In search of the neurobiological basis of decision making: explanation, reduction and emergence

The world, it seems, runs in parallel, at many levels of description. You may find that perplexing; you certainly aren't obliged to like it. But I do think we had all better learn to live with it (1).

The kind of naturalism I have in mind aims at bridging gaps between the sciences, not at universal reduction. Some important generalisations are likely to be missed when causal relationships are not accounted for in terms of lower-level mechanisms. Other valuable generalisations would be lost if we paid attention to lower-level mechanisms only. If we want bridges, it is so as to be able to move both ways (2).

Michele Di Francesco[a,b]
Matteo Motterlini[a,b]
Matteo Colombo[b]

[a] Faculty of Philosophy, Università "Vita-Salute" San Raffaele, Milan, Italy
[b] CRESA (Research Centre for Experimental and Applied Epistemology)
Università "Vita-Salute" San Raffaele, Milan, Italy

Corresponding author: Prof. Michele Di Francesco
Facoltà di Filosofia, Università Vita-Salute S. Raffaele
Palazzo Arese Borromeo, 20031 - Cesano Maderno
Milan - Italy
E-mail: difrancesco.michele@hsr.it

Accepted for publication:

Summary

In recent years, the study of decision making has provided a paradigmatic case of "crossbreeding" of different disciplines. The integration of economics, psychology and neurosciences within neuroeconomics calls for more accurate and comprehensive models of human rationality, which may be obtained by combining diverse theoretical approaches and experimental techniques. In this respect, neuroeconomics contributes to a naturalistic, brain-based, explanation of human agency. However, although contemporary naturalism insists on the unitary aspect of reality, we stress that supporting unitary study of nature is not the same as supporting a single, fundamental discipline to which all higher-order analyses could (or should) be reduced. We argue for integration, rather than reduction as the best approach to a naturalistic explanation of human decision making, and we claim that supporting epistemological pluralism does not mean being committed to any specific ontological position. However, we suggest that an "emergentist" ontology is the best candidate to integrate the epistemological analysis here endorsed.

KEY WORDS: decision making, emergentism, epistemological pluralism, neuroeconomics.

The philosophy of neurophilosophy: integration or reduction?

From the perspective of neurophilosophy, the search for the neural correlates of human decision making is a peculiar case. We refer, in fact, to the attempt to integrate evidence, concepts and tools from the fields of economics, psychology and neuroscience within the new domains of neuroeconomics (and neuroethics) (3-14) as a manifest realisation of the methodological ideal described more than twenty years ago by Patricia Churchland:

> *"Neuroscience and psychology need each other. Crudely, neuroscience needs psychology because it needs to know what the system does; that is, it needs high-level specifications of the input-output properties of the system. Psychology needs neuroscience for the same reason: it needs to know what the system does. That is it needs to know whether low-level specifications bear on the initial input-output theory, where and how to revise the input-output theory, and how to characterize processes at levels below the top." (15)*

A neuroeconomics research programme shows how a brain-based investigation of human behaviour may help higher-order theories to find new sources of validation, to modify significant assumptions, and even to reconfigure previously unquestioned categories (15). It is also a valuable tool for achieving a greater level of integration between the behavioural and cognitive sciences, on the one hand, and the neurosciences on the other, thereby contributing to a naturalistic explanation of human agency. All this is a major achievement for the neurophilosophical view which, on its own, is one of the best examples of a naturalistic approach to the study of human nature, since it suggests the possibility of greater integration between higher-level explanations of human behaviour, based on psychological concepts, and lower-level (brain-based) explanations of our cognitive performances. Contemporary naturalism insists on the unitary aspects of reality: "the world should be a unity in the sense of being amenable to a unified study which can be called the study of nature" (16) and the significant advances made in recent years by neuroeconomics and neuroethics bear witness to the progressive character of a research programme that aims to arrive at a unitary explanation of human agency. But to maintain that the world is a unitary place is not to claim that it is a monolith – or a pyramid, or whatever else may suggest that there exists just a small set of basic principles whose implementation can explain all the different features in which reality, in all its complexity, is structured. Put another way, supporting unitary *study* of nature does not amount to supporting a single *discipline* to which all higher-order analyses could (or should) be reduced. Perhaps, *integration*, *rather than reduction*, is the best approach for arriving at a naturalistic explanation of the human world (17-20).

The main aim of this paper is to suggest that real-science observation shows that *epistemological pluralism* is currently the best available approach if we want to grasp the intertheoretical relations in the growing field of neuroeconomics. Even though our main concern is epistemological, we shall also discuss its relevant metaphysical consequences with regard to the ontological relationship between brain processes and higher-order mental functions. We begin by illustrating the search for the neural basis of decision making, citing two case studies, the first regarding decisions made in conditions of *uncertainty*, the second regarding ethical dilemmas. We then deal with epistemological pluralism, and finally with the issue of the metaphysical import of brain-based explanations.

