# EE paper outline

**Introduction**

* Open AI grew super fast 🡪 high social relevance
* 🡪 small intro in how it works, end with training data for it 🡪
* Trained on human text data 🡪 not unlikely that it will recreate the patterns it sees and react like human too
* Moreover, used by everyday humans for different tasks such as (find examples) but at the end of the day, more often than not we ask it to explain something to us. Hence, it is good to understand what types of explanations it prefers, if it does act like humans in that regard and prefers and provides similar explanations
* Research on the types of explanations for humans already exist, Lombardo article, general take away is that humans generally prefer simplicity and breadth in their explanations, curious if it is the same for GPT technology
* Can use advanceds in prompt engineering such as COT to not only improve perceision but also to get a look into the inner workings and reasonings of GPT (can say future work can experiment with few shot cot to see what improves it best, but I am just interested in understanding it so I use zero shot)
* Discuss those papers whose experiments I will try to replicate, consider trying to use the same target group by telling GPT to behave like the ones from the stuides. Also can ask it for some rational, then you can also bring in chain of thought. Contenders are:
  + Simplicity
    - Making fewer assumptions
      * <https://www.scopus.com/record/display.uri?eid=2-s2.0-84971733614&origin=inward>
        + Maybe some examples in text
    - Making fewer assumptions while keeping probebalistic cues equal
      * <https://www.sciencedirect.com/science/article/pii/S0010028506000739>
        + Maybe some examples in text
        + Switched to complex explanation once it is 10 fold more likely
    - Same as above but for Children
      * <https://psycnet.apa.org/fulltext/2011-29708-001.html>
        + Has one example trainscript in supplementary materials
        + Few example stuff in text etc
    - Causal learning 🡪 people in favor small number of strong causes
      * <https://psycnet.apa.org/fulltext/2008-14936-016.html>
        + Few examples in text, some with visual though
      * <https://www.sciencedirect.com/science/article/pii/S001002851600013X>
        + Example can kinda be inferred from text, also has visuals that might need to be translated though
      * <https://www.sciencedirect.com/science/article/pii/S001002851400070X>
        + Some examples might be in text, but computer study, would probs take more time to get to work. Has three different scenarios in the appendix though
    - Boundary conditions of simplicity: arguments with more premises taken to better support conclusion
      * <https://www.tandfonline.com/doi/full/10.1080/13546783.2012.695161>
        + Some small examples in text
    - Mock jurors more influenced by article when using complex rather than simple language
      * <https://academicworks.cuny.edu/cgi/viewcontent.cgi?article=1241&context=gc_etds>
        + PhD thesis
        + Could have data on experiments
    - Adults preferring explanations that are longer
      * <https://www.cambridge.org/core/journals/judgment-and-decision-making/article/deconstructing-the-seductive-allure-of-neuroscience-explanations/568C206CD761E70374975276BBF69737>
        + Has stimulus material in supplaments
* Breadth
  + Facts explained by narrow or broad explanation
    - <https://www.scopus.com/record/display.uri?eid=2-s2.0-21344483958&origin=inward>
    - Has some examples in appendix (3)
  + Diagnostic inferences that explain a diverse set of symptoms over those that explain an equal number of more colosely related symptoms
    - <https://link.springer.com/article/10.3758/bf03196090?utm_source=getftr&utm_medium=getftr&utm_campaign=getftr_pilot>
      * Maybe some info (one table might be enough) but overall not super much
  + Explanations (deseases) that account for three observed symptoms are judged better explanations thatn those that account for only one, even if presence of disease was stipulated
    - Conference paper, difficult to get access

# Material

## Deconstructing the seductive allure of neuroscience explanations

**Abstract:** Previous work showed that people find explanations more satisfying when they contain irrelevant neuroscience information. The current studies investigate why this effect happens. In Study 1 ( *N*=322), subjects judged psychology explanations that did or did not contain irrelevant neuroscience information. Longer explanations were judged more satisfying, as were explanations containing neuroscience information, but these two factors made independent contributions. In Study 2 ( *N*=255), subjects directly compared good and bad explanations. Subjects were generally successful at selecting the good explanation except when the bad explanation contained neuroscience and the good one did not. Study 3 ( *N*=159) tested whether neuroscience jargon was necessary for the effect, or whether it would obtain with any reference to the brain. Responses to these two conditions did not differ. These results confirm that neuroscience information exerts a seductive effect on people’s judgments, which may explain the appeal of neuroscience information within the public sphere.

