

Ricerche

Are we Living an Illusion? Folk Intuitions on the Problem of Free Will

Silvia Felletti

Ricevuto: 22 novembre 2014; accettato: 24 marzo 2015

Abstract In recent years, philosophy has witnessed the birth and development of a new research program that has provoked both enthusiasm and strong criticism: *Experimental Philosophy* (or *X-Phi*). In this contribution, I will briefly examine the new field of experimental philosophy, its purposes and methodologies. I will then summarize some of the objections that have been raised against this research program, and the arguments with which experimental philosophers have used to counter these objections pointing to the usefulness of their studies for philosophy in general. I will then focus on a specific subject that has drawn the attention of experimental philosophers: the problem of free will, its compatibility with determinism, and the related issue of the role of intentionality in action generation. Free will is one of the most ancient and debated problems in philosophy, and also one of the issues on which experimental philosophy methods can be most fruitfully applied, the concepts of free will and of moral responsibility being pervasive in everyday life and grounded in common sense.

KEYWORDS: Free Will; Determinism; Experimental Philosophy; Folk Intuitions; Lay-theories.

Riassunto *Vivere un'illusione? Le intuizioni di senso comune sul libero arbitrio* – Recentemente la filosofia ha vissuto nascita e sviluppo di un nuovo programma di ricerca che ha suscitato entusiasmo e forti critiche: la *filosofia sperimentale*. In questa sede intendo illustrare la filosofia sperimentale, le sue finalità e i suoi metodi. Cercherò poi di riassumere alcune delle obiezioni sollevate contro questo programma di ricerca e gli argomenti dei filosofi sperimentali per replicare alle critiche e sottolineare l'utilità dei loro studi per la filosofia in generale. Mi concentrerò poi su un particolare tema che è all'attenzione dei filosofi sperimentali: la questione del libero arbitrio, la sua compatibilità con il determinismo e il ruolo dell'intenzionalità nella genesi dell'azione. Il libero arbitrio è uno dei più antichi e dibattuti problemi della filosofia e anche una delle questioni cui meglio si possono applicare i metodi della filosofia sperimentale, poiché libero arbitrio e responsabilità morale, concetto a esso collegato, sono molto presenti nella vita quotidiana ed estremamente radicati nel senso comune.

PAROLE CHIAVE: Libero arbitrio; Determinismo; Filosofia sperimentale; Intuizioni di senso comune; Teorie di senso comune.

S. Felletti - ISTC - Istituto di Scienze e Tecnologie della Cognizione - CNR, via San Martino della Battaglia, 44 - 00185 Roma (I) E-mail: silvia.felletti@istc.cnr.it (✉)



Experiments in philosophy

PHILOSOPHICAL ANALYSIS OFTEN REVEALS CONFLICTING intuitions about given issues. We seem to have good reasons to support a thesis, but also other equally strong reasons to sustain the opposite.

When facing such complex situations, philosophers have a long-standing tradition of resorting to thought experiments: they build imaginary scenarios expressly designed to elicit intuitions about a certain philosophical concept or issue, in order to use them as evidence in support of their hypothesis, or to undermine hypotheses which they take to be inadequate. The history of philosophy is full of famous thought experiments: Plato's allegory of the cave, the Cartesian evil genius, Searle's Chinese room, Chalmers' zombies, Putnam's brain in a vat, to name but a few. Dennett¹ has referred to them as "intuition pumps" – not without a hint of irony, with regard to some cases –, for these thought experiments can be very helpful in reasoning on philosophical problems or in understanding complex concepts, by transposing them in specific, exemplary cases.

There is, however, a potential problem with such a strategy: when philosophers reflect on a thought experiment, they often tend to consider their own intuitions as representative of common intuitions.² That is, they regard their intuitions as psychological universals, shared by the majority of people capable of correctly understanding the imaginary situation described in a thought experiment. Following Jackson,³ many philosophers assume that their intuitions can reflect the ordinary conception to the extent to which they are allowed to consider themselves "typical".

However, claims about supposed "common intuitions" could indeed be unwarranted. In recent years, this worry has been one of the major impulses for the birth of experimental philosophy, a research program that includes several different areas of research with the common objective of field-testing

folk intuitions and lay-theories about philosophical issues. Experimental philosophers (e.g., Joshua Knobe, Shaun Nichols, Thomas Nadelhoffer, Eddy Nahmias), subject people's intuitions on traditional philosophical issues such as consciousness, action, intentionality, or morality, to empirical testing, using tools that are typical of empirical research: scenarios, surveys, statistical analysis.

To those who may ask: "Where does the difference lie with social psychology?", an experimental philosopher will readily answer that there must not be any rigid dividing line between psychology and philosophy. Instead, the aim is to combine the tools of both the disciplines to develop interdisciplinary approaches to problems which are of both philosophical and psychological interest. Experimental philosophy studies are often co-authored by philosophers and psychologists, and published in philosophy as well as in psychology journals.

In their critique of the armchair philosophers' use of (what they consider to be) folk intuitions to sustain their theses, experimental philosophers hold that, in order to figure out what people find intuitive, the only legitimate way is to go and ask them. Otherwise, philosophers may often consider whatever their students or they themselves happen to find intuitive as "folk belief", with no guarantee of generality.

Experiments conducted so far in the field of experimental philosophy have shown that philosophers have often been wrong in their claims about how people think about given philosophically relevant problems – both because laymen may have different intuitions from experts (i.e., philosophers), and because laypersons may have different intuitions concerning a given issue, due to a variety of factors (e.g., culture, gender, social pressures).