Two case studies: the evidence and its epistemological implications

Economists have traditionally embraced comprehensive models in order to understand choice behaviour. According to these models, people choose between alternative courses of action on the basis of a rational evaluation of the outcomes of their decision. In spite of their mathematical tractability and formal rigour, these models regard the mind as a "black box" and focus on the outcomes rather than on the mechanisms by which decisions are reached in determinate contexts.

Psychological studies on judgement and decision making, however, have provided a wealth of evidence and cast serious doubt on the descriptive validity of the neoclassical "rational choice theory". Thus, the tentative integration of constructs from cognitive psychology led to the development of behavioural economics which fits better with "real world" phenomena, even though it is not as mathematically "clean" (21).

As part of this flow, we find the neurosciences searching for those brain processes that may underlie empirically observed economic and social behaviours. This is a first path that economists and neuroscientists can profitably follow together. Neuroscience can inform economic models, providing evidence and measures that support the variables and psychological parameters introduced in the field of behavioural economics. According to this view neuroeconomics is a subfield of both behavioural and experimental economics. In fact, neuroeconomics tests the empirical tenability of notions such as "fairness", "trust", "framing effect", "gratification delay", and "ambiguity aversion", using experimental tools and methods borrowed from the neurosciences [e.g. technologies for measuring brain activity such as functional magnetic resonance imaging (fMRI) and positron emission tomography (PET); studies of patients with brain lesions; recording activity of a single neuron at a time; active transcranial magnetic stimulation (TMS); psychophysiological recording; computational models].

One of the main achievements of this line of research is its suggestion that human behaviour is not the product of a unitary system, but is, rather, driven by the interaction of multiple specialised systems (22), and the usefulness of neuroeconomics in this sense is that it enables us to understand the heterogeneity of observed behaviour when standard economic models predict unique behaviour.

Benefits of another type, provided by neuroeconomics, concern the identification of neurobiological variables influencing choice and behaviour, which classical and behavioural theories neglect. The aim, in this case, is to make fresh predictions about observed behaviour on the basis of an understanding of brain mechanisms. Possibly, in this second sense, neuroeconomics represents a real breakthrough in decision-making research inasmuch as information about activation of specific brain areas makes it possible to take a step further towards understanding the nature of underlying decision-making processes. Thus, the biggest advantage of neuroscience could be not just that it lends precision to parameters in standard economic models, but also, and above all, that it locates a "middle ground" where it is possible to build models capable of "mediating" between the abstract criteria of rational theory and the concrete mechanisms of human behaviour[1].

To illustrate these methodological points we shall consider two recent case studies that addressed, respectively, the following questions: "Is there a neurobiological system which responds to different levels of uncertainty?"; "What are the neurobiological bases of moral judgement?" Our main concern in the next sections is to supply a straightforward illustration of the dynamics of such cross-scientific investigations.

Investigating uncertainty through the convergence of a host of different tools

According to most economic analyses, the variables that should be taken into account when making a decision in an uncertain-choice situation are the judged probabilities of possible outcomes and the evaluation of those outcomes. However, whereas in some decision-making situations, such as a game of poker, there are multiple possible outcomes that could occur with known or es-

[1] Notably, there is a third way in which economics and neuroscience fruitfully interact, suggesting new neurobiological models. Elegant and well-formalised economic models are valuable benchmarks for neuroscientists seeking to understand how neural structures encode information in order to facilitate decision making. Game theory and Bayesian models could be the basic instruments for studying the relationship between behaviour and the brain. The champion of this approach to neuroeconomics is Paul W. Glimcher. Glimcher argues "that economically based optimality modelling can be used to identify the computational goal of a behaviour and that the behaviour of real animals can be understood in terms of these goals." (23) Despite being of great interest, we are not taking this research programme into account here.

timable probabilities, in others, such as when considering the chance of a snowfall in Milan on a particular day, relevant information for judging the probabilities of possible outcomes is lacking. The first situation is a case of risk (i.e. uncertainty with known probabilities), the second a case of ambiguity (i.e. uncertainty with unknown probabilities). Although some ambiguity is present in most decision-making situations, standard economic theory does not distinguish sharply between decision making in conditions of risk and ambiguity, and tends to regard ambiguity as a special, more complex case of risk. According to this view, agents should not behave differently in the face of risk and ambiguity.

Hsu et al. (24) investigated ambiguity and risk using fMRI. Their work builds on a wealth of evidence provided by the field of behavioural economics which shows that people are likely to prefer risky options to ambiguous ones, even when this contradicts expected utility theory predictions. In this regard, the Ellsberg (25) paradox is paradigmatic. It shows that, typically, people are ambiguity averse. The expected utility theory states that all information relevant to decision making with respect to the likelihood of an event is just the judged probability that the event will occur. On the contrary, behavioural research shows that choices made in conditions of uncertainty strongly depend on how much relevant information is missing in the decision-making context. This contrast justifies the neuroeconomic search for neurobiologically based distinctions between risk and ambiguity.

