

Examples of “Thought Experiments” in Science and Philosophy

Phil 488/594; Winter 2012

[Science examples taken from James Brown *The Laboratory of the Mind*; Philosophy examples taken from Julian Baggini *The Pig that Wants to be Eaten*.]

Probably the most famous thought experiment is Galileo’s “Falling Bodies” experiment.

The point of the thought experiment is to demonstrate that all bodies, regardless of their weight, fall at the same speed. It begins by noting Aristotle’s view that heavier bodies fall faster than light ones. We are then asked to imagine that a heavy cannon ball is attached to a light musket ball. What would happen if they were released together?

Reasoning in the Aristotelian manner leads to an absurd conclusion. First, the light ball will slow up the heavy one (acting as a kind of drag), so the speed of the combined system would be slower than the speed of the heavy ball falling alone. But on the other hand, the combined system is heavier than the heavy ball alone, so it should fall faster. Contradiction... so the Aristotelian theory of falling bodies is destroyed, and the only remaining solution is that they all fall at the same speed.

[In the dialogue, Galileo has the character Simplicio (the Aristotelian mouthpiece) say, “So you have not made a hundred tests, or even one? Any yet you so freely declare it to be certain?” And Salviati (Galileo’s mouthpiece) says “Without experiment, I am sure that the effect will happen as I tell you, because it *must* happen that way.”]

Does Salviati have an *intuition* that the tethered balls would fall at the same speed as untethered balls? If not an intuition, then what is it?

The old transporter issue!

For Stelios, the teletransporter is the only way to travel. Previously it took months to get from the Earth to Mars, confined to a cramped spacecraft with a far from perfect safety record. Stelios's TeletransportExpress changed all that. Now the trip takes just minutes, and so far it has been 100% safe.

However, now he is facing a lawsuit from a disgruntled customer who is claiming the company actually killed him. His argument is simple: the teletransporter works by scanning your brain and body cell by cell, destroying them, beaming the information to Mars and reconstruction you there. Although the person on Mars looks, feels and thinks just like a person who has been sent to sleep and zapped across space, the claimant argues that what actually happens is that you are murdered and replaced by a clone.

To Stelios, this sounds absurd. After all, he has taken the teletransporter trip dozens of times, and he doesn't feel dead. Indeed, how can the claimant seriously believe that he has been killed by the process when he is clearly able to take the case to court?

Still, as Stelios entered the teletransporter booth once again and prepared to press the button that would begin to dismantle him, he did, for a second, wonder whether he was about to commit suicide. . .

Before his last teletransporter trip, did Stelios have the intuition that teletransportation was *not* murder? And that it was *him* and not a clone who ended up on Mars? In the booth before the final teletransporter trip, was he having an intuition that teletransportation of oneself might actually be suicide? And that it was a *clone*, and not Stelios *himself*, who ended up on Mars?

Newton on “Absolute Space”

It is indeed a matter of great difficulty to discover the true motions of particular bodies from the apparent motions; because the parts of that immovable space by no means come under the observation of our senses. Yet the thing is not altogether desperate.

For instance, if two globes, kept at a distance one from the other by means of a cord that connects them, were revolved around their common centre of gravity, we might, from the tension of the cord, discover the endeavour of the globes to recede from the axis of their motion. . . And thus we might find both the quantity and the determination of this circular motion, even in an immense vacuum, where there was nothing external or sensible with which the globes could be compared.

But now, if in that space some remote bodies were placed that kept always position one to another, as the fixed stars do in our regions, we could not indeed determine from the relative translation of the globes among those bodies, whether the motion did belong to the globes or to the bodies.

But if we observed the cord, and found that its tension was that very tension which the motions of the globes required, we might conclude the motion to be in the globes and the bodies to be at rest.

Is there an intuition here? And if so, what is it?

Newton on “Absolute Space”, again

Imagine the rest of the physical universe gone, only a solitary bucket partially filled with water remaining. The bucket is suspended by a twisted rope <Don't ask what it's tied to> and released. As the rope unwinds we notice three distinct states of the water/bucket system.

In state I, at the instant the bucket is released, there is no relative motion between the water and the bucket; moreover, the surface of the water is level. In state II, shortly after the bucket is released, the water and the bucket are in relative motion – that is, in motion with respect to one another. (The bucket is “moving faster” than the water.) The water is still level in state II. We reach state III after some time has passed; the water and bucket are at relative rest, that is, at rest with respect to one another. But in state III the water is *not* level – its surface is now concave.

