

Philosophy's Great Experiment

Prospect Magazine - March 2009 |

Philosophers used to combine conceptual reflections with practical experiment. The trendiest new branch of the discipline, known as x-phi, wants to return to those days. Some philosophers don't like it

David Edmonds Nigel Warburton

David Edmonds is a research associate at the Oxford Uehiro Centre for Practical Ethics. Nigel Warburton is a senior lecturer in philosophy at the Open University. Together they make the podcast Philosophy Bites: [http:// www.philosophybites.com](http://www.philosophybites.com)

Katja Wiech is a cheerful young German researcher who is fascinated by pain. She's discovered many things—for example, when devout Catholics are given electric shocks while looking at a picture of the Virgin Mary they feel less pain than atheists do when administered the same unpleasant treatment.

She works in a set of rooms at the end of a maze of corridors in the John Radcliffe Hospital in Oxford. In one room sits a magnetic resonance imaging (MRI) scanner. The magnet of this machine is so powerful it can seize a mobile phone from your hand, sending it flying through the air.

Her subjects lie flat on the scanner's bed, their head inside its white tube. A computer by their feet provides various stimuli—images, questions and so on—and is operated from an adjacent room divided off by a glass screen. The noise is very loud. There's a panic button if her subjects freak out.

Wiech is a neurologist. But here's the strange thing: she is working with philosophers. The caricature of a philosopher is of an otherworldly professor sitting in a comfy armchair in an Oxbridge college, speculating on the nature of reality using only his or her intellect and a few books. This has some basis in reality. Chemistry requires test tubes, history needs documents. In recent years, the main tool of the philosopher has been grey matter. The subject's evolution can be painfully slow, tiptoeing forward from footnote to footnote. But not always. Every so often a new movement overturns the orthodoxies of received opinion. We might just be entering one of those phases.

A dynamic new school of thought is emerging that wants to kick down the walls of recent

philosophy and place experimentation back at its centre. It has a name to delight an advertising executive: x-phi. It has blogs and books devoted to it, and boasts an expanding body of researchers in elite universities. It even has an icon: an armchair in flames. If philosophy ever can be, x-phi is trendy. But, increasingly, it is also attracting hostility.

Philosophers have always been informed by scientific research, history and psychology. Indeed, most of the giants of pre-20th century philosophy combined empirical and conceptual studies. Some drew on the research of others, while René Descartes and John Locke performed their own experiments; this was a time when science had not entirely split from philosophy. David Hume mixed reason with experience, including psychological and historical observations alongside more abstract reasoning—A Treatise of Human Nature was subtitled “Being an Attempt to Introduce the Experimental Methods of Reasoning into Moral Subjects.”

But for many philosophers today the idea of experimental philosophy still grates. Conceptual analysis has been a dominant strain of Anglo-American philosophy in the past 100 years. Philosophy of this kind considers not so much how things are, but rather how we think about them: the way we carve up the world, the frontiers of meaning, of what makes sense. But for the x-phi fan, empirical research is not a mere prop to philosophy, it is philosophy.

Under the x-phi banner it's possible to distinguish three types of activity. The first uses new brain-scanning technology, for which philosophers teaming up with neuroscientists, like Katja Wiech, to look for patterns of neuronal activity when subjects are presented with philosophical problems. In the second type, philosophers devise questionnaires to discover people's intuitions and go out in the street with the trusty clipboard. In the third, they conduct field experiments, observing how people behave in particular situations, often without their knowledge. All three aim to test the philosophers' assumption that they know from introspection what people are likely to say or believe. The traditional philosophical assertion, “we have strong intuitions that...” or “we can all agree that...” now have to be tested against the evidence. The idea of who “we” are is being challenged, for instance by surveys suggesting broad cultural differences about intuitions. The philosopher in his Oxford study may not share intuitions with the shopper down the road in Queen's Street, whose intuitions, in turn, may differ from those in Queen's Road, Hong Kong. Such research raises big issues about our moral education.