A typical experimental philosophy study will provide subjects with the description of particular scenarios especially designed for eliciting certain lines of thought, and will then ask participants questions that can be

reasonably taken to be indicative of, or express, their relevant intuitions on the given philosophical issue. This is often achieved by assessing and measuring the differences on people's intuitions that result from small changes either in the description of the scenario or in the way of putting the question. The results of the studies are analyzed statistically, and the differences in people's inclinations are explained in terms of psychological processes.

In a well-known study on intentionality in action,⁴ Joshua Knobe developed two versions of the same scenario representing the same actor doing the same action (a chairman deliberately trying to raise the profits for his company): in one version, the action had a positive side effect (improving the environment); in another version, it had a negative side effect (harming the environment). In both cases, the chairman was said not to care about whether the consequences of his actions were positive or negative. When asked about the intentionality of the effects on the environment, people showed two radically different patterns of response: in the harm condition, most of them (82%) said that the chairman brought about the side effect intentionally, whereas in the help condition only 23% of the subjects attributed intentionality to the positive side effect.

Also, the attribution of blame in the harm condition was significantly greater than the attribution of praise in the help condition. Since the side effect was the only element varying between the two scenarios, it very likely corresponds to the source of the discrepancy between intuitions. This phenomenon is known (even if not accepted without reserve) as the *side-effect effect*: when people have to attribute intentionality to the side effects of an agent's action, they tend to be influenced by the moral quality of the outcome in question.⁵

Experimental philosophy on free will

Let us now examine more closely one of

the issues which experimental philosophy has dealt with the most: free will, a concept that is both controversial and grounded in our common-sense.⁶ The problem of free will arises from two contrasting and yet deeply-rooted human beliefs: on the one hand, it seems natural to us that when we make decisions about alternative future behavior, we can typically decide either way, only if we want to.

Our acts follow our decisions coherently, and then we perceive those decisions as the main, if not the only cause of our behavior. On the other hand, it seems equally intuitive to us that every occurring event can be fully explained, or is entirely caused, by the causal chain of the past events that led to it. This second idea corresponds roughly to the philosophical concept of *determinism*,⁷ and even if it is an intuitively attractive idea and a largely accepted position in science, it clashes with the first intuition that it is possible for us, given the same past conditions, to choose one way or the other. Hence it seems that we have to opt for one of these two intuitions: either we reject determinism and preserve free will, or we endorse determinism and acknowledge free will as an illusion.

The causal role of intentions

The problem of free will can be hard or maybe even impossible to solve, for two main reasons. The first is that the existence of free will depends largely on our own ideas of it. Free will has been variously described as freedom from internal or external constraints,⁸ as the ability to choose between alternative possibilities or to do otherwise,⁹ as ultimate responsibility,¹⁰ as freedom to act on the basis of our beliefs and desires.¹¹ These concepts have been used in different ways to support or reject the thesis that we possess a free will.

The second problem is that even if we came to an agreement about the true nature of free will, which is unlikely to happen, it is doubtful that we could prove or disprove its

existence. Let us suppose, for the sake of argument, that we convene on the classical and widely accepted view of free will that the condition for us to be free is that we can act on the basis of our intentions and desires.

In this view, we are only free if our (conscious) intentions and desires can cause, or at least have a role in causing, our behavior, even if we assume desires to be determined. However, if action arises (only) from internal or external factors that are out of our control, then we cannot consider ourselves free. That is why so many scientists have deemed Libet's experiments on the timing of conscious will¹² suitable to sanction the death of free will (most notably, Daniel Wegner).¹³ By showing that the awareness of one's intention to move a finger arrives after the moment of the onset of the unconscious cerebral activity that leads to motor action, Libet's experiment proves that

the brain "decides" to initiate or, at least, to prepare to initiate the act before there is any reportable subjective awareness that such a decision has taken place.¹⁴

Then, our will can just be an illusion, not a cause of action but merely a «personal conscious feeling of such causing, forcing, or motoring».¹⁵ But the causal inefficacy of conscious intentions may be hard to prove: in fact, many have claimed that Libet's experiment only partially succeeded in accomplishing its task.

I intend to focus on a particular line of criticism¹⁶ that is based on a well-known distinction in the literature: the distinction between proximal and distal intentions. This argument states that Libet's experiment does not supply a valid proof against the possibility of a causal role of conscious intentions, as it only shows the inefficacy of intentions formed immediately prior to a motor behavior, that is, of *proximal intentions*. But Libet's experiment does not tell us anything about the role, if any, of *distal intentions*, that is, of intentions built far in advance of action. This

issue is particularly significant, especially in relation to our phenomenology of intentional action, since planning behavior and making general arrangements for the future is the most intuitive way in which we believe that conscious intentions can cause our behavior. Following this line of criticism, the problem of proving the existence or inexistence of free will boils down to proving the efficacy or inefficacy of conscious distal intentions.

In order to investigate people's perception of the role of consciousness in action generation, myself and Paglieri¹⁷ conducted a study that surveyed folk intuitions about proximal and distal intentions using the methods of experimental philosophy. Our hypothesis was that the early planning of intentions to act plays a fundamental role in determining ordinary intuitions about voluntary action. To test this hypothesis, we submitted different scenarios to participants describing actions where the two stages of distal and proximal formation of intentions were manipulated to appear either freely willed or rather externally and unconsciously determined. We found that the conscious vs. unconscious nature of intentions affected intentionality and responsibility attributions at both stages of action preparation, but the effect was significantly more marked for distal intentions than for proximal ones, thereby confirming the centrality of long-term planning in folk intuitions of intentionality.

Even for those who do not consider empirical data to be binding for philosophical theories, these results highlight a general point on the role of folk intuitions in philosophy. Claiming that a certain mental phenomenon, e.g. free will, is an illusion requires proving that people do not exhibit that particular trait in the way they think they do: this in turn demands an empirical grasp on folk intuitions regarding that phenomenon. In this case, if philosophers want to prove the illusory nature of free will, they must provide a theory (and, for empirically informed philosophers, also an empirical proof) of the inefficacy of distal intentions.

