Introduction

Gottfried Leibniz left behind a corpus of writings that is impressive both in size and in breadth: it contains discussions of an enormous array of areas of inquiry, both inside philosophy and out. One comparatively neglected portion of this corpus is Leibniz's contribution to what we would today call 'the philosophy of action.' That is unfortunate because Leibniz's discussions of agency are sophisticated and often compelling. In fact, they are far more compelling than one would expect, given the notorious eccentricity of Leibniz's metaphysics. To mention just a few examples of his eccentric metaphysical doctrines, take the claim that our minds do not, strictly speaking, interact with our bodies or the denial of causal interaction among finite substances. Or consider the doctrines that every substance perceives everything that happens in the entire universe and that the only ultimately real things are immaterial. Or take, finally, Leibniz's rehabilitation of substantial forms and teleology, which are rejected almost universally by modern philosophers. All of these idiosyncrasies may suggest that Leibniz's views on action are bound to be implausible and useless for advancing our understanding of agency. Yet, we will see that quite the opposite is the case.

The philosophy of action, as practiced in the twentieth and twenty-first centuries, analyzes a broad range of philosophical issues surrounding the notion of agency. These issues include the freedom of the will, shared agency, moral responsibility, the distinction between things that we do and things that merely happen to us, what it means to possess control over one's actions, and whether it is possible to act against one's better judgment. As this book shows, Leibniz discussed all of these topics. Many of his views on agency are directly relevant to present-day debates, and we can learn a number of things about agency from him.[1]

Leibniz's discussions of agency are more subtle and insightful than those of most (if not all) of his contemporaries. And this is no coincidence: his

[1] Analytic philosophers working in the philosophy of action are not typically aware of the fact that Leibniz made important contributions to their field. For instance, a recently published companion to the philosophy of action (O'Connor and Sandis 2010) contains chapters on the views of six early modern figures, but Leibniz is not among them. A notable exception is Susan Wolf, who explicitly describes part of her project as inspired by Leibniz (1990: 103).

metaphysical idiosyncrasies force him to pay particularly close attention to distinctions that his contemporaries simply take for granted. In many cases, this leads Leibniz to theories that are far less eccentric than their metaphysical foundations and that possess significant plausibility and explanatory power. One example is Leibniz's compatibilist theory of freedom. While other early modern compatibilists pay very little attention to internal impediments to free agency,[2] Leibniz's denial of inter-substance causation prompts him to take this type of impediment extremely seriously. This, in turn, makes his account of freedom, and his moral psychology more generally, far more powerful. The resulting theory of freedom is an intriguing combination of agent-causal views, the doctrine that being free means being determined by the good, and the doctrine that free actions have to flow from the agent's real self.

There are other cases where Leibniz arrives in familiar territory from extremely eccentric points of departure. One is the distinction between what we today call 'autonomous agency' and 'nonautonomous agency'; another is the closely related distinction between acting and being acted upon. Leibniz discusses these distinctions in terms of self-determination and end-directedness. Actions that we would describe as autonomous are self-determined and end-directed in a more demanding way than other actions. Similarly, active states differ from passive states in their self-determination and end-directedness. Leibniz's starting point is his idiosyncratic claim that finite substances do not, in metaphysical strictness, interact causally with each other. Because this doctrine makes it difficult to distinguish between activity and passivity, Leibniz is forced to be particularly attentive to the different ways in which states can originate in a finite substance. That scrutiny pays off: Leibniz manages to isolate factors within an agent that can undermine agency and autonomy. In fact, Leibniz's account resembles that of several prominent philosophers of action in our own day.

Some advantages of Leibniz's solutions arise directly from his peculiar metaphysical commitments and are hence unlikely to command broad appeal. For instance, some of the particularly attractive aspects of his theory of freedom depend on his doctrine that finite substances do not interact. Other advantages, however, do not depend on eccentric Leibnizian commitments. For example, Leibniz's accounts of control, weakness of will, compulsion, and moral responsibility – or, at least, their most central features – are compatible with a wide range of metaphysical systems. These accounts are more likely to be appealing to contemporary philosophers of action. Yet, it is fascinating and useful to explore both types of advantages. Studying Leibniz's theory of freedom, for instance, is helpful in part because it illustrates the costs of securing a particularly demanding type of independence from external determination.