It takes most people decades to reach guru status. But Joshua Knobe managed it within a few years of being awarded his PhD in philosophy from Princeton in 2006. He has an infectious

excitement for his research. In between his undergraduate and graduate days he published a few articles. One was about “intentionality”: when did people judge that behaviour was intentional? He and a collaborator tried to establish this by running a few experiments. Knobe says that his eureka moment occurred when a philosopher, Alfred Mele, responded to the article. Although he disagreed with Mele, the point was that Mele had “treated our work as a contribution to philosophy... I was too boneheaded to see for myself that the two disciplines [psychology and philosophy] could be brought together in this way.”

His work on intention soon attracted attention. Take one of his cases. A company chairman is told a new project will increase profits but **harm** the environment. He says, “I don't care about harming the environment. Let's start the new project. I just want to make as much profit as possible.”

Meanwhile another company chairman is faced with a similar choice, except this time it will **help** the environment. He says, “I don't care about helping the environment. Let's start the project. I want to make as much profit as possible.”

When asked whether the chairman intentionally harmed the environment in the first scenario, most people say “yes.” But did the chairman intentionally help the environment in the second scenario? Most people think not. This is weird. It led Knobe to conclude that people's moral judgements play a role in their concept of intentional action.

Another of Knobe's experiments—a collaboration with fellow philosopher Shaun Nichols—demonstrates x-phi's ambition, and how widely its methodology can be applied. The issue of free will is a perennial of western philosophy. Is the world entirely governed by causal laws and, if so, in what sense can humans be said to be free? Is moral responsibility compatible with a causally determined world? The range of possible responses is mind-bogglingly complex. But researchers, using surveys, now know what people think.

Unsurprisingly, perhaps, the majority of people turn out to be “non-determinists”—that is, they think that humans are free to choose. But science seems to reveal a world in which every event is explained in terms of prior causes and prevailing conditions, with no apparent room for free will. So, are we responsible for our actions even in a determinist world? Those who believe we are, and see no contradiction between our actions being causally determined and our having free will are known, in the jargon, as “compatibilists.”

Our rational response to determinism and free will is distorted; our emotional response leads us astray. If true, then they believe compatibilism loses some of its force.

In Oxford, philosopher Neil Levy is experimenting on a different topic, but along similar lines. Advances in medical science make it possible to improve ourselves both physically and mentally. Surveys show that we're more squeamish about cognitive changes: more of us object to using psychotropic drugs to improve our brain functioning (although we don't object to coffee), than, say, surgery such as tummy tucks or breast enlargements.

So why the difference? To access our intuitions, Levy and a collaborator use a technique called cognitive load, in which they ask questions while subjects are occupied on another task (such as subtracting three from 1,000, then three from 997, and so on). This supposedly allows a deeper and more accurate mapping of our intuitions because our cognitive capacities are otherwise engaged. The research is still at an early stage, but Levy suspects that most of us are intuitive dualists: we think mind and matter are distinct substances. This intuition contradicts the dominant view among philosophers of the mind, who believe (put simplistically) that there is no fundamental difference between mind and matter. If the philosophers are right and our intuitions just wrong it could change attitudes to cognitive and physical enhancements. Intuitions—even strongly held and apparently “natural”—can be misguided.

The 20th century saw an explosion in applied ethics—moral philosophers contributed to all manner of contemporary debates, from abortion to euthanasia, from the rules of engagement in war to the justification of punishment. In drawing out arguments, a traditional tool for the philosopher has been the thought experiment. These experiments tend to abstract from real cases to reveal the pertinent features of moral reasoning.

One of the most famous examples is the trolley problem. You are standing by a railway line when you see a train hurtling towards you, out of control; the brakes have failed. In its path are five people tied to the tracks. Fortunately, the runaway train is approaching a junction with a side spur. If you flip a switch you can redirect the train onto this spur, saving five lives. That's the good news. The not-quite-so-good news is that another person is tied down on the side spur of the track. Still, the decision's easy, right? By altering the train's direction only one life will be lost rather than five.

Call this Trolley A. Now vary the scenario a little. This time you're on a footbridge overlooking the railway track. You see the train hurtling towards you and five people tied to the rails. Can they be saved? Again, the moral philosopher has arranged it so they can. There's an obese man leaning over the footbridge. If you were to push him he would tumble over and squelch onto the track. He's so fat that his bulk would bring the train—Trolley B—to a juddering halt. Sadly,

the process would kill the fat man. But it would save the other five people. Should you shove him over? Again, apparently an easy decision. Surely you shouldn't. That would be an outrage. But what's the difference? Both cases involve killing one person to save five.