Compatibilism vs. Incompatibilism

Experimental philosophy studies have provided evidence for the claim that most people tend to reject the idea that our universe is deterministic.¹⁸ More precisely, they accept determinism when it is related to inanimate objects, but are not prone to do the same when it comes to agents, even at an early age.¹⁹

But what do people think about the relationship between free will and determinism? Assuming that people were to accept the truth of determinism, how would their belief in free will be affected? Those who feel the need to choose between these two positions are known as incompatibilists, incompatibilism being the thesis for which free will and determinism cannot be coherently accepted simultaneously. Compatibilists, on the other hand, claim that we could live in a deterministic universe and yet retain, in some meaningful way, our free will.

Several philosophers have claimed that most people, when faced with this problem for the first time, are inclined to be incompatibilists. Incompatibilist philosophers who defend this position claim that their view is commonsensical, while compatibilism is counter-intuitive. For example, Kane affirms that

most ordinary persons start out as natural incompatibilists. They believe there is some kind of conflict between freedom and determinism [...] Ordinary persons have to be talked out of this natural incompatibilism by the clever arguments of philosophers – who are only too happy to oblige.²⁰

Strawson writes that the libertarian form of incompatibilism, although impossible to sustain, is the spontaneous belief of common people who believe that things happen because of preceding causes, but also that we are irreducibly free and able to do otherwise, given the same conditions.²¹

On the other hand, compatibilists have also appealed to commonsense intuitions, claiming that people do not necessarily think about freedom in terms of the possibility of doing otherwise, as libertarians do. Frankfurt-style cases²² are especially designed to bring out the intuition that we can be free even if we are not able to act differently than we do, that is, simply acting according to our fully determined intentions and desires. Since compatibilists and incompatibilists cannot both be right in claiming that their theories match the pre-philosophical intuitions of the majority of people, there is of course a need to systematically test what is really the intuitive position on the issue, if any at all.

One of experimental philosophers' preferred ways to do this is presenting people with specific cases of deterministic scenarios and checking whether they consider freedom or responsibility as possible in those situations. To ensure that the observed intuitions are genuinely "folk", as opposed to philosophically informed, this type of survey must target people who lack specific knowledge of the relevant philosophical arguments about free will, determinism and moral responsibility.

Negative responses on the possibility of freedom and responsibility²³ in a deterministic scenario would indicate that subjects perceive a conflict between them, and thus consider determinism and freedom of will as incompatible, whereas positive responses would indicate a compatibilist position. Manipulating the scenarios and the questions addressed to the respondents could allow experimenters to identify the psychological factors that affect people's intuitions.

In a study by Nahmias and colleagues,²⁴ participants read about a deterministic scenario in which a supercomputer can predict future events on the basis of its knowledge of the past and all the laws of nature. The supercomputer then predicts that at a certain time a man named Jeremy will rob a bank. The prediction turns out to be correct, and Jeremy indeed robs the bank at the time indi-

cated by the supercomputer. Participants then were asked to judge whether Jeremy was morally blameworthy for what he had done: 76% of them said that he was.

To ensure that this result was not influenced by the negative moral valence of the action in question – that is, by the fact that we might be more inclined to assign responsibility to agents that are taken to be blameworthy – Nahmias and colleagues repeated the same test with a positive action (saving a child from a burning building) and with a neutral action (going jogging). In both cases, responses were similar to the first test: 68% of participants said Jeremy saved the child of his own free will, and 79% said he went jogging of his own free will.

Data from this study thus suggest that most ordinary people do not find incompatibilism intuitive or correct, and this, in the authors' opinion, undermines the claim that incompatibilism is the "natural" way of thinking about determinism and free will.

These results might yet be challenged in two regards: first, laypeople may be confused by the technical description of the deterministic reality, described by expressions such as "laws of nature" and "current state of the universe". Second, determinism is here depicted as predictability, which many philosophers do not consider to be as threatening to free action as, for example, causation is.

To solve these problems, Nichols and Knobe²⁵ conducted a similar study by building a scenario in which determinism was described as causal inevitability, in simple and non-technical terms. Their hypothesis was that most people really do have an incompatibilist position about determinism and free action, but that this position reveals itself only in hypothetical and theoretical reasoning. However, when faced with real, concrete actions, people let their affective reactions drive their judgments. In the experiment, each participant was given the description of a deterministic and an indeterministic universe (Universes A and B).

In Universe A, everything that happens is

completely caused by what happened before, including agents' decisions, from the beginning of time until present day. In Universe B, causation involves everything except human decisions, that are free and do not *have to* happen. Subjects were first asked whether they thought our own universe was more similar to Universe A or B: nearly all of them (over 90%) answered that the second universe was most like ours.

In the second question of the survey, participants were assigned either a concrete or an abstract condition. In the concrete condition, they were told that in Universe A lives a man who has killed his whole family in order to have a relationship with his secretary. Subjects were asked to say if this man is responsible for his actions: 72% of the respondents gave the *compatibilist* answer, saying that the man is fully morally responsible for his acts. In the abstract condition, the survey only asked whether it was possible for a person to be fully morally responsible for their actions in Universe A. In this condition, 86% of the subjects answered in the negative, thus endorsing an *incompatibilist* position.

According to the authors, this study indicates that one of the factors influencing people's contrasting intuitions on free will and determinism could be the abstract or concrete nature of the question, and the related level of emotion aroused by the story. Abstract problems trigger theoretical, "scientific" reasoning that drives people to draw the conclusion that determinism is incompatible with freedom and responsibility. On the other hand, real situations, especially when emotionally laden, lead us to assign praise and blame to actions, and then responsibility and freedom of action to the agents.