Thus, Hsu et al. (24) set out evaluate general uncertainty through investigation of neural systems; they used data both from an fMRI study and from neurological patients. The fMRI study involved three experimental contexts: the Card-Deck treatment, the Knowledge treatment and the Informed Opponent treatment. In all three cases, the focus was the response of different neural mechanisms to a degree of uncertainty in human decision making. The crucial factor manipulated in the experiment was the amount of information available. All three experimental contexts depict one condition where the subject lacked relevant information (ambiguity condition), which was instead available in the risk condition. Subjects made decisions between pairs of bets. In the ambiguity condition they chose between betting on one of two ambiguous options or accepting a certain payoff; in the risk condition, instead, subjects chose between two options with known or estimable likelihood, or taking a certain payoff.

Their results showed increased activation in the dorsomedial prefrontal (DMPFC) area, orbitofrontal cortex (OFC) and amygdala during the ambiguous condition as opposed to the risk condition. The level of ambiguity correlated negatively, instead, with the striatal system which showed greater activation in response to risk. Ambiguity aversion was correlated with higher right OFC activity.

The OFC is a portion of the prefrontal cortex located at the base of the frontal lobe. It is not easy to define its exact functional role because of difficulty characterising the behaviours for which it is responsible. Nonetheless, the OFC has been demonstrated to play a crucial role in the ability to process and react to affective (i.e. carrying positive/negative valences) and social information (26). The amygdala is an almond-shaped group of neurons located in the medial temporal lobes, close to the anterior portion of the hippocampus. Given its interconnections with the hippocampal system, the amygdala is mostly implicated in emotional learning and memory, especially in fear conditioning and in response to threating stimuli (27). The DMPFC is implicated in modulation of amygdala activity. In view of the fact that the OFC is bidirectionally connected with the amygdala, neuroscientists are beginning to investigate how these areas cooperate in order to bring out specific behavioural responses.

Dorsal striatum (caudate nucleus) activation was correlated with detection of reward-related stimuli. The caudate nucleus is a telencephalic structure located within the basal ganglia. Connected with cortical regions, it is highly innervated by dopamine neurons and is thought to play an important role not only in higher-order motor control, but also in learning and memory, particularly in reward anticipation and evaluation (28).

Hsu and colleagues' work revealed different activation time courses in the ambiguity versus the risk circuitry. The OFC and amygdala, indeed, reacted rapidly at the beginning of the trial; instead the dorsal striatum, building more slowly, peaked later. This difference may be seen as a strong indication of the existence of two neural systems. Whereas the ambiguity circuitry is part of an automatic, affective system that quickly proposes intuitive answers to problems as they arise, the risk circuitry is part of a controlled, slow system which evaluates inputs and allows deliberation.

Hsu et al. also tested card-deck problems on patients with OFC damage to establish whether the OFC plays a necessary role in the evaluation of uncertainty. It was actually found that the OFC-lesioned patients were insensitive to the level of ambiguity and risk involved in a choice, maybe because uncertain situations did not cause them emotional distress. It is worth noting that, ironically, risk and ambiguity neutrality is consistent with the prediction of rational choice theory.

Building on the choice behaviour displayed by the fMRI-investigated subjects, Hsu et al. constructed a nonlinear stochastic model of the subjects' choices from which they derived measures for ambiguity and risk aversion. It was then discovered that there actually exists a close link between the behavioural parameter for ambiguity aversion and right OFC activity. This suggests the presence of a biological variable which exerts a direct influence on behaviour, largely ignored in standard theory.

These data point to a general neural circuit which responds to the degree of uncertainty associated with a decision-making problem. The OFC and amygdala thus appear to be implicated in a lower-level, neurobiological mechanism responding to degrees of uncertainty. Even if partial and flawed, we know that this mechanism responds to uncertainty, alerts to unknown, potentially dangerous, consequences of a course of action, and warns cognitive resources to see whether more information can be derived from the environment.

This inquiry supports the behavioural economics distinction between risk and ambiguity, and introduces the concept of varying degrees of uncertainty, bypassed in standard theory. Moreover, because choice in conditions of uncertainty is a fundamental and widespread activity that takes place at many levels of human behaviour, the identification of neurobiological mechanisms of ambiguity aversion and their connection with psychological and economics research, may have potential implications for understanding and explaining phenomena such as preferences for familiarity and home-bias in assets (that is the tendency for investors to invest in a disproportionate amount of domestic equities), biased evaluations about (military, social and environmental) risk, people's aptitude for saving for retirement, companies' pricing of insurance, to touch on just a few concrete examples.