How do we account for the difference between state I and state III? In both states there is *no* relative motion of the water and bucket.

Newton's answer: In state I the water and the bucket are at absolute rest (i.e., at rest with respect to absolute space), and in state III the water and bucket are in absolute motion (i.e., in motion with respect to absolute space). The observed difference (flat vs. concave surface) can only be explained by absolute motion, and this requires absolute space.

Is there an intuition here? And if so, what is it? Isn't there something fishy going on about the cord and how it is attached to who-know-what??

Mary the Neurophysiologist

Mary knows everything there is to know about the colour red. As a scientist, it has been her life's work. If you want to know why we can't see infrared, why tomatoes are red or why red is the colour of passion, Mary is your woman.

All this would be unremarkable, if it weren't for the fact that Mary is an achromat: she has no colour vision at all. The world, for Mary, looks like a black and white movie.

Now, however, all that is to change. The cones on her retina are not themselves defective, it is simply that the signals are not processed by the brain. Advances in neurosurgery now mean that this can be fixed. Mary will soon see the world in colour for the first time.

As she walks out of her room she is amazed to discover that she in fact *didn't* know all there was to know about red. She now knows – *what red looks like*.

What do you think about this intuition that despite knowing all the *physical* things there are to know about red, there was still something further? That no scientific account of the world could give her this knowledge because no physical description of the world, however complete, can capture what goes on in our minds?

Think of Gettier-style cases when you read this thought experiment.

It was a very strange coincidence. One day last week, while Naomi was paying for her coffee, the man behind her, fumbling in his pockets, dropped his key ring. Naomi picked it up and couldn't help but notice the small white rabbit's foot dangling from it. As she handed it back to the man, who had a very distinctive, angular, ashen face, he looked a little embarrassed and said, "I take it everywhere. Sentimental reasons." He blushed and they said no more.

The very next day she was about to cross the road when she heard a screeching of brakes and then an ominous thud. Almost without thinking, she was drawn with the crowd to the scene of the accident. She looked to see who the victim was and saw that same white, jagged face. A doctor was already examining him. "He's dead."

She was required to give a statement to the police. "All I know is that he bought a coffee at that café yesterday and that he always carried a key ring with a white rabbit's foot." The police were able to confirm that both facts were true.

Five days later Naomi almost screamed out loud when, queuing once more for her coffee, she turned to see what looked like the same man standing behind her. He registered her shock but didn't seem surprised by it. "You thought I was my twin brother, right?" he asked. Naomi nodded. "You're not the first to react like that since the accident. It doesn't help that we both come to the same café, but not usually together."

As he spoke, Naomi couldn't help staring at what was in his hands: a white rabbit's foot on a key ring. The man was not taken aback by that either. "You know mothers. They like to treat their kids the same."

Naomi found the whole experience disconcerting. But the issue that bothered her most when she finally calmed down was: she had not told the police the truth!

Was Naomi's intuition that she had not told the police the truth correct? What was her reason for this intuition?

Einstein and Special Relativity

(Background info): Light is an oscillation in the electromagnetic field. A *changing* electric field gives rise to a magnetic field, and a *changing* magnetic field gives rise to an electric field. If a charge is juggled, it changes the electric field, which creates a magnetic field, which in turn creates an electric field, and so on. Maxwell's discovery: the wave travelling through the electromagnetic field with velocity c is light.

When he was only 16, Einstein wondered what it would be like to run so fast as to be able to catch up to the front of a beam of light. Would it be like running back along a pier to the beach at the same velocity as an ocean wave coming in? ... a hump in the water that is stationary with respect to the runner? It can't be like that since change is essential for a light wave – if either the electric or the magnetic field is static, it will not give rise to the other and hence there will be no electromagnetic wave.

(Einstein 1949): If I pursue a beam of light with the velocity c , I should observe such a beam of light as a spatially oscillatory electromagnetic field at rest. However, there can be no such thing, whether on the basis of experience or according to Maxwell's equations.

From the very beginning it appeared to me intuitively clear that, judged from the standpoint of such an observer, everything would have to happen according to the same laws as for an observer who, relative to the earth, was at rest. For how, otherwise, should the first observer know, i.e., be able to determine, that he is in a state of fast uniform motion?

One sees that in this paradox the germ of the special relativity theory is already contained.

Schrödinger and Quantum Mechanics

(Background info): In Quantum Mechanics, a physical system is represented by a vector ψ (the state vector or wave function) in a Hilbert space. Measurement outcomes correspond to the basis ('eigenvectors') of the space. The state of the system, however, may be a superposition of eigenvectors.