[2] See Gary Watson, who argues that this is a widespread problem among classical compatibilists (2004: 164).

Let me make a few remarks on the book's scope and methodology. First, it is not my goal merely to mine Leibniz's writings for doctrines that might be useful to contemporary philosophers of action. Instead, the book attempts to understand Leibniz on his own terms and within his historical context, unconstrained by the prospective utility of the resulting interpretation for contemporary philosophy. This often requires looking at Leibniz's predecessors and contemporaries. While I do think that many of Leibniz's views on agency are promising and helpful, defending their viability is not part of my project. Relatedly, the book does not explore in detail the similarities and differences between Leibniz's theory of action and the theories of more recent philosophers. Even though that is a worthwhile project, it would distract too much from my primary aim of providing an interpretation of Leibniz's views on agency in their philosophical and historical context. Hence, the book merely mentions some connections between Leibniz's doctrines and the contemporary philosophy of action, especially when these connections can help us understand Leibniz's views.

Given the book's methodology, it may seem worryingly anachronistic to talk of a Leibnizian philosophy of action in the first place. After all, Leibniz does not appear to use the term 'philosophy of action' himself, nor was it a commonly acknowledged subfield of philosophy until the twentieth century. Yet, that does not make this book's project anachronistic. It is clear, after all, that Leibniz answers many of the questions that we today associate with the philosophy of action; it is similarly clear that these questions are closely interrelated for Leibniz. Hence, treating Leibniz's answers to those questions as a unit is not only legitimate but also natural and useful. The term 'philosophy of action' is merely a convenient way of referring to that set of questions. However, using that term should not make us lose track of the fact that, for Leibniz, there are intimate connections between human agency and the changes that occur in animals and even in plants and other inanimate things. All substances act, in a broad sense, and understanding more primitive kinds of activity can help us understand human agency. As a result, significant portions of Chapters 1, 2, and 3 are concerned quite generally with the causation of changes in Leibnizian substances. The book thus goes beyond what we would today categorize under 'philosophy of action.'

One further methodological choice concerns the parts of Leibniz's corpus that the book takes into consideration. Leibniz's philosophical views undergo a number of changes in the course of his long career. While there is a substantial amount of controversy over the precise nature, significance, and timing of these changes, interpreters often divide Leibniz's career into three broad stages: the early period, the middle period, and finally the late, mature, or monadological period. The central focus of this book is the late period, which I take to begin around the middle of the 1690s. This is mainly because texts from the mature

period contain particularly sophisticated and detailed discussions of agency. It would, of course, be valuable and interesting to examine the development of Leibniz's philosophy of action, starting with his earliest philosophical writings. Yet, it is not possible to delve into such developmental questions here, given the breadth of topics that this book discusses, as well as the sheer size and complexity of Leibniz's corpus. While I occasionally bring up passages from earlier texts when they are particularly helpful – for instance, when they answer a question that the mature texts appear to leave open – my chief focus is on the late period.

The structure of the book is the following. Chapter 1 introduces readers to the book's protagonists: Leibniz's simple and mind-like substances, or monads. Leibniz describes the fundamental nature of monads in several different ways – for instance, as substantial forms, entelechies, primitive forces, and laws of the series. Even more perplexingly, he talks in many different ways about the changing modifications of monads and the relation of these modifications to their subjects. As one might expect, scholars disagree widely about the correct interpretation of these fundamental building blocks of Leibniz's philosophical system. Chapter 1 aims to untangle some of these issues in order to clear the way for the work of future chapters. It will also serve to introduce nonspecialists to the most central aspects of Leibniz's ontology. The chapter pays particularly close attention to appetitions and perceptions, the two fundamental types of monadic states. It argues that the distinction between these two kinds of states is important to Leibniz, that every state of a monad is efficiently caused by that monad itself, and that a monad's actions consist in its bringing about new perceptions. The chapter also examines the different types of appetitions and perceptions acknowledged by Leibniz, as well as the most plausible way to understand the influence of one monadic state on another.