In search of the neural correlate of "doing the 'right' thing"

Neuroeconomics experiments that involve evaluation of different scenarios and allow subjects to make choices with regard to hypothetical actions and outcomes provide a new perspective from which to consider the basis of moral judgement. The primary goal of so-called neuroethics is not to construct a new positive moral philosophy, but rather to identify neural structures of moral psychology in order better to understand how our moral attitudes work.

Typical neuroethics experiments involve fMRI scanning of subjects responding to dilemmas designed to elicit a definite moral attitude. Consider the well known trolley problem devised by Philippa Foot (29). A runaway trolley rushes towards five people who will be killed if it proceeds on its present course. The only way to save them is to hit a switch that will divert the trolley onto another set of tracks where it will kill only one person instead of five. Most people consider that in these circumstances it would be morally acceptable to hit the switch, thereby saving five people at the expense of one. Instead, in an alternative scenario in which pushing a person off a footbridge to stop the runaway trolley will prevent it from hitting the five people, most people prefer not to act even though, as in the previous scenario, there would still be five lives saved and only one lost.

But why, in this second case, would we tend not to act? What is the difference that motivates these different moral judgements? Beyond a utilitarian calculation (in these cases, saving more lives), the second scenario seems more emotionally charged than the first, because it involves a personal moral violation (i.e. actually being the one to push another person off a bridge) that results directly from the agent's will. Consequently, it seems that a difference in emotional response would explain the different patterns of judgment in these two scenarios. When we are directly involved in a situation our reaction is more instinctive, automatic, visceral; instead, "impersonal" moral judgements, in situations in which we are able to distance ourselves from direct involvement, are driven less by emotional response and more by "cold", cognitive processes.

Greene et al. (30), testing this two-systems hypothesis through an fMRI experiment, found that direct involvement in dilemmas implicating personal moral violation increasingly activated brain regions associated with emotion and social cognition (medial prefrontal cortex, posterior cingulate/precuneus and superior temporal sulcus/temporo-parietal junction). Instead, the brain regions usually associated with abstract reasoning and cognitive control (DLPFC, anterior cingulate cortex) were activated to resolve dilemmas in which utilitarian values led people to accept such personal moral violations. These findings support a theory of moral psychology that acknowledges crucial and sometimes competitive roles for both cognitive and emotional processes.

Nonetheless, the study by Greene et al. did not determine the specific contribution of emotion to moral judgement. Critically, it did not address the following problem: are emotional activations a cause or a consequence of moral judgements? In fact, fMRI data, taken alone, are not enough to answer this question. Neuroeconomic explanations, on the other hand, draw support from multiple lines of empirical work. One such line is that of clinical studies of patients with brain lesions. Although neuropsychological studies, too, may be unable to establish the precise performance of a certain structure, they can tell us whether a brain region performs a necessary role – be it direct or mediated by other processes – in the generation of normal behaviour.

Recently, Koenigs et al. (31) investigated subjects with focal damage to the ventromedial prefrontal cortex (VMPC) in order to assess whether emotional processes are causally implicated in moral judgement and whether they are necessary in the generation of typical moral evaluations. Patients with VMPC lesions showed conserved general intelligence abilities and capacity for logical reasoning and declarative knowledge of social and moral norms. Instead, their emotional responsivity and their social emotion processing were significantly jeopardised: their compassion, shame and guilt reactions were impaired and their anger and frustration poorly regulated. In view of this profile, what might VMPC subjects' attitudes towards moral judgement be? "If emotional responses mediated by VMPC are indeed a critical influence on moral judgement, individuals with VMPC lesions should exhibit an abnormally high rate of utilitarian judgements on the emotionally salient, or 'personal', moral scenarios (for example, pushing one person off a bridge to stop a runaway boxcar from hitting five people), but a normal pattern of judgements on the less emotional, or 'impersonal', moral scenarios (for example, turning a runaway boxcar away from five people but towards one person). If, alternatively, emotion does not play a causal role in the generation of moral judgements but instead follows from the judgements then individuals with emotion defects due to VMPC lesions should show a normal pattern of judgements on all scenarios" (31). The interesting result of this study was, precisely, the fact that compared with the controls' patterns of judgement on personal dilemmas, where social emotions are crucial in resolving possible moral conflicts, the VMPC subjects produced more utilitarian judgements. Impairment of the VMPC thus seems to interfere with the affective processing associated with typical emotional reactions to violation of others. What conclusions are worth drawing from this study?

First, these findings provide evidence that emotion mediated by the VMPC is necessarily involved in the generation of judgements in a definite class of moral dilemmas. In brief, the VMPC seems to be *causally involved* in the formation of a certain type of evaluation. Moreover, what is of greatest interest for our purposes is the methodological strategy that allowed this conclusion to be reached. In a nutshell: a correlation between neurobiological structure S and behavioural function F was observed in an fMRI experiment; subsequently, research on subjects with impairment of target structure S (lesioned or temporarily disrupted by repetitive TMS) focused on the causal involvement of S in F. Both of those studies showed that if S is damaged then F is considerably impaired. Therefore, S is necessary for F. Of course, all these inferences need to be strengthened by further empirical evidence from multiple, convergent lines of inquiry; nonetheless a nice starting point, beyond mere correlation, has been defined: we have formulated a clear hypothesis about *brain-based causal explanations* of moral behaviour.