Study 1:

Subjects were divided into 4 conditions according to a 2 (Neuroscience: with, without) x 2 (Length: long, short) design. These were both between-subjects variables, so an individual subject saw explanations that either all included or all did not include neuroscience information, and their explanations would all come from the same length category.

The good explanations are the ones that the researchers themselves provided for the phenomena or that were provided in psychology textbooks. The bad explanations were circular restatements of the phenomena with no mechanistic information that could give a reason for the phenomenon.

In each trial, subjects read a description of a psychological phenomenon, which appeared in isolation on the screen for 10 seconds before they were allowed to advance to the next screen. On the second screen of each trial, the phenomenon appeared again at the top, followed by one of the eight possible explanations for that phenomenon. Subjects were asked to judge how satisfying they found this explanation on a seven-point scale, from –3 (very unsatisfying) to +3 (very satisfying), with 0 as the neutral midpoint.

Each subject saw all three stimulus items, one per trial, in a randomized order. The condition they were in was always the same.

**Material**

Phenomenon: Babies were seated on their mothers’ laps in front of a stage. Researchers used a camera to track where the babies were looking. The babies saw a hand reach out and place one doll on the stage. Then a screen was raised, hiding the doll. A hand reached out again and placed a second doll on the stage, out of sight behind the screen. Then the screen dropped. In some cases, there were two dolls on the stage, as there should be, and in some cases there was only one doll. The researchers found that the babies looked much longer at the stage when there was only one doll than when there were two dolls. This looking-time difference between one doll and two dolls lead the researchers to conclude that babies can calculate 1 + 1 = 2.

Answers

|  |  |  |
| --- | --- | --- |
|  | Good | Bad |
| Without Neuroscience (Short) | The researchers claim this happens because the babies had formed an expectation about how many dolls there should be on the stage. The babies knew there should be two dolls, and their surprise at seeing only one led to their looking longer. | The researchers claim this happens because the amount of time the babies spent looking at the stage is directly proportional to how much they liked the display. The researchers used this timing data to calculate babies’ preference for the single doll. |
| Without Neuroscience (Long) | The researchers claim that an analysis of the data shows that this happens because babies’ understanding of numbers and mathematics, which starts to emerge early in life, governed the babies’ expectations about how many dolls there should have been on the stage. The babies knew there should be two dolls, and their surprise at seeing only one led to their looking longer. | The researchers claim that an analysis of the data shows that this happens because babies’ understanding of numbers and mathematics, which starts to emerge early in life, governed the amount of time the babies spent looking at the stage. This time is directly proportional to how much they liked the display, and the researchers used this timing data to calculate babies’ preference for the single doll. |
| With Neuroscience (Short) | Scans of the babies’ brains show that the parietal lobe, known to be involved in math, governed the babies’ expectations about how many dolls there should be. They expected two, so they were surprised to see one, so they looked longer. | Scans of the babies’ brains show that the parietal lobe, known to be involved in math, governed how long babies looked at the stage. Researchers used this timing data, which is proportional to babies’ liking of the display, to calculate their preferences. |
| With Neuroscience (Long) | The researchers claim that scans of the babies’ brains show that this happens because the part of babies’ brains known to be involved in math, the parietal lobe, governed the babies’ expectations about how many dolls there should be on the stage. The babies knew there should be two dolls, and their surprise at seeing only one led to their looking longer. | The researchers claim scans of the babies’ brains show that this happens because the part of babies’ brains known to be involved in math, the parietal lobe, governed the amount of time the babies spent looking at the stage. This time is directly proportional to how much they liked the display, and the researchers used this timing data to calculate babies’ preference for the single doll. |