But why is this so? What are the psychological mechanisms underlying such discrepancy in judgments?

Nichols and Knobe identify two possible explanations that could account for the experimental data. The first is that strong affective reactions can cause a bias in people's judgments of morality and responsibility. In

this case, people do have a tacit theory of morality, and they can apply its criteria efficiently in ordinary situations; but when strong violations of the moral norms that people implicitly accept are at stake, the resulting involuntary emotional reaction prevents them from using this resource. Affective activation is considered here as a disturbance in the correct reasoning for the attribution of responsibility.

Another possible explanation is that affect is not just a source of bias, but it is also one of the essential elements at the basis of the process by which we attribute praise and blame. In this sense, the cold, “scientific” theories of morality are not involved in our everyday-life attributions of moral responsibility.

Further studies are needed to establish which of these two alternative approaches (or a third one yet to be proposed) better explains the evidence. This discussion, however, exemplifies a fundamental question raised by experimental philosophy: how should we regard the intuitions elicited by the strong, emotional cases? Should we consider them as biased judgments, not reflecting the actual beliefs of laypeople, or should we take them as indicators of the folk conception of free will?

Researchers are divided on this point. Nichols and Knobe think that the intuitions in the high-affect cases are the result of a performance error due to the interference of emotions: when not biased by emotions, people tend to be libertarian about human choice, as they think human choices must not be determined in order for human behavior to be free. Eddy Nahmias, however, pleads differently.

Mechanical incompatibilism: Determinism vs. physicalism

Nahmias claims that the pre-philosophical position on free will and determinism is more often compatibilist, and that performance errors only occur when people mis-

takenly assume that determinism entails the inefficacy of the agents’ conscious deliberations and intentions. Even if it may seem that people have incompatibilist intuitions, this is only because they systematically confuse determinism with reductive mechanism (i.e., physicalism) – often because philosophers themselves conflate the two concepts in the first place.²⁶

While determinism only entails that the present state of a system is a consequence of its past states and of the laws of nature, reductive mechanism does not leave room for mental explanations of behavior: when people are led to think of humans as mechanisms, they are no longer able to apply an intentional stance to them, thus they cannot see their actions as fully intended and free. Nahmias and colleagues’ hypothesis is that the alleged intuitive incompatibilism is not a *pure incompatibilism* between free will and determinism, but only a *mechanical incompatibilism* between free will and physicalism. This is fully understandable given the intuitive nature of the conflict between a physicalist conception of the human brain and a mentalistic explanation of human behavior in terms of beliefs, desires and intentions.

Since some of the descriptions of deterministic scenarios in the surveys seem to suggest that our behavior is caused by forces that bypass our conscious mental life, incompatibilist intuitions might just arise because people

fear that certain reductionistic or mechanistic descriptions of decision-making and action conflict with freedom and responsibility.²⁷

But, as Nahmias points out, determinism does not entail fatalism, coercion by natural forces, reductionism or epiphenomenalism. Determinism is not necessarily the thesis that our brain mechanisms cause everything we do, while our mental states are just phenomenological noise. And, on the other hand, mechanism still allows indeterminism, as

causality could be explained through the laws of quantum physics. Determinism and mechanism are thus two distinct metaphysical concepts.

So, Nahmias' hypothesis is that ordinary people's apparent incompatibilism derives from the threat posed to free will by physicalism rather than by determinism, and arises when the descriptions of determinism erroneously suggest a bypassing of our conscious mental life by forces out of our control. Determinism *per se*, says Nahmias, does not constitute a threat for free will in the mind of people; but if people understand determinism in terms of mechanical or reductionistic causation, then it is not surprising if they turn out to be incompatibilists.

To test this hypothesis, Nahmias²⁸ developed an experiment in which two different scenarios corresponding to two different concepts of determinism were handed out to subjects divided in two conditions. In the first scenario, a universe where human behavior was the result of psychological states of mind was described, while the other represented a universe where the causes of human behavior were mechanical events in the brain. Both scenarios were clearly and explicitly deterministic in their description, but only the second one was also physicalist.

Participants in the two conditions were then asked whether they thought the agents living in the two universes (Ertans) could act of their own free will, and if they should be given credit or blame for their actions. Only few of the participants who read the neuro-reductionistic deterministic scenario thought that Ertans acted of their own free will (18%), and only 19% said that they deserve credit or blame for their actions, while a vast majority of those who read the non-reductionistic deterministic scenario judged Ertans to be free (72%), and attributed moral responsibility to them (77%).

Nahmias considers these data as evidence that people do not always regard causal determinism as a threat to freedom and responsibility. Instead, what they do find threaten-

ing is physicalism, since it suggests that our behavior is caused by forces that bypass our conscious mental life.

■ A new study on mechanical incompatibilism

In a recent study I conducted together with Fabio Paglieri²⁹ we tried to replicate Nahmias' results on mechanical incompatibilism introducing some methodological modifications to his design, as part of a broader study on folk intuitions on intentionality attributions. Here only our data on compatibilism vs. incompatibilism will be discussed.

The research included a pilot study on 51 subjects (23 females) recruited from a large metropolitan area, and a main study which included 200 participants recruited among the Roma Tre University students (135 females, with age between 16 and 66 years, mean: 22.19). The test included a first series of scenarios designed to prove intuitions on the role of conscious will in distal and proximal intentions; a questionnaire to assess the personal degree of impulsiveness (Italian version of the Barratt Impulsiveness Scale, see Patton, Stanford and Barratt, 1995),³⁰ and finally, a second series of scenarios to survey intuitions on mechanical incompatibilism.