Epistemological pluralism: a recipe for neuroeconomics

Our case studies show that, to reach its goal, neuroeconomics must succeed in integrating evidence, concepts and tools from the fields of economics, psychology and neuroscience. If we wanted to sketch a broad outline, the ingredients for this kind of integration would be:

- economics and cognitive psychology – which identify functional units (and classify behavioural phenomena);
- neurosciences – brought in to define structural units (i.e. neural phenomena);
- specialised experimental and analytical methods – to discover and justify structure/function relations;
- an interpretative framework – to connect, in a unified scenario, the different ingredients[2].

The result is a wide epistemological picture of what might be termed "explanatory pluralism". Explanatory pluralism is the idea that "simultaneously pursuing research at multiple analytical levels in science tends to aid progress at each of those levels" (32-34). We regard the above case studies as justification for taking explanatory pluralism as the perspective from which to evaluate properly the epistemology underlying neuroeconomics. A neuroeconomics research programme, indeed, benefits from separate inquiries that occur at multiple levels, combining concepts, methods and evidential resources from both social and biological sciences.

With regard to intertheoretical relations, by embracing explanatory pluralism we switch from a type of reductionism which aims at ontological economy and unification of science to a pluralistic emphasis on the increase of explanatory sources. In our view, therefore, explanatory pluralism constitutes the most suitable framework for studying the dynamic relations between economics and neurosciences, for it makes it possible to grasp the actual way these two fields inform each other.

Although the neurophilosophical perspective is commonly associated with reductionism of some sort, from the very beginning it placed the greatest emphasis on the dynamic aspect of intertheoretical relations, speaking of *co-evolution* of theories. "Co-evolution typically is […] interactive […] and involves one theory's being susceptible to correction and reconceptualization at the behest of the cohort theory. […] The heart of the matter is that if there is theoretical give and take, then the two sciences will knit themselves into one another" (3).

The relations between two theories can evolve over time because they can go through mutual adjustments and theoretical changes. Simply put, the idea is that two theories at different but adjacent levels (e.g. psychology and neurosciences) may inform and correct each other as they grow and mature further to discoveries and conceptual refinements. In the same way, explanatory pluralism does not overlook the role that interfield connections play in the progress of scientific practice, highlighting relations between structures and functions, causes and effects, parts and wholes, that may encompass different disciplines. It is remarkable, finally, that explanatory pluralism does not dispense with higher-level fields, rather it tends to preserve integrated and yet autonomous perspectives (20,32,33).

As a matter of fact, contemporary behavioural decision-making fields, such as economics and psychology, are not unified since they operate with diverse explanatory models in order to deal with distinct concerns. In fact, within each field we have well-developed methodologies which serve specific purposes. Nonetheless, a certain interpretation of the neuroeconomics approach could be misleading. Reflecting on neuroeconomics without methodological concern, one may wonder why the identification of the biological basis of economic behaviour, combined with powerful analytic tools like game theory and Bayesian models, will not supply a grand unification across the social sciences. Glimcher and Rustichini (35) seem to assign this very task to neuroeconomics. They hint that this emergent discipline should be properly understood as the attempt to "fuse" natural and social scientific fields "with the ultimate aim of providing a single (unified), general theory of human behavior" (35). But are they justified in maintaining this view of neuroeconomics? More radically, should there be a unique, systematic, universally applicable theory of human decision making?

The answer to these questions is what crucially defines the meaning and scope of the neuroeconomics research project. But instead of espousing the unificationist view, we claim that actual scientific practice suggests a very different interpretation; the mission of neuroeconomics is to integrate rather than unify knowledge from different fields (e.g. economics, cognitive and social psychology, neurosciences) and to apply that integrated knowledge to real world problems which currently do not fall in the domain of any one discipline. Consider, for instance, the way Hsu et al. (24) carried out their experiment. The behavioural findings associated with ambiguity aversion justified the distinction between decision making in conditions of ambiguity as opposed to risk. This distinction, in turn, rests upon the construct of psychological aversion to ambiguity that turned out to be positively correlated with amygdala and OFC activity. A neuroeconomics experiment like this should not be taken as an ultimate effort to reach a unified theory of human behaviour, instead it should be judiciously seen as an interfield endeavour, with the goal of making different levels of knowledge, of different kinds and from different fields, interact so as to be able to tackle the many questions that do not clearly belong to any exclusive theoretical domain.

---

[2] This kind of methodological pattern is not exclusive to neuroeconomics, but is shared by other research programs in cognitive neurosciences, for example in neurolinguistics. However, the major difference between distinct research programmes concerns the definition of functional units. In the case of neurolinguistics, for instance, functional units would be defined by linguistics and psycholinguistics.