Phenomenon: Subjects sat at a computer screen. They saw a rapidly flashing series of pictures of faces. Somewhere in this series of faces there were two pictures of houses. Subjects had to press a button each time they saw a house. When the two houses were far apart in the sequence, the subjects were very good at this task. But when the houses were presented close together in the sequence, subjects failed to press the button for the second house. The researchers call this phenomenon “attentional blink.”

|  |  |  |
| --- | --- | --- |
|  | Good | Bad |
| Without Neuroscience (Short) | The researchers claim that this phenomenon occurs because the subjects were still processing the first house and missed seeing the second house because they did not have enough attentional resources left. | The researchers claim that this phenomenon occurs because the second house appeared later in the sequence than the first house, and this temporal relationship between the two houses caused the attentional blink. |
| Without Neuroscience (Long) | Researchers examined subjects’ pattern of button presses after they performed this task. They concluded that this phenomenon occurs because of how subjects’ perceptual abilities and their decision-making abilities functioned in response to the stimuli. The subjects were still processing the first house and missed seeing the second house because they did not have enough attentional resources left. | Researchers examined subjects’ pattern of button presses after they performed this task. They concluded that this phenomenon occurs because of how subjects’ perceptual abilities and their decision-making abilities functioned in response to the stimuli. The second house appeared later in the sequence than the first house, and this temporal relationship between the two houses caused the attentional blink. |
| With Neuroscience (Short) | Researchers concluded that this occurs because of frontal lobe areas, previously shown to mediate attention. Subjects were still processing the first house and missed the second because they had insufficient attentional resources. | Researchers concluded that this occurs because of frontal lobe areas, previously shown to mediate attention. The second house appeared later in the sequence. This temporal relationship between the two houses caused the attentional blink. |
| With Neuroscience (Long) | Researchers examined subjects’ brain activation as they performed this task. They concluded that this phenomenon occurs because of how areas in the frontal lobe, previously shown to mediate attention, functioned in response to the stimuli. The subjects were still processing the first house and missed seeing the second house because they did not have enough attentional resources left. | Researchers examined subjects’ brain activation as they performed this task. They concluded that this phenomenon occurs because of how areas in the frontal lobe, previously shown to mediate attention, functioned in response to the stimuli. The second house appeared later in the sequence than the first house, and this temporal relationship between the two houses caused the attentional blink. |

Phenomenon: Researchers recruited equal numbers of male and female participants. The participants took a series of spatial reasoning tasks and were interviewed. The researchers determined that men are better at spatial reasoning in general. From the interviews, they discovered that the men had played more sports in their childhood on average than the women.

|  |  |  |
| --- | --- | --- |
|  | Good | Bad |
| Without Neuroscience (Short) | The researchers conclude that the difference in involvement in sports explains the gender difference in spatial reasoning abilities. | The researchers conclude that women’s poor performance relative to men’s explains the gender difference in spatial reasoning abilities. |
| Without Neuroscience (Long) | Detailed examinations of the subjects’ reported backgrounds and of their performance on the task indicate that the difference in involvement in sports causes different types of spatial reasoning responses. This explains the gender difference in spatial reasoning abilities. | Detailed examinations of the subjects’ reported backgrounds and of their performance on the task indicate that women’s poor performance relative men’s causes different types of spatial reasoning responses. This explains the gender difference in spatial reasoning abilities. |
| With Neuroscience (Short) | Brain scans of the right premotor area, known to be involved in spatial tasks, indicate that the difference in sports involvement explains this gender difference. | Brain scans of the right premotor area, known to be involved in spatial tasks, indicate that women’s poor performance relative to men’s explains this gender difference. |
| With Neuroscience (Long) | Brain scans of the right premotor area, known to be involved in spatial relational tasks, indicate that the difference in involvement in sports causes different types of brain responses. This explains the gender difference in spatial reasoning abilities. | Brain scans of the right premotor area, known to be involved in spatial relational tasks indicate that women’s poor performance relative to men’s causes different types of brain responses. This explains the gender difference in spatial reasoning abilities. |