Then, in Chapter 2, I explore a central commitment of Leibniz's metaphysics of action: the doctrine that all states of any substance originate within it, or arise "out of its own depths." All monads possess a far-reaching independence from other things, which Leibniz calls 'spontaneity.' This doctrine is undeniably radical. Many who encounter it – be it today or in Leibniz's own time – find it absurd, in part because it is difficult to square with the commonsensical distinction between acting and being acted upon. Yet, Leibniz can capture that difference by distinguishing three ways in which changes can originate in a subject; I call them 'metaphysical spontaneity,' 'agent spontaneity,' and 'rational spontaneity.' That threefold distinction, in turn, is tremendously important for Leibniz's philosophy of action, quite apart from helping him to distinguish actions from passions. In particular, it allows him to claim that some of the desires and emotions that occur in our minds are not ours in an important sense. They are external to our true selves. As a result, Leibniz – like several prominent philosophers of action today – can distinguish between situations in

which we act autonomously and situations in which we are controlled by desires or emotions that undermine our autonomy.

Chapter 3 investigates a second aspect of monadic independence: monads are not only the sources of their actions but also set the ends of those actions. For Leibniz, all monadic activity is immanently end-directed, or an instance of what is traditionally called 'final causation' or 'teleology.' The chapter first explores Leibniz's motivations for viewing teleology as ubiquitous. Next, it argues that there is a tight connection between spontaneity and teleology and that it is useful to distinguish three different types of teleology, parallel to the three types of spontaneity described in Chapter 2. I call them 'metaphysical teleology,' 'agent teleology,' and 'rational teleology,' respectively. This distinction, in turn, allows Leibniz to view end-directedness as ubiquitous without anthropomorphizing the least perfect monads and without trivializing the end-directedness of the most perfect actions. In fact, it allows him to make teleology in its most demanding form a crucial component of his accounts of freedom, control, and moral responsibility, as I argue in later chapters. Finally, Chapter 3 examines the lowest type of teleology in more detail. It argues, against the overwhelming majority of interpreters, that monads perform many of their actions simply because their natures prescribe these actions, not because these actions are or appear good.

Chapter 4 tackles another central issue in the philosophy of action: attributability, that is, the question of when an action is properly attributed to an agent. Here, one major obstacle for Leibniz is his endorsement of concurrentism: he accepts the traditional theistic doctrine that creatures require God's cooperation for all of their actions. Yet, when God acts together with a creature, the resulting action is supposed to be the creature's action alone, not God's. This is particularly important for sinful actions, which for theological reasons must not be attributed to God. Chapter 4 examines how it is possible for actions to be attributable only to the created agent, even though God and the creature are acting together. It argues that we can answer that question by taking seriously the roles that final and formal causation play in Leibniz's account of agency.

The fifth chapter focuses on a type of agency that has long been a central topic in the philosophy of action: free agency. Leibniz's theory of freedom is what we would today categorize as a version of agent-causal compatibilism. In other words, Leibniz believes that freedom is compatible with determinism, and he also holds that free actions are caused by agents rather than by events internal or external to agents. One goal of the chapter is to elucidate Leibniz's compatibilism. I will show that Leibniz's metaphysical commitments allow him to circumvent notorious shortcomings of other compatibilist theories. The chapter's second goal is to take a fresh look at the sense in which Leibnizian free actions are contingent. I argue, against many other interpreters, that Leibniz

does not ultimately take freedom to be incompatible with necessitarianism *simpliciter*. Instead, he takes it to be incompatible only with the kind of necessitarianism according to which everything is necessitated in a value-neutral way. Therefore, what Leibniz calls 'contingency' has an intimate connection to the most demanding type of final causation.