Philosophers have pondered this for over three decades. One possible explanation for our different intuitions in the two cases is this: in Trolley A, if you were to turn the train onto the spur and the person on the track were somehow to untie themselves and escape in time, you'd be delighted. Not only would you have avoided crashing into the five, but no one else would have got hurt. But with Trolley B, you need to lead the fat man to his death to save the five. It would be a noble sacrifice if the fat man jumped of his own accord. But if you push him you are using him as if he were an object.

The doctrine of double effect—which says that it may be acceptable to do something good when there is a foreseeable bad side effect, so long as this bad side effect is not intended—is much debated. The literature on runaway trains has become so vast that it's even been given a name of its own: “trolley-ology.” To an outsider it may all seem like harmless fun—crossword puzzles for philosophers. But it is designed to tease out whether we should ever sacrifice one person to save many and has numerous practical applications (for example, the issue of “collateral damage” in war).

Whether the doctrine justifies our conflicting intuitions over the trolley problem remains disputed. But what's interesting is the twist that experimental philosophers have brought to the debate. Trolley-ologists of the past assumed that their intuitions coincided with those of others, including non-philosophers—civilians, perhaps we should call them. But now there are easy ways to check. The BBC conducted an online poll in which 65,000 people took part. Nearly four out of five agreed that Trolley A should be diverted. Only one in four thought that the fat man should be shoved over the footbridge. (Nobody has yet looked for a link with the fact that nearly one in four Britons are obese.)

Neuroscientists and psychologists have also jumped on the trolleywagon. Brain scans allegedly indicate that when people are confronted with Trolley A, the part of the brain linked to cognition and reasoning lights up; whereas with Trolley B, people seem to use a section linked to emotion. The few people who are prepared to use the fat man as a buffer take longer to respond than those aren't, perhaps because they experience the emotional impulse and then reason their way out of it. Other experiments suggest people who have sustained damage to the prefrontal cortex, which is thought to generate various emotions, are far more likely than the rest of us to favour sacrificing the fat man.

Much of this work has been carried out in Harvard and Princeton. Meanwhile back in Oxford an Israeli philosopher, Guy Kahane, is poring over Katja Wiech's scans. On his monitor are images of the brain showing parts lit up like stars on a dark night. He has devised and tested subjects on a set of moral dilemmas and questions the Harvard findings. He is unconvinced that emotion is the driving force behind our judgments in these cases. But he too is using the tools of neuroscience and MRI scans to build his case.

There's a lot at stake. Peter Singer, the controversial utilitarian thinker and animal rights advocate, believes that while there are evolutionary explanations for why most of us recoil from pushing the fat man, reason should lead us to overcome our squeamishness. For him, there is no overriding moral difference between the two trolley examples, or between intentionally killing civilians in war and their deaths as a byproduct of a military objective. Other philosophers strongly disagree. If x-phi research could settle this debate, it would be quite an achievement.

Using state-of-the-art gadgetry to cast light on philosophical mysteries sounds like a breakthrough, and grand claims are being made on the basis of neuroscientific observations. But Raymond Tallis, a philosopher and medical scientist who used MRI machines for years to study strokes and epilepsy, is not so sure. He thinks that the accuracy and relevance of brain scanning has been overestimated. MRI technology is excellent for investigating physical damage to the brain, Tallis explains, but when it comes to more complex matters, such as localising particular thought processes, it is too crude. The data from these scans, for example, reflects average activity. When a section of a brain is illuminated this is because it is operating at a heavier load than usual compared with other areas. Changes happening over the whole brain are not picked up. And even sophisticated neural imaging cannot distinguish between physical pain and social rejection—they “light up” the same areas.

There's a more fundamental problem still, says Tallis. The magnetic tube can never replicate the real world—so answers given inside it are of limited value in predicting decisions that would be taken outside. The hypothetical scenarios presented to volunteers are ingenious but implausible. Even when suspending disbelief, subjects are not gripped by the same panic, indecision, fear and anguish that genuine moral dilemmas produce. Real decisions depend on the particular situation; ethical choices are not like T-junctions, where there are only two choices.