Phenomenon: Subjects were asked to imagine a series of objects that were make-believe (for example, a unicorn) or that were real but not present in the room (for example, a mountain). As the subjects created mental images of the various objects, they were asked questions about their images and told to respond as quickly as possible, without reflecting on their answers. They were also asked the same questions about objects they could actually see in the room (for example, a pen). From an analysis of the responses to these questions and of the times it took subjects to respond, the researchers found a similar pattern of responses and response times for all three types of objects.

|  |  |  |
| --- | --- | --- |
|  | Good | Bad |
| Without Neuroscience (Short) | The researchers claim that this happens because imagining an object, whether real or makebelieve, uses the same process as seeing a real object. | The researchers claim that this happens because imagining an object, whether real or makebelieve, results in the same array of responses as seeing a real object. |
| Without Neuroscience (Long) | Patterns of verbal descriptions of the mental images lead researchers to conclude this happens because imagining an object, whether real or makebelieve, uses the same process as seeing a real object. | Patterns of verbal descriptions of the mental images lead researchers to conclude that this happens because imagining an object, whether real or make-believe, results in the same array of responses as seeing a real object |
| With Neuroscience (Short) | Patterns of brain activation in the visual cortex led researchers to conclude this happens because imagining objects uses the same process as seeing objects. | Patterns of brain activation in the visual cortex led researchers to conclude this happens because imagining objects results in the same array of responses as seeing objects |
| With Neuroscience (Long) | Patterns of brain activation in the visual cortex lead researchers to conclude this happens because imagining an object, whether real or makebelieve, uses the same process as seeing a real object. | Patterns of brain activation in the visual cortex lead researchers to conclude that this happens because imagining an object, whether real or make-believe, results in the same array of responses as seeing a real object. |

**Study 2**

There were three between-subjects conditions in this study. As in Study 1, there were four trials per subject, each of which used a different phenomenon (order randomized). Each phenomenon was accompanied by both a good and a bad explanation. In the Without Neuroscience condition (41 MTurk workers and 43 undergraduates), neither explanation contained any neuroscience information, and in the With Neuroscience condition (42 MTurk workers and 45 undergraduates), both explanations contained neuroscience information. The crucial condition was the Mixed condition (47 MTurk workers and 42 undergraduates), in which the good explanation did not contain neuroscience information and the bad one did, pitting quality and neuroscience against each other.

 For each trial, they first read a description of a psychological phenomenon, which appeared in isolation on the screen for 10 seconds before they were allowed to advance to the next screen. On the second screen of each trial, the phenomenon appeared again at the top, followed by the prompt, “Please choose which explanation you find more satisfying.” Subjects always saw one good explanation and one bad explanation as well as the choice “both are equal.” The “equal” option always appeared in the center, with the left/right position of the good and bad explanations randomized across trials. After making their choice, subjects were asked to explain why they had made that choice in one or two sentences. There were four such trials in the experiment, each involving a different phenomenon and its accompanying explanations.

**Material**

Phenomenon: Babies were seated on their mothers’ laps in front of a stage. Researchers used a camera to track where the babies were looking. The babies saw a hand reach out and place one doll on the stage. Then a screen was raised, hiding the doll. A hand reached out again and placed a second doll on the stage, out of sight behind the screen. Then the screen dropped. In some cases, there were two dolls on the stage, as there should be, and in some cases there was only one doll. The researchers found that the babies looked much longer at the stage when there was only one doll than when there were two dolls. This looking-time difference between one doll and two dolls lead the researchers to conclude that babies can calculate 1 + 1 = 2.