Measurement outcomes are always eigenvalues (= magnitudes that are associated with eigenstates and not with superpositions). Question: what happens when a measurement changes a superposition to an eigenstate? **Realist:** The physical system was always in one of the particular states – superpositions merely reflect ignorance of which. **Copenhagen Interpretation:** In a state of superposition, reality itself is indeterminate – measurements *create* the reality by putting the system into one of the eigenstates.

One can set up quite ridiculous cases. A cat is penned up in a steel chamber, along with the following diabolical device (which must be secured against direct interference by the cat): in a Geiger counter there is a tiny bit of radioactive substance, *so* small, that *perhaps* in the course of one hour one of the atoms decays, but also, with equal probability, perhaps none; if it happens, the counter tube discharges and through a relay releases a hammer which shatters a small flask of hydrocyanic acid. If one has left the entire system to itself for an hour, one would say that the cat still lives *if* meanwhile no atom has decayed. The first atomic decay would have poisoned it. The ψ -function of the entire system would express this by having in it the living and the dead cat (pardon the expression) mixed or smeared out in equal parts.

(Guess what? Schrödinger was a realist).

Judith Thomson (1971) as related by Brown. She argues for the moral permissibility of abortion in spite of granting (for the sake of the argument) that the fetus is a person with a right to life.

Imagine a great violinist who has some very unusual medical condition. There is only one cure, and it consists in being hooked up to *you* for nine months. Your biological makeup is the one and only one in the world which will help the violinist. In the night, unknown to you and unknown to the violinist (who is in a coma and will remain so for nine months, and thus is 'innocent'), he is attached to you by the Society of Music Lovers.

What should you do? Are you morally required to go through the nine months – an enormous sacrifice – to save the violinist's life? The answer, pretty clearly, is no. Yes, the violinist is an innocent person with a right to life and you are the one and only person in the world who can save the violinist's life. But you are not morally obliged to make the sacrifice (though you would be a moral hero if you did).

The analogy with the fetus is obvious. Thomson grants that it is an innocent person with a right to life and the pregnant mother is uniquely capable of bringing it to term. What the thought experiment does is distinguish two concepts which easily get run together: *right to life* and *right to what is needed to sustain life*. The fetus and the violinist have the former, but they do not have the latter. Having a right to life does not imply having a right to the use of another's body

Some related comments:

From Brown pp. 14-15:

Often in psychology or linguistics people are asked what they think about such and such. For example, someone might be asked to consider the sentence *Colourless green ideas sleep furiously* and to decide whether it is grammatically in order. Such a process, naturally enough, is often called a thought experiment. However, it is not a thought experiment as I am considering them here; rather it is a *real* experiment about thinking. The object of the psycholinguistic experiment is thought itself, whereas the object of a thought experiment (in my sense) is the external world and thinking is the *method* of learning something about it.

From Norman Daniels “Reflective Equilibrium” in *Stanford Encyclopedia of Philosophy*

The method of reflective equilibrium consists in working back and forth among our considered judgments (some say our “intuitions”) about particular instances or cases, the principles or rules that we believe govern them, and the theoretical considerations that we believe bear on accepting these considered judgments, intuitions, principles, or rules, revising any of these elements wherever necessary in order to achieve an acceptable coherence among them. The method succeeds and we achieve reflective equilibrium when we arrive at an acceptable coherence among these beliefs. An acceptable coherence requires that our beliefs not only be consistent with each other (a weak requirement), but that some of these beliefs provide support or provide a best explanation for others. Moreover, in the process we may not only modify prior beliefs but add new beliefs as well. There need be no assurance the reflective equilibrium is stable – we may modify it as new elements arise in our thinking. . . . We arrive at an optimal equilibrium when the component judgments, intuitions, principles, and theories are ones we are un-inclined to revise any further because together they have the highest degree of acceptability or credibility for us.

From Michael DePaul & William Ramsey “Introduction” in their *Rethinking Intuition*, 1998.

Numerous studies of the patterns of inductive inference that people use and judge to be intuitively plausible have revealed that people are prone to commit various fallacies. Moreover, they continue to find these fallacious patterns of reasoning to be intuitively acceptable upon reflection. Arguably, therefore, the rules of inductive inference that best capture our intuitive judgments, that is, the rules that are in reflective equilibrium for us, are simply unacceptable. Similarly studies of the “intuitive” heuristics ordinary people accept reveal various gross departures from empirically correct principles. Once again, there seems to be no reason to place a strong credence in intuition.