Some philosophers quietly dismiss the movement as a cynical step by researchers to appear

cutting edge and to tap into scientists' funding. Interdisciplinary research can be a shrewd career move: it can, as Tallis notes, allow you to "rise between two stools." David Papineau, professor of the philosophy of science at King's College London, says that philosophers who want to know about the real nature of categories like mind, free will, moral value and knowledge should on occasion abandon their armchairs and pay attention to relevant findings. But that doesn't mean that they should be in the street handing passersby questionnaires: "I don't see that they'll learn anything worthwhile from asking ordinary people what they think about these things."

A philosophical problem is not an empirical problem, a fact is not an interpretation, an "is" is not an "ought," a description of how we actually behave and think is not a rationale for how we should behave and think. Yet despite the critics, the clipboards and scanners are multiplying, with sometimes surprising effects on ancient debates.

In the past few decades there has been a renewed interest in Aristotelian ethics and the notion that ethics is a matter of cultivating virtue. Many recent papers in moral psychology stress the ways situations and unconscious influences affect what we do. These seem more reliable predictors of our actions than our underlying character. There's a link here with behavioural economics, which stresses our irrational and often hidden impulses.

Moral philosophy appears to be especially fertile ground for combining the conceptual and the empirical. Princeton philosopher Kwame Anthony Appiah, in his recent book *Experiments in Ethics*, cites some experiments demonstrating the degree to which situations affect how we behave. Aristotelian virtue theorists stress consistency across situations: an honest person is likely to be honest when presented with different temptations in different circumstances, a compassionate person compassionate wherever appropriate, and so on. Is this, though, the way things really are? Empirical research suggests not. People asking for change for a dollar got a much better response outside a pleasant-smelling bakery than a neutral-smelling hardware store; unwitting subjects in an experiment who found a dime in a phone booth were far more ready to help someone pick up dropped papers than those who hadn't had that tiny piece of good luck.

Situations have a bigger influence on how we behave than we think they do. Perhaps, then, rather than worrying so much about character building in an Aristotelian vein we should be making people more aware of how easily apparently irrelevant factors can shape what we do. As Appiah asks: "Would you rather have people be helpful or not? It turns out that having little nice things happen to them is a much better way of making them helpful than spending a huge amount of energy on improving their characters."

Is this all a storm in a common room? The repercussions of the experiments cannot be so easily dismissed. Think of the impact on political liberalism. At the heart of liberalism is the idea that an educated adult is and should be capable of choosing how he or she lives. But if, for example, situations affect us more than the reasons we give for our actions, and we use those reasons to rationalise them retrospectively, this assumption may need revision. This branch of x-phi might be nudging us towards Nietzsche's view that what we take to be the inexorable conclusions of clear rational thought are nothing but reformulations of our innermost desires—disguised as the products of logic. We are not as in control of our thoughts as we thought. Nietzsche fully grasped how profoundly unsettling this notion was.

Experiments in moral psychology may be making back-to-Aristotle ethics less plausible. But in another sense, the experimental philosophy enterprise is eminently Aristotelian. In Raphael's famous painting, *The School of Athens*, Plato points up to the otherworldly realm: true reality, the world of the suprasensual Forms that can be understood only by pure thought. Aristotle, however, is reaching out to the world in front of him. X-phi looks like it's here to stay, and contemporary philosophy should surely take notice.

Discuss this article at [First Drafts](#), Prospect 's blog

Questions:

1. Is experimentation new to philosophy? Explain.
2. What are the three main branches of x - phi?
3. Did each chairman intentionally harm the environment?
4. Are you a compatibalist? Explain.
5. How does 'cognitive load' help brain imaging? How may this affect ontology?
6. How would you decide to act in the Trolley problem?
7. What parts of the brain light up for Trolley A and Trolley B?
7. How does the 'doctrine of double effect' apply to war?
8. What does Peter Singer think?

9. Briefly explain three criticisms of x- phi.

10. What does Aristotle's virtue ethics stress. What do experiments say about this?