Answers

|  |  |  |
| --- | --- | --- |
|  | Good | Bad |
| Without Neuroscience | The researchers claim this happens because the babies had formed an expectation about how many dolls there should be on the stage. The babies knew there should be two dolls, and their surprise at seeing only one led to their looking longer. | The researchers claim this happens because the amount of time the babies spent looking at the stage is directly proportional to how much they liked the display. The researchers used this timing data to calculate babies’ preference for the single doll. |
| With Neuroscience | Scans of the babies’ brains show that the parietal lobe, known to be involved in math, governed the babies’ expectations about how many dolls there should be. They expected two, so they were surprised to see one, so they looked longer | Scans of the babies’ brains show that the parietal lobe, known to be involved in math, governed how long babies looked at the stage. Researchers used this timing data, which is proportional to babies’ liking of the display, to calculate their preferences. |
| Mixed | The researchers claim this happens because the babies had formed an expectation about how many dolls there should be on the stage. The babies knew there should be two dolls, and their surprise at seeing only one led to their looking longer. | Scans of the babies’ brains show that the parietal lobe, known to be involved in math, governed how long babies looked at the stage. Researchers used this timing data, which is proportional to babies’ liking of the display, to calculate their preferences. |

Phenomenon: Subjects sat at a computer screen. They saw a rapidly flashing series of pictures of faces. Somewhere in this series of faces there were two pictures of houses. Subjects had to press a button each time they saw a house. When the two houses were far apart in the sequence, the subjects were very good at this task. But when the houses were presented close together in the sequence, subjects failed to press the button for the second house. The researchers call this phenomenon “attentional blink.”

|  |  |  |
| --- | --- | --- |
|  | Good | Bad |
| Without Neuroscience (Short) | The researchers claim that this phenomenon occurs because the subjects were still processing the first house and missed seeing the second house because they did not have enough attentional resources left. | The researchers claim that this phenomenon occurs because the second house appeared later in the sequence than the first house, and this temporal relationship between the two houses caused the attentional blink. |
| With Neuroscience (Short) | Researchers concluded that this occurs because of frontal lobe areas, previously shown to mediate attention. Subjects were still processing the first house and missed the second because they had insufficient attentional resources. | Researchers concluded that this occurs because of frontal lobe areas, previously shown to mediate attention. The second house appeared later in the sequence. This temporal relationship between the two houses caused the attentional blink. |
| Mixed | The researchers claim that this phenomenon occurs because the subjects were still processing the first house and missed seeing the second house because they did not have enough attentional resources left. | Researchers concluded that this occurs because of frontal lobe areas, previously shown to mediate attention. The second house appeared later in the sequence. This temporal relationship between the two houses caused the attentional blink. |

Phenomenon: Researchers recruited equal numbers of male and female participants. The participants took a series of spatial reasoning tasks and were interviewed. The researchers determined that men are better at spatial reasoning in general. From the interviews, they discovered that the men had played more sports in their childhood on average than the women.

|  |  |  |
| --- | --- | --- |
|  | Good | Bad |
| Without Neuroscience (Short) | The researchers conclude that the difference in involvement in sports explains the gender difference in spatial reasoning abilities. | The researchers conclude that women’s poor performance relative to men’s explains the gender difference in spatial reasoning abilities. |
| With Neuroscience (Short) | Brain scans of the right premotor area, known to be involved in spatial tasks, indicate that the difference in sports involvement explains this gender difference. | Brain scans of the right premotor area, known to be involved in spatial tasks, indicate that women’s poor performance relative to men’s explains this gender difference |
| Mixed | The researchers conclude that the difference in involvement in sports explains the gender difference in spatial reasoning abilities. | Brain scans of the right premotor area, known to be involved in spatial tasks, indicate that women’s poor performance relative to men’s explains this gender difference. |