From Scholz, Pelletier, Pullum “Philosophy of Linguistics”
in *Stanford Encyclopedia of Philosophy*:

The preferred method of gathering data for linguistic theories by Essentialists [= Chomskian-influenced linguists] is informal elicitation, including elicitation from oneself. Many Emergentists [a particular type of anti-Chomskian linguistics] cast the debate in terms of whether linguistic intuitions should *ever* count as evidence for linguistic theorizing. And many Essentialists cast it in terms of whether anything *but* linguistic intuitions are ever really needed to support linguistic theorizing.

The debate focuses on the Essentialists’ notion of a mental grammar, since linguistic intuitions are generally understood to be a consequence of tacit knowledge of language. Emergentists who deny that speakers have innate domain-specific grammars or competence have raised a diverse range of objections to the use of reports of intuitions as linguistic data. For example, Tomasello says “The data that are actually used in Generative Grammar analyses are almost always disembodied sentences that analysts have made up ad hoc, rather than utterances produced by real people in real discourse situations.”

Furthermore, the practice of Essentialists of collapsing various kinds of acceptability judgments under the single label ‘intuitions’ masks possibly important differences... Collections of linguists’ reports of their own judgments are usually criticized by Emergentists as “arm-chair data collection,” or “data collection by introspection”.

On the other side, Essentialists tend to deny that usage data is adequate evidence. “. . . the standard methodology of hypothesis formation and empirical verification via judgment elicitation . . . has yielded good, replicable results, ones that could not have been obtained by using other data-gathering methods, such as corpus-based research.”

From Baggini's introduction:

The purpose of 'thought experiments' is to strip away the things that complicate matters in real life in order to focus clearly on the essence of a problem.

A real-life ethical dilemma will always be complicated by contingent, context-specific factors. Take the general issue of whether eating meat is morally wrong. When you consider the issue, multiple factors come into play. Some animals will have been factory farmed, some humanely reared, some caught wild. Some animals will have been raised on land that was once rain-forest, others will have freely grazed on open pastures. Some meat will be organic, some will be genetically modified, some will have been shipped from the other side of the world. Deciding the ethical rights and wrongs requires untangling all these multiple factors, and weighing up the different considerations accordingly.

Thought experiments can help because they aim to isolate the key variables, the specific factors under examination, to see what difference they, and they alone, make to our understanding of the world. So we imagine a case where the particular issue of concern is the only one to differ between scenarios. If we're worried about how we treat farm animals, let us imagine what difference good treatment, and good treatment alone, makes. If our intentions are under scrutiny, we can ask what difference does it make if the chicken in your kiev died in an accident whereas mine had its neck wrung intentionally, but before that they had lived identical lives.

We can simply stipulate that all other things are equal, so the only question we need to settle is the core moral one.

From a review of Daniel Kahneman *Thinking, Fast and Slow* (2011) by Freeman Dyson in *The New York Review of Books* (Dec. 22, 2011):

Intuition is the name we give to judgments based on the quick action of [primitive systems of the brain that make fast decisions]. This system makes quick judgments and takes action without waiting for our conscious awareness to catch up with it. The most remarkable fact about [this system] is that it has immediate access to a vast store of memories that it uses as a basis for judgment. The memories that are most accessible are those associated with strong emotions, with fear and pain and hatred. The resulting judgments are often wrong, but in the world of the jungle it is safer to be wrong and quick than to be right and slow.

Jennifer Nagel “Epistemic Intuitions” *Philosophy Compass* v. 2 pp. 792–819 (2007):

The expression ‘epistemic intuition’ is sometimes used very broadly, as a label for any immediate (or not explicitly inferential) assessment of any claim of interest to epistemologists; for example, philosophers sometimes speak of having an intuition that knowledge entails belief, or that knowing something entails that one couldn’t easily have been wrong about it. The nature and basis of these more general judgments is an interesting (and large) topic in its own right, as is the even broader question of the value of intuitions in philosophical debates outside of epistemology. The present article restricts its focus to epistemic evaluations of particular cases . . . Evaluations of the particular kind that will matter here occur when one is presented with a description of some real or hypothetical subject’s situation, where the description doesn’t already label the subject’s condition as one of knowledge, or justified belief, or whatever epistemological notion is being tested. Clear epistemic intuitions arise when the subject’s described condition plainly appears to fall on one side or the other of some significant divide in epistemology, such as the divide between knowledge and ignorance, or justified and unjustified belief.

Jesse Prinz “Empirical Philosophy and Experimental Philosophy” in J. Knobe & S. Nichols *Experimental Philosophy* (OUP, 2008).