In this last task, we assessed participants' intuitions on the relationship between determinism and free will, in terms of compatibilism vs. incompatibilism in the two conditions of deterministic vs. physicalist reality. We used the following double-universe scenarios, inspired by Nahmias' experiment, with only minor modifications³¹ and translated in Italian:

Imagine that there is a universe similar to ours, in which there are two planets, called Neuroearth and Psychoearth, that are similar to our Earth in many respects. On these planets, the landscape and life are very similar to ours, and there are advanced life forms (respectively, Neu-

roearthlings and Psychoearthlings) who look, talk, and behave very much like we do. However, the science of these planets has advanced far beyond ours: on Neuroearth there is an advanced knowledge in the field of neuroscience, on Psychoearth in that of psychology. The two planets are very far apart, and their inhabitants have no contact with each other, nor are they aware of each other's existence.

On Neuroearth, neuroscientists have discovered exactly how Neuroearthlings' brain works. They have found that every single decision taken and action performed by Neuroearthlings is completely caused by the particular chemical reactions and neurological processes that occur in the brain at that time, and that these chemical reactions and neurological processes in the brain are caused entirely by preceding events, including Neuroearthlings' genetic heritage and the environment in which they have lived. So, every action performed by Neuroearthlings is completely caused by the particular chemical reactions and neurological processes that occur in their brain at that time, and these neural events are in turn caused by genetics and environment.

On the other hand, on Psychoearth, psychologists have discovered exactly how Psychoearthlings' mind works. They have found that every single decision taken and action performed by Psychoearthlings is completely caused by the particular thoughts, desires and plans they have at the moment, and that these thoughts, desires and plans are completely caused by earlier events, including Psychoearthlings' genetic inheritance and the environment in which they have lived. So, every action performed by Psychoearthlings is completely caused by the particular thoughts, desires and plans they have at the moment, and these mental events are in turn

caused by genetics and environment.

While Nahmias presented different groups of participants with only one deterministic scenario, either physically reductionistic or non-reductionistic (between-subjects design), we submitted both scenarios to our participants, and probed their intuitions on both of them, with a within-subjects design. Our purpose was to ensure the *robustness* of the original result: presenting both the physically reductionistic (Neuroearth) and simple deterministic (Psychoearth) version within the same scenario makes it clear that they are equally deterministic, and thus should reduce any tendency to answer differently to them, if what matters is determinism. If, on the other hand, as Nahmias argues, what matters is physicalism, then the difference noted in the original study should emerge also in our within-subjects design.

Another methodological difference regarded how questions were formulated: in addition to the questions about free will and responsibility of Neuroearthlings and Psychoearthlings,³² our subjects were questioned on the plausibility of the scenarios – namely, they were asked whether they thought it would be possible for us to obtain the same complete knowledge of the mind or brain as described in the scenarios.

Moreover, participants in our study responded on a 5-point Likert scale (ranging from “certainly not” to “absolutely yes”, plus “I don't know”) instead of giving a yes/no/I don't know answer as in Nahmias' original study.³³ This was done to test the original result in a more challenging version of the experimental design, that allow to prove the robustness of mechanical incompatibilism with graded judgments instead of discrete choices (intentional/unintentional, responsible/non-responsible).

We tested 200 participants (mean age = 22.19, N of females = 135) recruited in Rome among university students. Participants had no professional knowledge on philosophical issues of free will and determinism.³⁴ Partici-

pation in both studies was on a voluntary basis, and subjects did not receive any compensation for taking part to the experiment. Testing took place in November and December 2013. Participants were tested right before class, while seated in medium sized groups within a large lecture room: the experimenter and the teacher made sure no communication went on during testing, and participants located next to each other always received different versions of the test with a different order of presentation of the scenarios, to minimize mutual influence. The order of presentation of the descriptions of Psychoearth and Neuroearth was counter-balanced across subjects, to control for ordering effects.

The results we obtained confirmed Nahmias' thesis: participants' judgments about free will and responsibility were significantly higher in the simple deterministic scenario rather than in the mechanical one ($F(1, 199) = 58.03, p < .0001$). There was a higher willingness to ascribe free will and responsibility to Psychoearthlings (mean attributed free will: 3.47, mean attributed responsibility: 3.62) than to Neuroearthlings (free will: 2.83, responsibility: 3.09) which indicates stronger compatibilist intuitions in the deterministic universe. Regarding the perceived plausibility of the scenarios, all participants considered the scenarios reasonably plausible (mean Neuroearth: 3.50, mean Psychoearth: 3.46), but we found an effect of the participants' gender: while female participants tended to consider the non-physicalist scenario as more plausible than the physicalist scenario, males considered them to be equally plausible ($F(1,197) = 5.75, MSE = 3.91, p = 0.0174$). This suggests that – at least among our participants – females seem to be less prone to consider mental events as reducible to brain events.

Aside from confirming Nahmias' result, this replication in fact also strengthened it: as mentioned, presenting both scenarios to all participants made the similarities between the descriptions entirely obvious. The fact

that the effect survives even under these conditions speaks strongly in favor of its significance.

An interesting way of looking at data on folk intuitions like those provided by this study has been pointed out by Nahmias and colleagues:³⁵ folk intuitions can be used to *settle presumptive probative obligations* – in other terms, which side of a philosophical dispute has to start providing evidence in its favor. As they put it,

if a philosophical theory does turn out to be privileged by the endorsement of the folk, that would seem to position the burden of proof on the shoulders of those who argue *contrary* to folk intuitions.³⁶

Whichever position turns out to be supported by the majority of people, would then enjoy “squatter's rights”. This was exactly the point that both incompatibilists and compatibilists tried to exploit in claiming their respective positions to be commonsensical – a dispute that only experimental philosophy can help to settle.

■ What can experimental philosophy offer to philosophy?