Chapter 6 examines Leibniz's responses to three notoriously difficult problems: (a) the problem of explaining in what sense free agents have control over their actions, (b) the problem of explaining ostensibly weak-willed actions, and (c) the problem of distinguishing weak-willed from compelled actions. Leibniz explicitly discusses the notion of control – or, as he usually calls it, 'mastery' – and, this chapter argues, he manages to make room for a meaningful and desirable type of control. For Leibniz, we possess control to the extent that our rational judgments and rational desires are able to influence our actions. He acknowledges that we sometimes lack the ability to control our actions – namely, when our passions are so powerful that they would outweigh even the strongest rational desire. Yet, Leibniz insists, there are indirect ways to make our rational desires succeed: we can take steps ahead of time that drastically reduce the influence of the passions. Some of the resources that allow Leibniz to give a convincing account of control also allow him to acknowledge a form of weakness of will. That is surprising because he holds that all intentional actions are determined by what the agent perceives as good. Moreover, Leibniz can capture the difference between weakness and compulsion – a hard problem for determinists.

The seventh and final chapter addresses two questions concerning moral agency: what it takes to be a moral agent and what it takes to be morally responsible for particular actions. Moral agency is intimately connected to many of the concepts investigated in previous chapters, though these connections are less straightforward than one might initially think. One particularly important result of this chapter's discussion is that teleology is central to Leibniz's notion of moral agency. Another important result is that agents are morally responsible for some of their unfree actions. Finally, a particular kind of ability to do otherwise is required for moral blame but not for moral praise.

Some general themes will emerge in this book. One important theme is that the will, appetitions, and teleology are key players in Leibniz's discussions of agency. In order to understand what it means to act and what differentiates different types of activity, we need to look not just at cognition but also at appetition. This goes against existing scholarship on Leibniz's moral psychology; with a few notable exceptions, other scholars focus almost exclusively on perceptions and the intellect. Another, related general theme of the book is that Leibniz is less of an intellectualist than commonly thought. While this comes up in several chapters, it becomes particularly clear in Chapter 6, where I argue

that knowing what is best, all things considered, is often insufficient for doing the right thing. Leibniz acknowledges that taming our irrational passions requires us to be extremely resourceful. For instance, we sometimes need to distract ourselves, cultivate beneficial passions, and use sensory images to our advantage. He is far less optimistic than one might initially expect about our intellect's ability to take on the passions directly.

1 Monads and Their Actions

The protagonists of this book are monads, or Leibnizian simple substances. Despite their simplicity, it is surprisingly difficult to figure out what exactly these monads are and how they act. Leibniz describes the fundamental nature of monads in a number of different ways: as substantial forms, entelechies, primitive forces, and laws of the series, to name just a few. And as if that were not confusing enough, it is unclear how exactly Leibniz understands the changing modifications of monads and the relation of these modifications to their subjects. It is no surprise, therefore, that there is no consensus about the correct interpretation of these fundamental building blocks of Leibniz's philosophical system. The present chapter aims to address these issues in order to shed light on some of the most basic features of monadic agency and clear the way for the work of future chapters. Of course, I cannot – and do not need to – answer all of the numerous questions about Leibniz's fundamental ontology. Instead, I focus on the questions that are directly relevant for my interpretation of Leibniz's philosophy of action.

One can learn a lot about what monads are and how they act by looking into the reasons that Leibniz cites for rejecting a purely mechanistic view of nature. These reasons include accounting for the reality, unity, and activity of natural things. Because understanding these reasons proves helpful for my interpretation of monadic agency, I will examine them briefly. Next, I turn to some of the ways in which Leibniz characterizes monads: his claims that they are similar to Scholastic substantial forms and Aristotelian first entelechies, that they consist in primitive force, that their only internal qualities are perceptions and appetitions, and that they contain their entire histories. I will also sketch some of the most important differences between the three types of monads that Leibniz distinguishes, that is, bare monads, nonrational souls, and minds. Finally, I consider the types of causation that are involved in monadic activity. All of these elements will become important later in the book, though they are also interesting in their own right.

