Introduction

Bernard Feltz, Marcus Missal and Andrew Sims

What is the relationship between mind and body? Even in antiquity, the contrast between Platonic and Aristotelian conceptions of the relationship between soul and body anticipate this anthropological debate, which has recently been taken up under the banner of a “naturalisation of consciousness.” In the modern period, the dialogue between Descartes’s neo-Platonism and Spinoza’s monism—and later, the Kantian critique—can be interpreted as a continuation of these themes. The passionate discussions in the 20th century between Sartre and Merleau-Ponty about phenomenological “Being-in-the-World” offer a new approach to that which is specifically human. At the heart of all these traditions, language and the capacity for meaning-making play a decisive role. Paradoxically, it is at the very moment when philosophy discovers the corporeal dimension of humanity—in dialogue with the natural sciences, furthermore—that the ubiquity of language in all behaviour is freshly brought into view. Paul Ricoeur and Jürgen Habermas are exemplary of such a position. A human being is a corporeal entity but human behaviour is unintelligible in abstraction from the role of language.

Recent developments in neuroscience have led to a fresh perspective on this set of issues. The relationship between mind and body is not only a philosophical matter. A close dialogue with the experimental sciences is not only possible but indeed also necessary. In recent decades, new methods for studying the brain have produced important theoretical advances and led to the rapid development of various disciplines, among which are neuroscience and the cognitive sciences. The philosophy of neuroscience has itself grown considerably. In this context, one influential line of thinking tends toward the thesis that free will is pure illusion and that the principle of causation in all its rigor leads inexorably to the rejection of a concept of free will likely to contribute to an understanding of human behaviour.

This conducts us to propose two introductory developments : free will and causality.

1 Free Will

There are several epistemological positions that yield different conceptions of free will. Beginning from his work in cognitive psychology, Daniel Wegner

(2002) constructs what he calls a “theory of apparent mental causation” which explains the conscious experience of our own causal efficacy during action as a post-hoc reconstruction. On the basis of Libet’s (1985) experiments, he argues that the conscious experience of acting is pure illusion. It is language—in an a posteriori reorganisation—that inscribes a particular action in the autobiographical memory. The subject then acquires an illusory sense of herself as the author of that action.

In the same line, in the course of a discussion of theories of self-organisation, decision theory, and cognitive models, Henri Atlan (2011) proposes a neo-Spinozist interpretation of human behaviour which is characterised by its total determinism. For him there is a kind of freedom that is linked to the impossibility of predicting behaviour, but this impossibility is purely epistemological. Human behaviour is deeply determined and freedom consists in making this determination our own. In this context, free will proper has no place and is qualified as a “necessary illusion” in that, while we must act as if our decisions were efficacious, this efficacy is fictive on account of the principle of total determinism. Both of these perspectives just discussed are characterised by the conception of a deterministic world and the conception of scientific explanation in terms of causation. Free will is understood as calling into question a principle of causal continuity that these conceptions imply.

This introduction is not the place for a detailed discussion of each of these sorts of models. Simply put, our working hypothesis is that the concept of free will remains pertinent without coming into conflict with any scientific practice that requires the concept of causation in its explanations. Language, as an emergent process of brain activity, contributes to the development of behaviour which count as unified actions taking place over long intervals of time and which makes it possible to conceive of an acting body which is both corporeal and yet aptly described as “free.”

More precisely, intentional action in human beings is a function of both the initial conditions of the distributed neural networks that are involved (that is, the “brain state”: emotions, physiological state, autobiographical and implicit memory) and the circumstances and events at some moment (rest, activity, social interactions, etc...). The expression “free will” is used in order to describe this situation of interaction between an agent whose nervous system reacts to a prior set of events according to both its state at that point and according to an intentional logic. Intentional action refers explicitly to an operation that involves the capacity to represent a future state of the world. The concept of free will refers both to behaviour that has this intentionality and for which it is not entirely determined. Our hypothesis is that the use of language is an activity that allows such an operation.

To defend this conception, we have to analyze more precisely the relation with causality.

