Chapter 6
Philosophical Thought Experiments, Intuitions, and Cognitive Equilibrium

Grace Paterson

Feb 6 2012

Three Main Claims

1. Thought experiments are processed differently from more abstract reasoning and thus may evoke different sorts of responses from us.

2. These responses may actually conflict with responses to the same material presented in a different manner. As a result, achieving cognitive equilibrium may not always be possible.

3. When the thought experiment becomes the dominant way of thinking about a problem they can be highly persuasive. This is desirable in some areas of philosophy, such as moral philosophy.

Cognitive Underpinnings

- There is plenty of psychological data about how "tasks with the same formal structure but different contents may prompt different rates of success, presumably because the alternative framings activate different processing mechanisms." (118)
- Gendler will provide a brief survey of such research (some of which we have discussed before)
- These cases can be contrasted with philosophical examples but have the advantage of being a little more clear, a little less ambiguous.
    - We know what the right and wrong answers are and so can focus on how well people are actually reasoning about them.

Thorndike

- E.L. Thorndike, 1922 "The Effect of Changed Data on Reasoning"
  - Subjects were given simple algebra problems
  - Different groups were given the problem using different symbol sets
  - Result: error rates were significantly higher among those given "complicated" symbols as compared to those given "simple" symbols
- This is consistent with experience, I think (try reading *Principia Mathematica*)

Wilkins' Experiment

- Minna Cheves Wilkins, 1928 on syllogistic reasoning
- Subjects were asked to judge whether conclusions followed from premises
- Syllogisms were presented in different ways:
  1. With concrete, familiar terms
     - "Some of the girls in the chorus wear their hair braided; all the girls in the chorus wear their hair bobbed; therefore..."
  2. With abstract symbols
     - "All x's are z's; all x's are y's; therefore..."
  3. With nonsense terms
     - "No juritobians are cantabilians; no cantixianti are cantabilians; therefore..."
  4. With familiar terms where subjects had "antecedent views about their relations"
     - "If New York is to the right of Detroit; and Chicago is to the left of New York; then..." (119)

Wilkins' Results

- Results showed reasoning with abstract symbols was harder than familiar symbols
- Except in the case of certain fallacies
- We can be both aided and mislead by our own preconceptions
- This prompted further study of such "interference effects"
- Such experiments come mainly in two flavours:
  1. Syllogism Tasks
  2. Wason Selection Tasks

Syllogism Tasks

- As with Wilkins' original experiment, subjects are given syllogistic premises and asked if a particular conclusion logically follows
- Problems vary along two dimensions:
  1. valid vs invalid reasoning
  2. plausible vs implausible conclusions
- Results of these types of experiments:
  - *Valid* inferences are less likely to be judged valid if *implausible*
  - *Invalid* inferences are less likely to be judged invalid if *plausible*
- So people associate validity with plausibility, and invalidity with implausibility. Mixing these up increases the chance of error.
  - This is called *belief bias*

Wason Selection Tasks

- Due to Wason, 1966
- Shows how bad people are at understanding material implication
- Subjects are shown 4 cards: A; D; 3; 7
- Each card has a number on one side and a letter on the other side.
- Told "If there is an A on one side, then there is a 3 on the other"
- Question: Which cards do you need to turn over to verify if that statement is true?

Wason Results

- Answer: you need to turn over the A and the 7
- But fewer than 10% of subjects get this right!
- Most people turn over some combination of the A and the 3 (but don't touch the 7)

Wason Variants

Some things have been found to dramatically improve people's performance:

- Change the phrasing of the statement to "if there is an A on one side, there is not a 7 on the other" and most people get it right!
  - This is called *matching bias* because people match their response to the stimulus. Whatever is explicitly mentioned is more likely to be turned over.
- Replace the symbols with words related to social rules.
  - We now have cards with ages on one side and drinks on the other
  - Subjects are shown Beer; Coke; 21; 16
  - Q: what do you need to turn over to verify the truth of "if a person is drinking beer, then the person must be at least 21 years of age."? (this experiment was clearly performed in the US)
  - A: Beer and 16. Most people get this right.