Those who argue for a unificationist view of neuroeconomics take for granted that decision-making models and the neurobiological functioning of the brain can actually be compatible. But, *a priori* we cannot rule out the possibility that two different fields have developed models which are individually coherent but incommensurable or mutually inconsistent. Therefore, it does not appear to be a wise strategy to declare that one has great expectations of a new-born scientific programme, when there is no epistemological guarantee that the different streams of data can actually converge.

In fact, unificationists, apart from alluding to payoffs that would follow from unity of different disciplines, do not really support their claim. Presumably they assume that unification enhances explanatory strength, providing broad, comprehensive laws.

Recently, however, a large body of research from different fields firmly challenges the epistemological ideal that scientific theories should be judged in relation to their ability to express universal laws. Philosophers (Hacking, Cartwright, Dupré, Suppes), historians (Galison), and sociologists (36-43) of science claim that research practice in the social and natural sciences suggests that there is not – and probably cannot be – any such a thing as a universal law for all the phenomena of a determinate class. The laws most likely to be true are numerous and diverse, complicated and limited in scope. As Cartwright persuasively put it: "the laws that describe this world are a patchwork, not a pyramid" (38).

Brain-based explanations and emergent properties

Our epistemological conclusions demand an ontological complement; does a patchwork model of scientific explanation imply a patchwork picture of the world? A picture showing that the world we inhabit is not a unified place of universal order, but rather a basically "dappled" one? Could it be, therefore, that there exists no totally accurate descriptive account of any bit of reality?[3] (38). In our view, to be an epistemological pluralist does not mean having to adopt such a radical view. The kind of pluralism we are arguing for here leaves room to co-evolution, intertheoretical relations, and many forms of integration among different levels of inquiry.

We are not claiming then that the methodological and epistemological reading of the neuroeconomic findings we proposed in the previous paragraph demands a pluralistic metaphysical view. We are not inferring ontological conclusions from epistemological premises[4]; we are only claiming that our analysis is compatible with them. This compatibility is an important fact, however. Crudely put: neuroeconomics does not have its own, single metaphysical view; rather there are (at least[5]) two views, which we shall refer to as *reductionism* and *emergentism*. The reductionist approach claims that behind present epistemological variety there is ontological unity. Reductionism looks for a unique, systematic, universally applicable theory of human decision making and takes the neurobiological level to be ontologically fundamental. From this perspective, when we are able to propose, as in the previously presented findings of Koenigs et al. (31), *brain-based* causal explanations of moral behaviour, we are in fact describing the *real causes* of behaviour – offering more fundamental causal explanations than higher-level explanations (such as psychological and economic ones). In contrast with this view, emergentism claims that the world is structured on many different levels, each of which exhibits different kinds of causally efficacious properties. These properties are emergent in the sense of being "novel" and "irreducible" to the "basic", lower-level properties. In order to clarify the idea that there are many ontological perspectives compatible with neuroeconomic research, we may compare the following alternative metaphysical views:

- *Reductive physicalism:* physics is the key to true metaphysics. Physical explanations are the real ones. Any other higher-order explanation, in principle, can be deduced from physical explanations *plus* the definitions of higher-order concepts in terms of lower-order ones (45-47). Being nearer to physics, brain-based explanations are more fundamental than psychological or socio-cultural ones.
- *Emergentism:* reality is structured on different levels, and when a given level reaches a certain degree of complexity new causal powers emerge. So, higher-order psychological phenomena emerge when a suitable set of brain processes occur, but they are not reducible to the lower-level processes; rather they possess their own causal powers whose instantiation falls under higher-level laws and generalisation[6] (44,47-51).

As long as we are prepared to settle for the methods and results of neuroscience research, both these metaphysical views are acceptable. The reductive perspective may be associated with a view of intertheoretical relations that can be labelled "brain fundamentalism". Brain fundamentalism sees brain-based explanations, as opposed to higher-level (e.g. psychological or economic) explanations, as playing the fundamental role in our explanatory practice. The increasing success of scientific research into brain-based causal explanations of human behaviour (of the kind described earlier) may encourage brain fundamentalism. According to this view, authentic biological explanations about our nature should be contrasted with psychological, social and cul-

[3] Cartwright argues powerfully for this picture of the world. She states that the empirical success of our best scientific models results from these being idealised into theories distantly abstracted from the messiness of reality. According to Cartwright, the world is messier than our theories depict it; "as appearances suggest, we live in a dappled world, a world rich in different things, with different natures, behaving in different ways". Thus it may be our "dappled" world itself that rejects the ideal of unity of science, which would best account for what we observe in nature.

[4] Not yet, at least; see later on.

[5] We choose two radical alternatives such as reductionism and emergentism just to strengthen our point (that they are both compatible with our epistemological reading of neuroeconomics). Of course, there are other views – starting from non-reductive physicalism (44).

