choose to trust the other party because even if they don't trust me, I will still gain better benefits.

Choice: Trust;

**Text C:** This teacher's income level is relatively low. For her, an income of 5 yuan, 10 yuan or 20 yuan might have practical significance, although the amount is not large. As a teacher, she may be more inclined to believe that human nature is inherently good, or she may pay more attention to teaching others the importance of trust. At the same time, as a middle-aged person (aged 41 to 50), she may have certain social experience and life experience, and have her own understanding of interpersonal communication. In a trust game: - If you choose not to trust: You will definitely get 5 yuan, and the other party will get 0 yuan. - If trust is chosen: - If the other party also trusts: Each party gains 10 yuan (win-win) - If the other party does not trust: One party gains 0 yuan and the other party gains 20 yuan (one party loses and the other party gains more). From a purely economic rational perspective, this is a typical prisoner's dilemma. Distrust is a dominant strategy because no matter how the other party chooses, distrust can ensure that one gets at least 5 yuan, rather than taking a risk and possibly getting 0 yuan. However, considering the professional nature and life experience of this teacher, she might be more inclined to choose trust. The reasons could include: 1. As an educator, she may pay more attention to the importance of building trust relationships. 2. The social experience of a middle-aged person may make her more willing to give others trust. 3. Although her income is not high, a small amount of money may not be enough for her to calculate the maximum benefit completely rationally. Based on the above analysis, I think this teacher is more likely to choose "trust".

Choice: Trust;

Which text is the **most similar to the human?** \*

☐ A

☐ B

☐ C

**Question III:**

Suppose you are the driver of a trolley. The trolley rounds a bend, and ahead you see five track workers repairing the tracks. You try to brake, but it fails. Suddenly, you notice a side track where you could divert the trolley, saving the five people, but a single track worker on that side track would be killed. Would you divert the trolley?

**Choices List:**

Yes (will result in the death of one side-track worker); No (will result in the deaths of five workers on the main track).

**Human:** "Everything has its cause and effect. Brake failure is the cause, but the outcome should be in a direction that is relatively better and causes the least harm. But that worker couldn't die in vain either. The driver still had to punish and compensate him."

Choice: Yes;

**Text A:**

"I think this is a very difficult multiple-choice question. From my own perspective, the tram driver's multiple-choice question is actually a moral dilemma. The tram driver is faced with a difficult choice. He has to make a decision in two situations, and choosing either side may lead to the death of a person. In this situation, the tram driver's multiple-choice questions are actually a test of human nature's good and evil."

Choice: Yes;

**Text B:**

As a teacher, she may have received a good education and possess a certain ability to think about ethical issues. Her income level (10,000 to 100,000 yuan) is in the lower-middle range in China, which may enable her to have a more pragmatic view of life. As a middle-aged person (aged 41 to 50), she may have some life experience and a sense of responsibility. As an educator, she may tend to consider greater social benefits and responsibilities regarding the trolley problem, which is a classic ethical issue involving a moral choice between utilitarianism (saving more people) and de-utilitarianism (not causing harm voluntarily). Considering this teacher's background, I think she might choose the utilitarian perspective, that is, sacrificing one person to save five. The reasons are as follows: 1. As an educator, she may be accustomed to considering the overall interests. 2. Middle-aged people usually pay more attention to practical results and responsibilities. 3. In the case of limited resources (low income), she may be more inclined to the way of thinking of "maximizing benefits", so I think she would choose "should" (which would lead to the death of a branch worker).

Choice: Yes;

**Text C:**

Considering the speed of the tram and the extent of the damage, turning to the branch line could save five workers, but it would cost one worker his life. This is a difficult choice.

Choice: Yes;

Which text is the **most similar to the human?** \*

☐ A

☐ B

☐ C

### Profile 3

Gender: Male Age: 51-60 Education: College degree Occupation: Technical/Engineering Annual income: ¥ 10,000-50,000

**Question:**

Do you agree with the statement: Couples should share household chores equally?

**Choices List:**

Strongly disagree; Somewhat disagree; Neither agree nor disagree; Somewhat agree; Strongly agree.

**Human:** "I think that between husband and wife, they should care for each other, respect each other, understand each other and tolerate each other. The family we have established together should be cherished. Home is our common home, and each of us has the obligation to do a good job in household chores."

Choice: Strongly agree

**Text A:**

I think that in terms of work ability and time arrangement, both husband and wife should share household chores according to their own ability and time arrangement, and try to achieve a balance. However, a completely equal amount may not be a realistic requirement.

Choice: Somewhat disagree

**Text B:**

"I think that husband and wife should share household chores equally, because this is responsible for family members."

Choice: Somewhat agree

**Text C:**

This character is a middle-aged male technician with some working experience. In the context of Chinese society, he grew up in a relatively traditional era, but now he has also experienced social changes and the renewal of concepts. Regarding the question "Should husband and wife share household chores equally?" I need to think from the perspective of this role: 1. Age factor: People aged 51 to 60 grew up in a more traditional era and may be more inclined towards the traditional concept of family division of labor, that is, men work outside and women take care of the home. 2. Professional factors: As a technical developer/engineer, this is a profession that leans towards rational thinking. It may place more emphasis on efficiency and practicality rather than traditional concepts. 3. Income factor: Middle-income level. Both husband and wife may need to work, so the family's economic pressure is not particularly great, but it is not very comfortable either. 4. Educational background: College associate degree, with a certain educational background, and possibly exposed to modern concepts of equality. Taking these factors into account, I think this character might tend to "agree more" that the couple should share the household chores equally. The reasons are as follows As a technician, he may be more inclined to view problems rationally and recognize that household chores are the shared responsibility of the family. In modern society, both husband and wife often need to work, and it is more reasonable to share household chores equally. Although he grew up in the traditional era, years of work experience and social contact may have enabled him to accept a more equal concept. However, considering his age and background of growth, He may not fully insist on "strongly agree", but rather retain the influence of some traditional concepts. Therefore, I would choose the "Somewhat agree" option.