Wason Variants

- Gendler points out in the end-notes that there have been multiple methods discovered to improve performance on Wason selection tasks:
    - Using concrete and meaningful terms
    - Presenting the task in terms of rule-violation rather than truth/falsity
    - Embedding the task in a role-playing scenario
    - Relating the two components (sides of the cards) meaningfully
- Most of these have to do with making the reasoning more concrete and less abstract.
- Q: Could belief bias also be exploited here? What if the phrase being verified were "if a person is drinking beer, then the person must *not* be at least 21 years of age." Would that change things?

Jelly Beans

Studies have also shown that people are bad at probabilistic reasoning:

- ▶ Subjects were shown bowls containing mixes of red and white jelly beans. Some contained 100 beans total, some contained 10 beans total

- ▶ Each bowls was *clearly labelled* with the percentage of red jelly beans in it.

- ▶ Subjects were told that they would get a dollar for every time they blindly drew a red jelly bean

- ▶ They were then given a choice between two bowls to draw from:
    1. Contained a lower percentage of red beans but more beans overall
    2. Contained fewer beans total but a higher proportion of them were red.

Jelly Beans Results

- Clearly you should always choose (2)
- ...But many subjects did silly things such as choosing the 9:100 bowl over the 1:10 bowl!
- This boggles my mind!
- Subjects said things like "I picked the ones with more red jelly beans because it looked like there were more ways to get a winner, even though I knew there were also more whites, and that the percents were against me"(122)
- So even though they *knew* it was the wrong move, they still did it.

Linda the Bank Teller

Here is another famous example!

- Subjects presented with a blurb: "Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations'.'

- They are then given eight statements and asked to rank them in order of likelihood. Two such statements are:
    1. Linda is a bank Teller (T)
    2. Linda is a bank Teller and is active in the Feminist movement (T+F)

Linda the Bank Teller Results

- Clearly T is more probable than T+F
  - After all, T+F *implies* T
- Yet many people rank T+F above T anyway!
- *I know that [T+F] is least probable, yet a little homunculus in my head continues to jump up and down, shouting at me–'but she can't just be a bank teller; read the description'* –Stephen Jay Gould (not a dumb guy)

Characterizing Homunculi

- In all of these experiments, most subjects "feel the pull" of the wrong answer although they may be able to override it
  - Speed and accuracy improve for everyone when conditions are made more bias friendly
- A popular approach to explaining this is the *dual systems* approach (which we have seen already)
  - "According to such accounts, there are at least two clusters of subsystems involved in mental processing– one associative and instinctive, operating rapidly and automatically; the other rule-based and regulated, operating in a relatively slow and controlled fashion." (123)
  - You know, left-brain/right-brain type stuff

Two Systems

What Gendler cares about in all this:

> ...when content is presented in a suitably concrete or
> abstract way, this may result in the activation or
> fortification of a representational schema that was
> otherwise inactive or subordinate. The result of this may
> be to evoke responses that run counter to those evoked
> by alternative presentations of relevantly similar
> content.(124)

In other words, *different presentations* of the *same content* may
activate *different cognitive systems* and may therefore evoke
different intuitions.

Implications for Thought Experiments

Assuming it is correct, what does all this tell us about the use of thought experiments in philosophy?