[6] We are not claiming that emergentism is the best way of expressing the "dappled world" model described before (speaking of different domains of reality is not the same as speaking of different levels of reality). Again, we chose it because a) it is a naturalistic perspective, and b) the alternative between reductive physicalism and emergentism is one of the most illuminating in the current debate on the ontology of mind and cognition.

tural narratives. The former should be considered closer to the true ontological structure of reality – and in this sense the causal transactions described at biological level are closer to the true causal ordering of the world. Since the biological causal explanation is more fundamental than psychological or economic ones, to appeal to alternative explanations is, if not illegitimate, at least less interesting or less illuminating. To quote Philip Dick's novel, non-biological explanation is destined to be only our "penultimate truth" (52).

We believe, however, that the adoption of epistemological pluralism to reconstruct intertheoretical relations prevents any direct derivation of brain fundamentalism. While this latter view may depend on further philosophical assumptions, it is certainly not based on scientific practice as it is, since this practice is compatible with a different metaphysical scenario – such as emergentism. Indeed, in spite of a natural tendency to consider the brain level as more fundamental, epistemological pluralism resembles a causal pluralism very close to emergentism[7] (53,54).

As we said before, this brand of causal pluralism holds that nature is structured in many levels of reality, each of which exhibits different kinds of properties that may be causally efficacious. According to this view, we have reasons to believe that there are non-microphysical regularities within the physical and social world that are disclosed by the special sciences. These regularities allow us to speak of emergent properties which are causally responsible and which describe a proper level of organisation of the world.

From this perspective there are many ways of understanding and explaining people's behaviour. The physical and neurobiological approach is one, the psychological another, and so on. This latter approach does not accept that biology offers special access to the causal explanation of human behaviour, but it is perfectly compatible with the integrationist strategy we described before. Where decision making is concerned, brain-based mechanisms are parts of a wider picture, extending form neuro- to psychological and socio-cultural reality, but they are not metaphysically fundamental.

Finally, this kind of pluralism is not in conflict with an integrated and naturalistic conception of human actions. We are not compelled to limit ourselves by observing that the world "runs in parallel at many levels of description" (1); we may go further and look for connections between (really existing) levels. Indeed, we are fully vindicated in adopting a naturalistic stance that "aims at bridging gaps between the sciences, not at universal reduction" (2).

References

1. Fodor J. Special sciences: still autonomous after all these years. In: Fodor J ed In Critical Condition. Cambridge, MA. MIT Press 1998:9-24
2. Sperber D. Explaining Culture, Oxford, Blackwell 1996, p. 99.
3. Camerer CF, Loewenstein G, Prelec D. Neuroeconomics: how neuroscience can inform economics. Journal of Economic Literature 2005;XLIII: 2004;9-64
4. Glimcher PW, Rustichini A. Neuroeconomics: the consilience of brain and decision. Science 2004;306:447-452
5. Kenning P, Plassmann H. Neuroeconomics: an overview from an economic perspective. Brain Res Bull 2005; 67:343-354
6. McCabe K. Neuroeconomics. In: Nadel E ed Encyclopedia of Cognitive Science. New York; Nature Publishing Group, Macmillan Publishing 2003:294-298
7. Sanfey AG, Loewenstein G, McClure SM, Cohen JD. Neuroeconomics: cross-currents in research on decision-making. Trends Cogn Sci 2006;10:108-116
8. Zak PJ. Neuroeconomics. Philos Trans R Soc Lond B Biol Sci 2004;359:1737-1748
9. Farah MJ. Emerging ethical issues in neuroscience. Nat Neurosci 2002;5:1123-1129
10. Farah MJ. Neuroethics: the practical and the philosophical. Trends Cogn Sci 2005;9:34-40
11. Hauser M. Moral Minds. London; Little Brown 2006
12. Illes J (ed). Neuroethics: Defining the Issues in Theory, Practice and Policy, Oxford; Oxford University Press 2006
13. Marcus SJ (ed). Neuroethics: Mapping the Field. New York; Dana Press 2002
14. Moreno JD. Neuroethics: an agenda for neuroscience and society. Nat Rev Neurosci 2003;4:149-153
15. Churchland PS. Neurophilosophy. Cambridge, MA; MIT Press 1986
16. Honderich T. The Oxford Companion to Philosophy. Oxford; Oxford University Press 1995
17. Darden L, Maull N. Interfield theories. Philosophy of Science 1977;44:43-64
18. Bechtel W (ed). Integrating Scientific Disciplines. Dordrecht; Martinus Nijhoff 1986
19. McCauley RN. Reduction: models of cross-scientific relations and their implications for the psychology-neuroscience interface. In: Thagard P ed Handbook of the Philosophy of Psychology and Cognitive Science. Amsterdam; Elsevier 2007:105-158
20. Bechtel W, Hamilton A. Reductionism, integration, and the unity of the sciences. In Kuipers T ed Philosophy of Science: Focal Issues (Volume 1 of the Handbook of the Philosophy of Science). New York; Elsevier 2007
21. Camerer CF, Loewenstein G, Rabin M (eds). Advances in Behavioral Economics. Princeton; Princeton University Press and Russell Sage Foundation Press 2003
22. Kahneman D. Maps of bounded rationality: a perspective on intuitive judgment and choice. http://nobelprize.org/nobel_prizes/economics/laureates/2002/kahneman-lecture.html
23. Glimcher PW. Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics. Cambridge, MA; MIT Press 2003
24. Hsu M, Bhatt M, Adolphs R, Tranel D, Camerer CF. Neural systems responding to degrees of uncertainty in human decision-making. Science 2005;310:1680-1683
25. Ellsberg D. Risk, ambiguity, and the Savage axioms. Quarterly Journal of Economics 1961;75:643-669
26. Bechara A, Damasio H, Damasio AR. Emotion, decision making, and the orbitofrontal cortex. Cereb Cortex 2000;10:295-307
27. LeDoux J E. The Emotional Brain. New York; Simon & Schuster 1996
28. Preuschoff K, Bossaerts P, Quartz SR. Neural differentiation of expected reward and risk in human subcortical structures Neuron 2006;51:381-390
29. Foot P. The Problem of Abortion and the Doctrine of the