An observation often directed to experimental philosophy is that ordinary people lack the epistemic background to express philosophically relevant views on philosophical topics. According to this *expertise objection*,³⁷ even when it comes to common-sense concepts like those of free will or morality, philosophers can use those concepts with a level of precision and expertise that ordinary people cannot quite reach. For this reason, it seems obvious that philosophers should not be interested in folk intuitions about philosophical issues. They can safely ignore them, in much the same way as physicists can safely ignore folk intuitions about physics.³⁸

But experimental philosophers do not claim that common people's intuitions – even when clearly aligned towards a certain

position – should be considered as a definitive proof or confutation of a philosophical theory. Even if it is true that one of the different projects of experimental philosophy³⁹ consists of the assessment of *what* people's intuitions are, no experimental philosopher ever aimed to use such data for settling philosophical disputes. To put it simply, truth – at least in philosophical matters – is not a matter of consensus. Instead, the experimental philosophy project comes in response to the need to verify philosophers' claims about what is "ordinary belief", "common intuition" or "pre-philosophical theory". If philosophers want to appeal to the intuitiveness of a thesis as evidence in their theorizing, they should also want to be sure their own intuition about other people's intuitions is correct.⁴⁰

Furthermore, experimental philosophers claim that even if surveying folk intuitions does not directly aim at challenging or confirming philosophical theories, certain philosophical questions would not even exist without common sense intuitions. Consider, for example, the problem of moral responsibility, that arises only because people, for a number of reasons that it is interesting to investigate, consider themselves and others as worthy of praise or blame for their behavior. In such a case, consulting folk intuitions can indeed be helpful for formulating philosophical theories that can match with our everyday idea of what constitutes moral responsibility. In other words, in certain domains intuitions rightfully belong to the set of relevant facts that an adequate philosophical theory has to account for. In those areas, experimental philosophy stands to armchair philosophy as fact checking stands to good journalism.

Besides, there are at least two more reasons why knowing laypeople's ideas on philosophical questions can be useful for philosophy. First, folk intuitions are not influenced by an intense philosophical training and so may be more spontaneous than those of philosophers. The problem is not that lack of

spontaneity implies worse intuitions, but that philosophers themselves often present their own intuitions as being representative of laymen views on a certain philosophical issue.⁴¹ Second, it is just when folk intuitions differ from those of philosophers that it becomes interesting in its own right to investigate the relevant differences between the two ways of thinking, and to search for explanations of them.

Ultimately, the debate on the usefulness (or lack thereof) of experimental philosophy boils down to the role of intuitions in philosophical analysis. If we want intuitions to do any heavy lifting in philosophical arguments, then we better make sure we get our facts straight, and experimental philosophy has the unquestionable merit of providing a method for discerning which claims about ordinary intuitions are well-founded and which are not.⁴² Sure, the hype about it may have gotten out of hand on some occasions – but also the criticisms raised against it were often excessive, as it is wont to happen when claims about "a new approach" are made. Granted, no substantial boundary is apparent between experimental philosophy and social psychology, but nor it is needed: in fact, arguably a huge pay-off of experimental philosophy is precisely to get social psychologists more interested in philosophical topics, and to bring more philosophical analysis into psychological papers – in short, a virtuous increase in true interdisciplinary collaboration.

An interesting interpretation of the debate on experimental philosophy comes from Rose and Danks, who suggest that

the major objections to the use of intuitions by experimental philosophers are actually expressions of antinaturalism, concerns about the cognitive science, or both.⁴³

In their view, experimental philosophy, rather than being a brand new discipline, is a new instantiation of the ancient tradition of

philosophical naturalism. The only difference is that philosophers have started to conduct experiments by themselves, or in close collaboration with psychologists and social scientists, instead of just appealing to external empirical evidence. Thus, the usefulness of data coming from experimental philosophy's studies depends solely on how empirical data are relevant to that particular philosophical issue. From this point of view, even those who are not willing to grant experimental philosophy the status of "new discipline" may acknowledge its contribution to an empirically informed philosophy.

But experimental philosophy does not limit itself to surveying intuitions: another project of experimental philosophy aims to discover *why* people think the way they think, that is, what are the psychological mechanisms that produce a given intuition or folk theory. The increasing amount of data coming from experimental psychology about the many biases and prejudices involved in almost every judgment, decision or reasoning process should give pause to philosophers, since experts are certainly not immune to automatic processes and cognitive distortions. Armchair philosophers could indeed be prone to privileging theories that accord with their own intuitions: lacking proper empirical analysis of such intuitions, this tendency could be problematic, insofar as such intuitions largely depend on philosophically irrelevant factors.

Indeed, folk intuitions can be the result of a large set of factors that form or influence our subjective idea of the world. Studying in depth what factors influence people's judgments is extremely important, so as to find out what leads us to have the intuitions we have. In their experimental philosophy manifesto,⁴⁴ Knobe and Nichols claim that the first major goal of experimental philosophy is to rehabilitate a practice that has been abandoned for far too long: the pursuit of the underlying sources of our beliefs. If, for example, we were to find out that a certain intuition is driven by a particular bias or emotion-

al condition, we then would know more about the cognitive mechanisms responsible for that intuition.

Thus, experimental results on folk intuitions do not have to have a direct impact on the work of philosophers in order to be granted philosophical relevance. It is sufficient that they can have a legitimate indirect impact in terms of being a guide to discovering the underlying psychological processes that generate people's intuitions.

Acknowledgements

I would like to thank two anonymous reviewers for the helpful comments and suggestions on an earlier version of this paper. I wish to thank Fabio Paglieri for supervising my experimental research and commenting on a previous version of this paper. I am also grateful to the Institute of Cognitive Sciences and Technologies (ISTC-CNR) in Rome for supporting my research, and to Francesco Ferretti, Mario De Caro and Massimo Marraffa for their guidance throughout my studies. This research was funded by the Italian Ministry for Education, University and Research (MIUR), as part of the research project *L'amministrazione della giustizia in Italia: il caso della neurogenetica e delle neuroscienze* [*The administration of law in Italy: the case of neurogenetics and neurosciences*].