- A theory may be correct even if there are cases where our intuitions conflict

- This is *disturbing* because we must accept cognitive disequilibrium as part of philosophical life

- It is *liberating* because we aren't beholden to our intuitions as the ultimate standard

Example: Trolley Problems

A famous thought experiment in moral philosophy (Thomson 1985)

- Version 1:You are driving a trolley and are about to run over a group of five people, but at the last minute you notice that you can turn the trolley onto a track on which there is only one person who will be hit. Should you turn the trolley?
    - Most people answer "yes"
- Version 2: This time you are standing on a bridge above the track and see an out of control trolley about to hit five people. There is a fat man (large enough to stop the trolley) on the bridge as well and you could push him over to stop the trolley from hitting the five men. Should you push the man?
    - Most people answer "no"

Example: Trolley Problems

- In the abstract these cases are the same (sacrifice one life for five), however we don't process them in the same way!
    - Version 1 activates "higher cognitive" regions of the brain
    - Version 2 activates "emotional/social" regions of the brain
- Do we wish to integrate these distinctions into our moral philosophy?

More Problems for Moral Intuitions

Gendler points out that "parallel differences can be evoked by what seem clearly to be morally irrelevant variables." (126)

- For instance, another experiment changed the fat man scenario in two different ways:
    1. "Chip Ellsworth III could be thrown off a bridge to stop a trolley hurtling towards 100 members of the Harlem Jazz Orchestra."
    2. "Tyrone Peyton could be thrown off to save 100 members of the New York Philharmonic"

- Judgements of the moral acceptability of these scenarios were politically motivated. Liberal subjects found (1) more acceptable than (2)

- How can the orchestra in question possibly change the moral acceptability of this action?

It Gets Worse

- We can change people's responses using unconscious priming too.
  - unscrambling sentences related to patriotism (multiculturalism) causes subjects to have responses consistent with conservatives (liberals.)
- This extends to epistemology as well:
  - "Jonathan Weinberg and colleagues have discovered that subjects' willingness to attribute knowledge in ambiguous cases 'increases after being presented with a clear case of non-knowledge, and...decreases after being presented with a clear case of knowledge' "(127)
  - This harkens back to chapter 5.

Fake Barns Revisited

- Remember the original fake barns case? Classically there is a preamble in which Henry clearly knows he is seeing a barn.
  - Usual preamble: Henry drives by a barn (no facades around) and says "that's a barn". Does he know it is a barn? yes.
- What if instead the preamble gave a case where Henry clearly does *not* know he is seeing a barn.
  - Different preamble: Henry drives past a barn facade and mistakenly says "that's a barn". Does he know it is a barn? no! It isn't even true!
- Given Weinberg's result, might we have seen some 30 years of epistemology turn out differently?

Using Thought Experiments Properly

- ▶ Gendler still believes that thought experiments have a place in philosophy
- ▶ The fact that they can activate different cognitive systems from abstract reasoning is exactly what makes them powerful
- ▶ In particular, thought experiments may help us work around *first person exceptionalism*
  - ▶ The tendency to view oneself as special, exceptional, or even superior to others.
  - ▶ "among the most widespread and pervasive of our tendencies toward bias" (129)
  - ▶ This can cause us to be blind to the effects of our own actions

Using Thought Experiments Properly

- A well constructed thought experiment can help us re-evaluate our own actions by framing the situation in terms of other people and then making the parallels explicit
- This makes it easier to apply the non-first-person-exceptional intuition to even ourselves.
- These sorts of techniques show up in Kant, Rawls, and even the Bible
- Gendler thinks this is a good use of thought experiments

Thought Experiments Redeemed

Viewed in this light, moral and political philosophy have a secondary task that runs alongside the task of ascertaining what morality demands, namely, that of providing the reader with resources that enable her to make the perspective shift that the moral stance requires at the moment of moral decision-making. (...) It is this role, I want to suggest, that is played by some of the most famous thought experiments in moral and political theorizing. (131)

Wrapping Up

- We process content differently based on how it is presented to us
- For this reason we cannot expect to reach cognitive equilibrium about all problems and all foreseeable scenarios
- Thought experiments should therefore *not* be used as a means to find answers to philosophical questions as when we "search our intuitions"
- But they *may* be used to present (existing) arguments persuasively. Such techniques help us consider problems in the "right" way