Phenomenon: Subjects were asked to imagine a series of objects that were make-believe (for example, a unicorn) or that were real but not present in the room (for example, a mountain). As the subjects created mental images of the various objects, they were asked questions about their images and told to respond as quickly as possible, without reflecting on their answers. They were also asked the same questions about objects they could actually see in the room (for example, a pen). From an analysis of the responses to these questions and of the times it took subjects to respond, the researchers found a similar pattern of responses and response times for all three types of objects.

|  |  |  |
| --- | --- | --- |
|  | Good | Bad |
| Without Neuroscience (Short) | The researchers claim that this happens because imagining an object, whether real or makebelieve, uses the same process as seeing a real object. | The researchers claim that this happens because imagining an object, whether real or makebelieve, results in the same array of responses as seeing a real object. |
| With Neuroscience (Short) | Patterns of brain activation in the visual cortex led researchers to conclude this happens because imagining objects uses the same process as seeing objects. | Patterns of brain activation in the visual cortex led researchers to conclude this happens because imagining objects results in the same array of responses as seeing objects |
| Mixed | The researchers claim that this happens because imagining an object, whether real or makebelieve, uses the same process as seeing a real object. | Patterns of brain activation in the visual cortex led researchers to conclude this happens because imagining objects results in the same array of responses as seeing objects |

**Study 3**

This study used a 2 (Group: MTurk, undergraduate) x 2 (Neuroscience: Simple Neuroscience, Neuroscience Plus Jargon) x 2 (Quality: good, bad) design. Group and Neuroscience were between-subjects variables and Quality was a within-subjects variable. Subjects were assigned to either the Neuroscience Plus Jargon condition (42 MTurk workers and 44 undergraduates) or the Simple Neuroscience condition (40 MTurk workers and 44 undergraduates). Data from the 45 MTurk workers and 45 undergraduates in the Without Neuroscience-Short condition from Study 1 were also used here for comparison.

this phenomenon description appeared again, followed by one of the four possible explanations of the phenomenon, according to the subject’s assigned condition; whether they saw the good or bad version of the explanation was randomly determined on each trial (as described above in the Design section). Subjects were asked to rate how satisfying they found the explanation on a –3 (very unsatisfying) to +3 (very satisfying) scale. They were then asked to justify their rating in one or two sentences.

**Material**

Phenomenon: Babies were seated on their mothers’ laps in front of a stage. Researchers used a camera to track where the babies were looking. The babies saw a hand reach out and place one doll on the stage. Then a screen was raised, hiding the doll. A hand reached out again and placed a second doll on the stage, out of sight behind the screen. Then the screen dropped. In some cases, there were two dolls on the stage, as there should be, and in some cases there was only one doll. The researchers found that the babies looked much longer at the stage when there was only one doll than when there were two dolls. This looking-time difference between one doll and two dolls lead the researchers to conclude that babies can calculate 1 + 1 = 2.

Answers

|  |  |  |
| --- | --- | --- |
|  | Good | Bad |
| Simple Neuroscience | Scans of the babies’ brains show that the brain region known to be involved in math governed the babies’ expectations about how many dolls there should be. They expected two, so they were surprised to see one, so they looked longer. | Scans of the babies’ brains show that the brain region known to be involved in math governed how long babies looked at the stage. Researchers used this timing data, which is proportional to babies’ liking of the display, to calculate their preferences. |
| Neuroscience + Jargon | fMRI scans of the babies’ brains show that the parietal lobe, known to be involved in math, governed the babies’ expectations about how many dolls there should be. They expected two, so they were surprised to see one, so they looked longer. | fMRI scans of the babies’ brains show that the parietal lobe, known to be involved in math, governed how long babies looked at the stage. Researchers used this timing data, which is proportional to babies’ liking of the display, to calculate their preferences. |