Notes

¹ D.C. DENNETT, *Elbow Room: The Varieties of Free Will Worth Wanting*, MIT Press, Cambridge (MA) 1984.

² Concerning the role and status of intuitions in "armchair philosophy" see, for instance, A.R. BOOTH, D.P. ROWBOTTOM (eds.), *Intuitions*, Oxford University Press, Oxford 2014.

³ F. JACKSON, *From Metaphysics to Ethics*, Oxford University Press, Oxford 1998.

⁴ J. KNOBE, *Intentional Action and Side Effects in Ordinary Language*, in: «Analysis», vol. LXIII, n. 279, 2003, pp. 190-194.

⁵ Whether or not the side-effect effect is determined by moral considerations, as originally sug-

gested by Knobe (2003), is far less obvious, though. Machery showed that the effect obtains even when the side effects are not regarded as morally relevant by participants (see E. MACHERY, *The Folk Concept of Intentional Action: Philosophical and Experimental Issues*, in: «Mind & Language», vol. XXIII, n. 2, 2008, pp. 165-189), and it can be explained without making recourse to moral concepts (on this, see also F. HINDRIKS, *Normativity in Action: How to Explain the Knobe Effect and its Relatives*, in: «Mind & Language», vol. XXIX, n. 1, 2014, pp. 51-72). More generally, the rich literature spawned by Knobe's seminal paper has uncovered a variety of specific results, which are not always easy to reconcile with each other: recent attempts to propose a coherent synthesis can be found in Pettit and Knobe (see D. PETTIT, J. KNOBE, *The Pervasive Impact of Moral Judgment*, in: «Mind & Language», vol. XXIV, n. 5, 2009, pp. 586-604), Hindriks (see F. HINDRIKS, *Normativity in Action*, cit), and Paprzycka (see K. PAPRZYCKA, *The Omissions Account of the Knobe Effect and the Asymmetry Challenge*, in: «Mind & Language», forthcoming).

⁶ For a discussion on the role of experimental philosophy on this topic, see S. NICHOLS, *Experimental Philosophy and the Problem of Free Will*, in: «Science», vol. CCCXXXI, n. 6023, 2011, pp. 1401-1403.

⁷ For the sake of simplicity, here I conflate two different versions of determinism: epistemological determinism, i.e. the thesis that every event can be fully explained by reference to previous facts, and ontological determinism, i.e. the claim that every event is entirely caused by past occurrences. The reason why this simplification is justified in this context is because both versions of determinism are potentially problematic for free will.

⁸ This is what is called "classical compatibilism", and includes philosophers like Thomas Hobbes, David Hume, John Stuart Mill, Donald Davidson.

⁹ H. FRANKFURT, *Alternate Possibilities and Moral Responsibility*, in: «Journal of Philosophy», vol. LXVI, n. 23, 1969, pp. 829-839.

¹⁰ R. KANE, *Free Will: New Directions for an Ancient Problem*, in: R. KANE (ed.), *Free Will*, Blackwell, Oxford 2002, pp. 222-248.

¹¹ For an exhaustive account of the past contributions and of the recent research on free will, see R. KANE (ed.), *The Oxford Handbook of Free Will*, Oxford University Press, Oxford 2011.

¹² B. LIBET, C.A. GLEASON, E.W. WRIGHT, D.K.

PEARL, *Time of Conscious Intention to Act in Relation to Onset of Cerebral Activity (Readiness-potential): The Unconscious Initiation of a Freely Voluntary Act*, in: «Brain», vol. CVI, n. 3, 1983, pp. 623-642.

¹³ D.M. WEGNER, *The Illusion of Conscious Will*, MIT Press, Cambridge (MA) 2002.

¹⁴ B. LIBET, *Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action*, in: «Behavioral and Brain Sciences», vol. VIII, n. 4, 1985, p. 536.

¹⁵ D.M. WEGNER, *The Illusion of Conscious Will*, cit., p. 3.

¹⁶ See E. NAHMIA, *When Consciousness Matters: A Critical Review of Daniel Wegner's The Illusion of Conscious Will*, in: «Philosophical Psychology», vol. XV, n. 4, 2002, pp. 527-541; M. SLORS, *Conscious Intending as Self-programming*, in: «Philosophical Psychology», vol. XXVIII, n. 1, 2015, pp. 94-113.

¹⁷ S. FELLETTI, F. PAGLIERI, *The Illusionist and the Folks. On the Role of Conscious Planning in Intentionality Judgments*, submitted to «Philosophical Psychology», under review.

¹⁸ S. NICHOLS, J. KNOBE, *Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions*, in: «Nous», vol. XLI, n. 4, 2007, pp. 663-685.

¹⁹ S. NICHOLS, *The Folk Psychology of Free Will: Fits and Starts*, in: «Mind & Language», vol. XIX, n. 5, 2004, pp. 473-502.

²⁰ R. KANE, *Responsibility, Luck, and Chance: Reflections on Free Will and Indeterminism*, in: «The Journal of Philosophy», vol. XCVI, n. 5, 1999, pp. 217-240, here 217.

²¹ G. STRAWSON, *Freedom and Belief*, Oxford University Press, Oxford 1986.

²² H. FRANKFURT, *Alternate Possibilities and Moral Responsibility*, cit.

²³ Judgments of moral responsibility are also surveyed in such cases, being the concept of free will intimately related to that of moral responsibility, e.g. because it is a necessary condition for it.