[7] So perhaps we are close to inferring ontological conclusions from epistemological premises.

Double Effect in Virtues and Vices. Oxford; Basil Blackwell 1978
30. Greene JD, Nystrom LE, Engell AD, Darley JM, Cohen JD. The neural bases of cognitive conflict and control in moral judgment. Neuron 2004;44:389-400
31. Koenigs M, Young L, Adolphs R et al. Damage to the prefrontal cortex increases utilitarian moral judgements. Nature 2007;446:908-911
32. McCauley RN, Bechtel W. Explanatory pluralism and heuristic identity theory. Theory and Psychology 2001;11: 736-760
33. van Eck D, De Jong HL, Schouten MKD. Evaluating new wave reductionism: the case of vision. The British Journal for the Philosophy of Science 2006;57:167-196
34. Schouten MKD, De Jong HL. Reduction, elimination, and levels: the case of the LTP-learning link. Philosophical Psychology 1999;12:237-262
35. Glimcher & Rustichini. Neuroeconomics: The Concilience of Brain and Decision. Science 2004;306: 447-452.
36. Dupré J. The Disorder of Things: Metaphysical Foundations of the Disunity of Science. Cambridge, MA; Harvard University Press 1993
37. Cartwright N. How the Laws of Physics Lie. Oxford; Oxford University Press 1983
38. Cartwright N. The Dappled World: A study of the Boundaries of Science. New York, Cambridge University Press 1999
39. Galison, P. Image and Logic: A Material Culture of Microphysics, Chicago, University of Chicago Press 1997
40. Galison P, Stump D (eds). The Disunity of Science; Stanford, Stanford University Press 1996
41. Hacking I. Representing and intervening. Cambridge; Cambridge University Press 1983
42. Latour B. Science in Action: How to Follow Scientists and Engineers through Society. Cambridge, MA; Harvard University Press 1987
43. Suppes P. The plurality of science. In: Asquith P, Hacking I eds Proceedings of the Biennial Meeting of The Philosophy of Science Association 1978 (Vol. 2). East Lansing, MI; University of Chicago Press: 3-16
44. Kim J. Mind in a Physical World. Cambridge, MA; MIT Press 1998
45. Horgan T. Supervenience and cosmic hermeneutics. Southern Journal of Philosophy 1984; 22 (Spindel Conference Supplement on Supervenience):19-38
46. Jackson F. From Metaphysics to Ethics A: Defence of Conceptual Analysis Oxford; Oxford University Press 1998
47. Crane T. The significance of emergence. In: Gillett C, Lower B eds Physicalism and its Discontents. Cambridge; Cambridge University Press, 2001:207-224
48. McLaughin B. The rise and fall of British emergentism. In: Beckermann A, Flohr H, Kim J eds Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism. Berlin, New York; De Gruyter 1992:49-93
49. Kim J. Making sense of emergence. Philosophical Studies 1999;95:3-36
50. Crane T. Elements of Mind. Oxford; Oxford University Press 2001
51. Di Francesco M. Filling the gap, or jumping over it? Emergentism and naturalism. Epistemologia 2005;XXVIII:93-120
52. Dick PK. The Penultimate Truth. London; Vintage 1964
53. Burge T. Mind-body causation and explanatory practice. In: Heil J, Mele A eds Mental Causation. Oxford; Clarendon Press 1993:97-120
54. Rudder Baker L. Metaphysics and mental causation. In: Heil J, Mele A eds Mental Causation. Oxford; Clarendon Press 1993:75-95