2 Causality

From the philosophical point of view, the study of the concept of causality started with Aristotle (384 BC–322 BC) who proposed four different types of causes (material, formal, efficient, final; *Physics* II and *Metaphysics* V 2). During the following centuries, the concept of causality has continued to be interpreted in Aristotelian terms. David Hume initiated the modern approach of causality. He recognized the importance of causal beliefs for human understanding. However, he convincingly demonstrated that causality itself is not observable. Describing colliding objects, David Hume wrote: “When we consider these objects with the utmost attention, we find only that the one body approaches the other; and the motion of it precedes that of the other without any sensible interval” (Hume 1739). The argument of David Hume seems logically impossible to contradict: a necessary connection between events cannot be observed or measured. Only contiguity and succession can be observed. Causality seems indispensable to human understanding but could not be founded rationally and causal inferences are made on the basis of non-causal co-variations. If we follow David Hume’s philosophy, the mind is a white sheet of paper and only learned associations can form the base of human knowledge. Immanuel Kant considered Hume’s conception of causality as deeply unsatisfactory. In Kant’s approach, causality is an a priori category of understanding, a logical necessity for the possibility of experience. Categories of understanding are *a priori* features of the mind. Therefore, for Immanuel Kant, the mind is not a white sheet of paper. The British philosopher Bertrand Russell tried to clause the debate by declaring the concept of causality obsolete: “The law of causality, I believe, like much that passes muster among philosophers, is a relic of a bygone age, surviving, like the monarchy, only because it is erroneously supposed to do no harm” (Russell 1912). However, we suggest that simply giving up the concept of causation at the macroscopic level is unsatisfactory. More specifically, the concept of causation is central to the notion of free will. Indeed, free decisions could cause behavior if humans enjoy free will and this question is central in modern philosophy.

The modern concept of causality has been deeply influenced by physics and psychology during the xxth century and has a deep impact on causality in neurosciences.

According to the physicist Max Born (1949): “Causality postulates that there are *laws* by which the occurrence of an entity B of a certain class depends on

the occurrence of an entity A of another class, where the word ‘entity’ means any physical object, phenomenon, situation, or event. A is called the cause, B the effect.” The concept of “lawlike” necessity is important in the contemporary approach to causation. Moreover, Max Born added that the cause should precede (or at least be simultaneous with) the effect and that there must be some sort of spatial contact between the cause and effect (even if it is by way of a chain of intermediaries). Albert Michotte (1949/1963), profoundly influenced by Immanuel Kant, considered the possibility that humans actually “perceive” causality directly through the activation of an encapsulated specific brain detector receiving a particular pattern of spatio-temporal inputs (Wagemans et al. 2006). Michotte used abstract visual stimuli, such as shapes that moved and collided in various ways, and made detailed manipulations of their spatial and temporal properties. His subjects responded with verbal descriptions of the resulting “scenes,” and Michotte determined whether they thought there was a causal percept (“object A caused object B to move”) or not. Michotte concluded that humans perceive causality as a Gestalt, similar to the way they perceive shape, motion, or other fundamental qualities in the world.

Michotte’s results have been replicated in contemporary experiments (for review, see Scholl and Tremoulet 2000). Whatever their interpretation, Michotte’s experiments and those of his followers clearly show the prevalence of causal judgment in psychology and behavior (Badler et al. 2010; Badler et al. 2012). Suggesting that causality is an illusion is epistemologically counterproductive. Similarly, idea that causal beliefs are elaborated on the basis of passive observations of covariations suffers from obvious limitations. Indeed, readings of a drop of atmospheric pressure on a barometer covaries with storms occurrence. However, nobody will claim that manually changing the reading of a barometer could cause a storm. Genuine causation must be distinguished from spurious. The modern approach to causality inference that is emerging can be thought of in terms of graphs and probabilities. The fundamental idea is that a cause raises the *probability* of occurrence of an effect. Making causal hypotheses is very similar to elaborating a scientific theory from experimental data (Glymour 2001).

More recently, any works are specifically oriented to causality in neuroscience. Craver distances himself from “law-like” necessity causality and defends a mechanistic conception of causality. To explain is to show multilevel mechanisms conducting the transition from state 1 to state 2. In the same line, Woodward proposes an interventionist concept of causality where the articulation between levels of organization in the brain is essential.

These diverse conceptions of causality are present in this book. Each author dialogues with one or other conception in order to think the possibility of free will.

3 Content

The relation between free will and causality is an important focus of this book. That is why we organize it into three parts. In the first part, “Intention and Consciousness,” the objective is to consider a priori theories of the meaning of intentional action in light of our increasing knowledge about the architecture of cognition, and to probe intuitive ideas about the relationship between control, intention, and consciousness. The compatibility of intention with efficient causality is also analysed. In the second part, “Libet-Style Experiments,” while Libet’s famous experiments are generally considered as defending a causality which reject free will, we would like to reconsider these experiments in light of the variety of ways in which they have been instantiated as well as the sorts of theories which they are intended to refute. The third part, “Causality and Free Will,” aims to clarify the ways in which language has an impact on human behaviour, and in the way that it allows a rich scope of flexibility and planning that would otherwise be out of reach. The relation with causality is the main topics, first in articulation with mental causation, finally in the context of emergence.