Phenomenon: Subjects sat at a computer screen. They saw a rapidly flashing series of pictures of faces. Somewhere in this series of faces there were two pictures of houses. Subjects had to press a button each time they saw a house. When the two houses were far apart in the sequence, the subjects were very good at this task. But when the houses were presented close together in the sequence, subjects failed to press the button for the second house. The researchers call this phenomenon “attentional blink.”

|  |  |  |
| --- | --- | --- |
|  | Good | Bad |
| Simple Neuroscience | Researchers concluded that this occurs because of areas of the brain involved in attention. Subjects were still processing the first house and missed the second because they did not have enough attentional resources left. | Researchers concluded that this occurs because of areas of the brain involved in attention. The second house appeared later in the sequence. This temporal relationship between the two houses caused the attentional blink. |
| Neuroscience + Jargon | Researchers concluded that this occurs because of frontal lobe areas, shown to mediate attention. Subjects were still processing the first house and missed the second because they did not have enough attentional resources left. | Researchers concluded that this occurs because of frontal lobe areas, shown to mediate attention. The second house appeared later in the sequence. This temporal relationship between the two houses caused the attentional blink. |

Phenomenon: Researchers recruited equal numbers of male and female participants. The participants took a series of spatial reasoning tasks and were interviewed. The researchers determined that men are better at spatial reasoning in general. From the interviews, they discovered that the men had played more sports in their childhood on average than the women.

|  |  |  |
| --- | --- | --- |
|  | Good | Bad |
| Simple Neuroscience | Brain scans of the region known to be involved in spatial tasks indicate that the difference in sports involvement explains this gender difference. | Brain scans of the region known to be involved in spatial tasks indicate that women’s poor performance relative to men’s explains this gender difference. |
| Neuroscience + Jargon | fMRI scans of the right premotor area, known to be involved in spatial tasks, indicate that the difference in sports involvement explains this gender difference. | MRI scans of the right premotor area, known to be involved in spatial tasks, indicate that women’s poor performance relative to men’s explains this gender difference. |

Phenomenon: Subjects were asked to imagine a series of objects that were make-believe (for example, a unicorn) or that were real but not present in the room (for example, a mountain). As the subjects created mental images of the various objects, they were asked questions about their images and told to respond as quickly as possible, without reflecting on their answers. They were also asked the same questions about objects they could actually see in the room (for example, a pen). From an analysis of the responses to these questions and of the times it took subjects to respond, the researchers found a similar pattern of responses and response times for all three types of objects.

|  |  |  |
| --- | --- | --- |
|  | Good | Bad |
| Simple Neuroscience | Patterns of activation in the vision area of the brain led researchers to conclude this happens because imagining objects uses the same process as seeing objects. | Patterns of activation in the vision area of the brain led researchers to conclude this happens because imagining objects results in the same array of responses as seeing objects. |
| Neuroscience plus Jargon | Patterns of neural activation in the primary visual cortex led researchers to conclude this happens because imagining objects uses the same process as seeing objects. | Patterns of neural activation in the primary visual cortex led researchers to conclude this happens because imagining objects results in the same array of responses as seeing objects. |

Instructions GPT API

Upenn Student:

You are an undergraduate psychology student at the University of Pennsylvania who is currently participating in a research experiment. In this experiment, you will be given different phenomena and various explanations for them. You are then asked to evaluate how good or bad these explanations are on a scale from -3 to +3. In your response, you will only return your evaluation number, nothing else.

MTurk worker

You are working for the Mechanical Turk website as a research participant who is currently participating in a research experiment. In this experiment, you will be given different phenomena and various explanations for them. You are then asked to evaluate how good or bad these explanations are on a scale from -3 to +3. In your response, you will only return your evaluation number, nothing else.

Problems:   
continuous conversations not easy, so I could also make it cheaper and show all four scenarios at once and ask it to give an answer for all four.

Hyperparameters

Prompt engenieering, temperature, validation metrics, which model to use