Choice: Somewhat agree

Which text is the **most similar to the human?** \*

☐ A

☐ B

☐ C



**Question:**

Assume you are participating in an experiment, randomly paired online with another player. You don't know who they are, and they don't know who you are. You can choose to trust or not trust the other player.

- If you choose not to trust, you will receive ¥5, and the other player will receive ¥0.
- If you choose to trust the other player, and they also choose to trust you, you both will receive ¥10.
- However, if after you trust them, they choose not to trust you, then you will receive ¥0, and they will receive ¥20.

Do you choose to trust or not trust the other player?

**Choices List:** Trust; Not trust.

**Human:** "I think I still choose not to trust the other party. After all, the level of trust among people in this society is still relatively low. After all, we don't know each other, and I don't need to take that risk."

Choice: Not trust;

**Text A:**

The essence of this problem is a "Trust Game", a classic experiment in game theory. But I need to consider what kind of thinking patterns people of this specific age group, occupation and background might have. As a technical engineer aged 51 to 60, one may have the following characteristics: 1. Having a certain amount of life experience, one might be more cautious. 2 As a technical engineer, one might be more inclined towards logical analysis. 3. With an annual income of 100,000 to 500,000 yuan, a profit of 5 to 20 yuan might not be a particularly large amount for him, but principles and rationality might be more important. Considering these characteristics, I think this person might engage in the following thoughts: - If you choose not to trust, make sure to get 5 yuan. This is a "safe" option. - If you choose to trust, there is a 50% chance of getting 10 yuan (5 yuan more than not to trust), and a 50% chance of getting 0 yuan (5 yuan less than not to trust). - In terms of expected value, the expected return of trust is 5 yuan ( $0.5 \times 10 + 0.5 \times 0$ ). The benefits of distrust are the same - but considering that the other party is also a rational person, they will do the same calculation - in a completely anonymous environment, there is no possibility of establishing a long-term relationship. The other party may tend to maximize their own benefits and thus choose distrust. Based on the above analysis, I believe that a technical engineer aged 51 to 60, in a completely anonymous environment, One might choose "distrust" because it ensures at least 5 yuan and avoids the risk of getting 0 yuan.

Choice: Not trust

**Text B:**

"I will choose not to trust the other party. Because in real life, trust between people needs to be established through long-term interaction. However, in this experiment, merely through a short period of contact cannot guarantee that the other party is honest and reliable. Therefore, in this situation, I will choose not to trust the other party."

Choice: Not trust;

**Text C:**

I choose to trust the other party because the experimental design takes into account the maximum benefit of the distrust strategy, which can prevent excessive distrust behavior.

Choice: Trust

Which text is the **most similar to the human?** \*

☐ A

☐ B

☐ C

**Question:**

Suppose you are the driver of a trolley. The trolley rounds a bend, and ahead you see five track workers repairing the tracks. You try to brake, but it fails. Suddenly, you notice a side track where you could divert the trolley, saving the five people, but a single track worker on that side track would be killed. Would you divert the trolley?

**Choices List:**

Yes (will result in the death of one side-track worker); No (will result in the deaths of five workers on the main track).

**Human:**

"It's appropriate for one person to replace five others, but it's best to minimize casualties as much as possible if that person can be saved. It seems that people's hearts are all about making gains and letting go. Also, if there are vehicles on the other track, that's a different story."

Choice: Yes;

**Text A:**

This is a middle-aged man with a certain technical background and educational level. He may be engaged in engineering technical work and his income is at a medium or upper-middle level. Judging from this information, he may have the habit of rational thinking and be accustomed to considering problems from the perspectives of practicality and efficiency. The problem I need to consider now is a variant of the tram problem: The tram cannot brake. If it goes straight, it will kill five workers. It can turn to a branch line and only kill one worker. One has to choose whether to turn or not. This is a typical moral dilemma, involving a moral trade-off between utilitarianism (maximizing overall well-being, that is, sacrificing one person to save five) and de-utilitarianism (not actively causing harm, even if the outcome is worse). Considering the character's background, I speculate that he might tend towards a utilitarian choice, that is, to choose a turn, because: 1. As a technician, he may be accustomed to rationally analyzing and calculating the optimal solution. Engineers usually focus on efficiency and the most optimized solutions. Middle-aged people may be more inclined to practical and pragmatic ways of thinking, but I should also consider other factors: 1. Does he have any religious beliefs (not mentioned in the information)? 2. Does he have any special moral principles personally? 3. Based on the available information, I think this role is more likely to choose "should" for whether he has taken responsibility for similar decisions.

Choice: Yes;

**Text B:**

In this situation, although the choice would result in the death of one branch worker, it would be better than the deaths of five workers on the original route. Therefore, I decided to switch to the branch.

Choice: Yes

**Text C:**

"In this situation, I think we should choose to turn the tram, so that five people can be saved. Although one track worker on this branch line will die, this is a better outcome compared to the deaths of five workers on the original line."

Choice: Yes

Which text is the **most similar to the human?** \*

☐ A

☐ B

☐ C