Specifically, in “Perceptual Decision-Making and Beyond,” Andrew Sims and Marcus Missal extend models of perceptual decision-making in psychophysics in order to elaborate a theory of intentional action that does not rely on the propagation of content from abstract propositional attitudes to sensorimotor representations in the concrete moment of action (e.g. Pacherie 2008). Instead, this model conceptualises intentional action as a process in which quasi-perceptual representations bias the evolution of a “decision variable” into a state space which represents the sensorimotor consequences of a particular outcome. This is intended to be an alternative to the sort of causal theory of action that links action to causation by propositional attitudes.

Markus Schlosser is more optimistic that the traditional picture of action and control can be retained, and illustrates this by walking us through a challenge posed to that picture by dual-process theories of cognitive architecture. In his piece “Dual-System Theory and the Role of Consciousness in Intentional Action” he carefully distinguishes between various kinds of control and guidance, and concludes that the traditional picture can be preserved given qualifications about the role of consciousness.

Nahmias, Allen, and Loveall ask their participants the question: "When do Robots have Free Will?" They do so in order to further probe the importance that attributions of phenomenal consciousness have to ascriptions of free will. Their guiding hypothesis is that phenomenal consciousness matters because for an agent to be free and responsible requires that agent to care about one's choices and their consequences, and that care requires the capacity to feel emotion. Their results provide tentative support for this hypothesis.

The second part of the book contains two chapters that revisit themes in the empirical literature. In his "Free Will and Neuroscience," Alfred Mele considers Libet-style arguments against free will in light of recent updated instances of these studies. He considers two specific arguments for the nonexistence of free will that he takes to be refuted and concludes that recent studies do not do anything to salvage them.

Then, in "Why Libet-Style Experiments Cannot Refute All Forms of Libertarianism," László Bernáth argues that such experiments are able to serve as evidence against forms of libertarianism that do not make metaphysical distinctions between types of decisions. However, he claims that there are a class of libertarian positions for which they are powerless: those that restrict the set of free decisions in a way that rules out their testing in existing paradigms (though he suggests ways in which those paradigms might be modified).

In "Actions and Intentions," Sofia Bonicalzi argues that recent findings in cognitive neuroscience militate against a proposition-style causal theory of action. Instead, she claims, we are better off thinking of action as the product of complex interactions between a number of different systems. Under such a scheme, intentions are not plausibly context-independent, inherently causal, discrete entities. On the basis of specific Libet's experiment interpretations, she suggests that neuroscience can play a constructive role with respect to basic concepts in the philosophy of action.

Finally, the third part of the book returns to the articulation of language and causality in agency and free will in human beings. Anna Drozdewska argues that the problem of mental causation is of central importance in the free will debate, despite the fact that it is often missing from discussions in the extant literature. In "The Mental, the Physical, and the Informational" she suggests that the right approach in this context will be to consider the causal role that information can play in the brain. She motivates the view that this may provide a new approach to the problem of causal exclusion.

Last, in "Free Will, Language, and the Causal Exclusion Problem," Bernard Feltz and Olivier Sartenaer address a similar theme, by considering the ways in which the use of language might instantiate emergent causal powers that produce downward-causal effects. In doing so they bring recent ideas about

diachronic causation into contact with the neuroscience of learning and philosophy of language.

4 Opening Perspectives

The question of free will admits of a diverse number of positions. Even in this book, which brings together contributions by authors largely open to the possibility of free will, it would be hazardous to propose general conclusions to which all could agree. However, in our role as editors of this book, we would like nonetheless to propose some final thoughts on the relation between language and causality with respect to the question of free will.

From a social point of view, it seems difficult to defend the idea that language lacks causal power. We need just be reminded that science is language, and that the products of science—technology—are unthinkable without the causal efficacy of language. Law, economics, political science, rhetoric, all of these are equally languages for which it seems superfluous to argue for their efficacy. Now, if language has causal power from this social point of view, the monist presupposition requires that we posit its efficacy at the individual level as well. If one wonders about the causal efficacy of language at the level of the individual, then in a certain sense the question is *how* to understand this efficacy and not *if* there is any such efficacy. Language operates just as much on the individual level as it does on the social.