²⁴ E. NAHMIA, S. MORRIS, T. NADELHOFFER, J. TURNER, *Surveying Freedom: Folk Intuitions about Free Will and Moral Responsibility*, in: «Philosophical Psychology», vol. XVIII, n. 5, 2005, pp. 561-584.

²⁵ S. NICHOLS, J. KNOBE, *Moral Responsibility and Determinism*, cit.

²⁶ E. NAHMIA, J. COATES, T. KVARAN, *Free will, Moral Responsibility, and Mechanism: Experiments on Folk Intuitions*, in: «Midwest Studies in Philo-

sophy», vol. XXXI, 2007, pp. 214-242.

²⁷ E. NAHMIA, *Folk Fears about Freedom and Responsibility: Determinism vs. Reductionism*, in: «Journal of Cognition and Culture», vol. VI, n. 1-2, 2006, pp. 215-237, here p. 217

²⁸ *Ibidem*.

²⁹ S. FELLETTI, F. PAGLIERI, *The Illusionist and the Folks*, cit.

³⁰ Italian version of the Barratt Impulsiveness Scale, see J.H. PATTON, M.S. STANFORD, E.S. BARRATT, *Factor Structure of the Barratt Impulsiveness Scale*, in: «Journal of Clinical Psychology», vol. LI, n. 6, pp. 768-774.

³¹ In response to feedbacks we received after the pilot study, we decided to remove references to upbringing, among the causes of action mentioned in the psychological deterministic scenario (Psychoearth). Upbringing, in fact, appeared to be confusing for several participants, that believed it was the key concept for determining the difference between the scenarios.

³² Questions about free will and determinism were: (1) Do you think Psychoearthlings/Neuroearthlings act by their own free will?; (2) Do you think we could ascribe praise or blame to Psychoearthlings/Neuroearthlings deeds?

³³ E. NAHMIA, J. COATES, T. KVARAN, *Free Will, Moral Responsibility, and Mechanism*, cit., also use a 6-point rating scale for a similar study.

³⁴ We carried our tests on bachelor students who have not yet attended a philosophy course. Moreover, when we returned to explain the aim of our study and illustrate our partial results, we asked them whether they were familiar with philosophical concepts such as free will or determinism, and they all replied in the negative. Anyway, we considered their very age and role of students as a sufficient condition to consider them as “non-professional”.

³⁵ E. NAHMIA, S. MORRIS, T. NADELHOFFER, J. TURNER, *Surveying Freedom*, cit.

³⁶ *Ivi*, p. 564.

³⁷ T. WILLIAMSON, *Philosophical Expertise and the Burden of Proof*, in: «Metaphilosophy», vol. XLII, n. 3, pp. 215-229.

³⁸ While this expertise defence against the philosophical significance of experimental philosophy results is still debated in the literature, empirical results suggest that, as far as intuitions on free will and moral responsibility are concerned, experts (i.e. philosophers) are no less biased than laymen – therefore undermining the expertise defence

(see E. SCHULZ, E. T. COKELY, A. FELTZ, *Persistent Bias in Expert Judgments about Free Will and Moral Responsibility: A Test of the Expertise Defense*, in: «Consciousness and Cognition», vol. XX, n. 4, 2011, pp. 1722-1731).

³⁹ See E. NAHMIA, J. COATES, T. KVARAN, *Free Will, Moral Responsibility, and Mechanism: Experiments on Folk Intuitions*, cit.

⁴⁰ Sommers, in his recent “sympathetic critique” (sic) of experimental philosophy studies on free will and moral responsibility (see T. SOMMERS, *Experimental Philosophy and Free Will*, in: «Philosophy Compass», vol. V, n. 2, 2010, pp. 199-212), argues that, while fact checking appeals to intuition is a noble call, its significance ultimately depends on the role such appeals play in the philosophical theory in question. With respect to the debate on compatibilism vs. incompatibilism, Sommers notes that incompatibilists use claims on the intuitiveness of incompatibilism (the polemical target of the experimental philosophers, in this case) simply as a rhetorical device to introduce the notion, without assigning it any serious role in their arguments in favour of incompatibilism. So, Sommers suggests, experimental philosophy may very well prove these claims to be wrong, and yet fail to have an impact (other than providing interesting information on empirical facts) on the philosophical arguments of the incompatibilists.

⁴¹ The following are relevant examples with respect to incompatibilism: «In my experience, most ordinary persons start out as natural incompatibilist» (R. KANE, *Responsibility, Luck and Chance*, cit., p. 217); «We come to the table, nearly all of us, as pretheoretic incompatibilists» (L. EKSTROM, *Libertarianism and Frankfurt-style Cases*, in: R. KANE (ed.), *The Oxford Handbook of Free Will*, Oxford University Press, Oxford – New York 2002, pp. 309-322, here p.310); «Most of us start off by making an important assumption about freedom» (see T. PINK, *Free Will: A Very Short Introduction*, Oxford University Press, Oxford – New York 2004, p. 12).

⁴² On the question of whether philosophers really have to rely on intuitions as evidence, see H. CAPPELEN, *Philosophy without Intuitions*, Oxford University Press, Oxford 2012; T. WILLIAMSON, *Philosophical Expertise and the Burden of Proof*, cit.

⁴³ D. ROSE, D. DANKS, *In Defense of a Broad Conception of Experimental Philosophy*, in: «Metaphilosophy», vol. XLIV, n. 4, 2013, pp. 512-532, here

p. 520. I thank an anonymous reviewer for this suggestion.

⁴⁴J. KNOBE, S. NICHOLS, *An Experimental Philoso-*

phy Manifesto, in: J. KNOBE, S. NICHOLS (ed.) *Experimental Philosophy*, Oxford University Press, Oxford - New York 2008, pp. 3-14.