How do we discover what our intuitions are? Presumably, we introspect. Intuitions are presented to us as mental states that become accessible to consciousness and available for reporting. An idea comes before the mind, along with a felt sense of confidence in that idea, and we report the result. Introspection can be described as a kind of observation. . .

. . . armchair conceptual analysis can be characterized as an introspective memory retrieval process. As such, it can be regarded as a form of observation. And this leads to the question, how good a form of observation is it. I don’t think it’s a bad method. But introspection is not necessarily the best method. First, . . . what we recall often depends on beliefs, expectations, norms, context, and other factors. It is prone to confabulation and distortion. . . We might assume that introspection reveals a rule that is already present in the mind, when in fact we are actually drawing inferences from specific cases in the course of what appears to be an innocent retrieval process. The results of introspection are often variable across individuals. In philosophy, debates often collapse into intuition mongering because defenders of opposing views are equally confident about conflicting intuitions. People in different cultural settings may have conflicting intuitions, but this variation is missed by contemporary Anglophone philosophers who are predominantly Western, highly educated, politically liberal, white and male.

This does not entail that we should give up on introspection. Far from it. But it does follow that we need to be cautious about philosophical intuitions, and we should be open to the use of other methods that help reveal the content of our concepts.

Jonathan Weinberg, Shaun Nichols & Stephen Stich “Normativity and Epistemic Intuitions” *Philosophical Topics* v. 29, pp. 429–460 (2001).

According to Epistemic Romanticism, knowledge of the correct epistemic norms is implanted within us in some way, and with the proper process of self-exploration we can discover them. . . .

There are various ways in which the basic idea of Epistemic Romanticism can be elaborated. The family of strategies that we want to focus on all accord a central role to what we will call *epistemic intuitions*. Thus, we will call this family of strategies *Intuition-Driven Romanticism* (or IDR). As we use the notion, an epistemic intuition is simply a spontaneous judgment about the epistemic properties of some specific case – a judgment for which the person making the judgment may be able to offer no plausible justification. . . .

The examples we have mentioned so far are hardly the only examples of IDR. Indeed, we think a plausible case can be made that a fair amount of what goes on in normative epistemology can be classified as IDR. Moreover, to the extent that it is assumed to have normative implications, much of what has been written in descriptive epistemology in recent decades also counts as IDR. For example, just about all of the vast literature that arose in response to Gettier’s classic paper uses intuitions about specific cases to test proposed analyses of the concept of knowledge.

Saul Kripke *Naming and Necessity* Harvard UP (1972/1980).
The “Gödel Case”.

Suppose that Gödel was not in fact the author of [Gödel’s] theorem. A man called “Schmidt”... actually did the work in question. His friend Gödel somehow got hold of the manuscript and it was thereafter attributed to Gödel. On the [descriptivist] view in question, then, when our ordinary man uses the name ‘Gödel,’ he really means to refer to Schmidt, because Schmidt is the unique person satisfying the description ‘the man who discovered the incompleteness of arithmetic.’ ... But it seems we are not. We simply are not.

Saul Kripke “The Jonah Case” (also in *Naming and Necessity*).

Suppose that someone says that no prophet ever was swallowed by a big fish or a whale. Does it follow, on that basis, that Jonah did not exist? There still seems to be the question whether the Biblical account is a legendary account of no person or a legendary account built on a real person. In the latter case, it’s only natural to say that, though Jonah did exist, no one did the things commonly related about him.

Plato's *Theaetetus* 201a–c, as retold by Jennifer Nagel
“Epistemic Intuitions” *Philosophy Compass* v. 2 pp. 792–
819 (2007).

When Plato's character Theaetetus proposes defining knowledge as true belief, Socrates tells him a story about a talented lawyer who had to defend a difficult case. Facing charges arising from some violent incident, the lawyer's client could provide him with little firm evidence concerning the event, and no neutral eye-witnesses to call. The client actually happened to be innocent. However, since the lawyer had limited time to present his case in court, he decided that charming the jury with rhetoric would probably be more effective than trying to instruct them about what really happened. This tactic worked well, and the jury became convinced that the defendant was innocent.

Responding to Socrates' story, Theaetetus readily grants that in this case the members of the jury do not *know* that the defendant is innocent, despite having a true belief on this point. Socrates and Theaetetus both take this case to refute the theory that knowledge is simply true belief.

If you also find it easy and natural to evaluate the jury in Plato's story as lacking knowledge, you can be described as sharing Theaetetus's epistemic intuition on this case.

□