On this point, the contributions on language in this collection demonstrate that it is possible to think about the effect of language on the brain while respecting the principle of causal closure. Such a result is important since it allows us to give language a decisive place in our thinking about free will and to bring about a rapprochement between certain philosophers of language and the experimental work of contemporary neuroscientists. In decision-making processes, language is not epiphenomenal. It is perfectly coherent to defend that language has a causal influence in decision-making processes.

However, this causal efficacy of language does not correspond to free will. For example, some thinkers inspired by structuralism defended the idea that language itself determines behaviour through the unconscious (in Lacan's (1966) "Return to Freud") or determines collective behaviour through ideology (as in Althusserian (1970) Marxism), and without the knowledge of the persons concerned. So in order to intervene in the debate over freedom on the basis of language, one needs to go beyond its efficacy.

In contrast with structuralism, Habermas (2007) develops the idea of a productive language in culture that gives rise to meaning: this is what he calls

“objective mind.” The human subject, by learning language, becomes part of this cultural dynamics and becomes able to participate in social conversation, becomes capable of inventiveness, creation, and novelty. The latter is what Habermas calls “subjective mind.” For Habermas, then, language is not an oppressive structure but an open and dynamic structure that allows each individual within it to choose the meaning of her existence. To be free is to behave according to the system of meanings that we have chosen and helped to construct. Not only is language causally efficacious then, but it brings with it the potential for novelty.

Habermas's philosophy of language leads us to defend the idea that the causal power of language allows for a wider efficacy of the processes that involve the self-determination of human behaviour. Language, through which the individual participates in objective mind, allows each person to construct a system of meaning by which he gives sense to the world and orients himself behaviourally. Neural plasticity and learning as implemented in neural networks are the mechanisms which allow to understand the effect of language on the brain. These mechanisms bring us back to these processes themselves as the conditions of possibility for the production of language. Such a perspective makes it possible to meet the difficulties related to the problem of causal exclusion. It produces reconciliation between neuroscience and a conception of the human being as free.

References

- Althusser, L. (1970). *Idéologie et appareils idéologiques d'État* (notes pour une recherche). *La Pensée*, 151, 3–38.
- Atlan, H. (2011). *Selected Writings: On Self-Organisation, Philosophy, Bioethics, and Judaism*. New York: Fordham University Press.
- Badler J, Lefèvre P, Missal M. (2010). Causality attribution biases oculomotor responses. *J Neurosci*. 2010 Aug 4;30(31):10517–25.
- Badler JB, Lefèvre P, Missal M. (2012). Divergence between oculomotor and perceptual causality. *J Vis*. 2012 Jan 1;12(5):3. DOI: 10.1167/12.5.3. Print 2012.
- Born, M. (1949) *Natural philosophy of cause and chance*. Oxford: Clarendon.
- Craver, C.F. (2007). *Explaining the Brain. Mechanism and the mosaic unity of neuroscience*. Clarendon Press, Oxford.
- Glymour, C. (2001), *The Mind's Arrows : Bayes nets and causal graphical models in psychology*. Cambridge, MA : MIT Press.
- Habermas, J. (2007). The language game of responsible agency and the problem of free will: how can epistemic dualism be reconciled with ontological monism? *Philosophical Explorations*, 10, 13–50.

- Hume, D. (1739/1987). *A Treatise of Human Nature*, Ed 2 (Mossner EC, ed). Oxford: Clarendon.
- Lacan, J. (1966). *Écrits*. Paris: Seuil.
- Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioural and Brain Sciences*, 8, 529–566.
- Michotte, A. (1949/1963). *The Perception of Causality*, (Miles TR, Miles E, transl). New York: Basic Books.
- Pacherie, E. (2008). The Phenomenology of action: a conceptual framework. *Cognition*, 107, 179–217.
- Russell, B. (1913). On the Notion of Cause, *Proceedings of the Aristotelian Society*, Volume 13, Issue 1, 1 June 1913, 1–26.
- Scholl, B.J., Tremoulet, P.D. (2000). Perceptual causality and animacy. *Trends Cogn Sci* 4:299–309.
- Wagemans, J., van Lier R., Scholl B.J. (2006). Introduction to Michotte's Heritage in Perception and Cognition Research. *Acta Psychol* (Amst) 23:1–19.
- Wegner, D.M. (2002). *The Illusion of Conscious Will*. Cambridge: MIT Press.
- Woodward, J. (2003). *Making Things Happen*, Oxford University Press.