1 The Fundamental Nature of Monads

Let us start at the beginning – the very beginning, in fact. For Leibniz, it all begins with God and his ideas. Leibniz's God is all-powerful, all-knowing, and perfectly good: he can do whatever is metaphysically possible, knows everything that there is to know, and wills only what he recognizes as best. God's power and goodness will take center stage in later chapters where I discuss divine freedom and the contingency of the created world. For now, however, let us focus on God's knowledge, or his intellect. God's omniscience, according to Leibniz, "encompasses every idea and every truth, that is, all things – simple or complex – that can be an object of the understanding" (CD 13). In other words, God eternally possesses ideas of all metaphysical or logical possibilities, as well as knowledge of all necessary truths (see COE 21; letter to Morell, September 29, 1698, A 1.16.164). As a matter of fact, Leibniz views the divine intellect as the source or ground of all possibilities and necessary truths: without God's intellect, nothing would be possible or true (M 46; CD 7f.; T 184; 189). We must be careful, however, not to confuse this dependence on the divine intellect with the Cartesian doctrine that God creates eternal truths. Leibniz's God does not ground these truths by willing that they be true, but merely by having an intellect that contains them.

Because God's intellect contains everything that can be known, it also contains knowledge of the goodness and badness of all different possibilities: the ideas in God's intellect "represent to him the good and the evil, the perfection and the imperfection, the order and the disorder, the congruity and the incongruity of possibles" (COE 21). As a result, the divine intellect can also compare different possibilities and judge them with respect to their goodness. And that is precisely what God does in order to figure out what to create: he judges, based on his perfect knowledge of all possibilities, which possible world is the best.[1] Because of God's perfect goodness, the world that he judges to be best is the world that he subsequently creates.

What exactly does God create when he creates a world? In Leibniz's mature writings, the answer is that God creates an infinity of monads – that is, an infinity of simple, immaterial, mind-like substances.[2] Monads are the fundamental building blocks of the created world, and everything else that has a place in Leibniz's ontology depends on, or is grounded in, them. One crucial aspect of the relationship between the divine intellect and finite monads is that each finite monad corresponds exactly to an idea in God's intellect. As already seen, the

[1] How precisely God does this is controversial and complicated. Yet, for present purposes, we need not worry about the details.

[2] God himself is a monad, in fact (see a letter to Bierling, August 12, 1711, G 7:502; supplement to a letter to Des Bosses, February 15, 1712, LDB 233f.), though, of course, he differs from created monads in a number of important ways.

divine intellect eternally contains ideas of all possibilities, which include ideas of possible finite substances. God's ideas of these possible substances, moreover, contain information about every change that these substances will (or would) undergo if they are created. In the middle period, Leibniz usually calls these ideas 'complete concepts'; in his mature writings, he more frequently calls them 'possibles' or 'essences.' When God creates the world, he actualizes some of these possibles, namely the ones that together constitute the best of all possible worlds.

One significant complication for understanding the fundamental nature of created monads is that the details of Leibniz's theory of substance, or at least the terminology he uses to express it, are not entirely stable, even within the mature period. In some texts, Leibniz identifies monads with entelechies or substantial forms (e.g. M 18; 63; T 396); in others, he describes monads as made up of entelechies or substantial forms together with passive force, or with matter (e.g. ONI 11). Despite this apparent instability, however, a large portion of Leibniz's ontology remains the same. As we will see later, he thinks throughout the mature period that finite substances possess, or are,[3] primitive active and passive forces,[4] from which all of their modifications arise, and that their only fundamental internal modifications are appetitions and perceptions. The ways in which he distinguishes rational monads from lower monads, as well as his characterization of different types of appetition and perception, appear to be stable as well, at least in the most important respects.

1.1 Reality, Unity, and Activity

The essay "New System of Nature" is Leibniz's first published account of his mature views. It is an autobiographical account of the process that led him to realize that there must be something over and above matter, namely true unities.[5] He lists three closely related reasons for positing immaterial unities: (a) matter is not in itself fully real, (b) matter lacks unity, and (c) matter lacks activity. The first two reasons are closely related; Leibniz argues that "a multitude can derive its reality only from *true unities*," but that matter, since it is

[3] Even though sometimes Leibniz talks of substances as possessing primitive forces, that is probably misleading. There is good evidence that, strictly speaking, substances *are* primitive forces. I argue this in Jorati (forthcoming b). However, I will bracket this complication for the purposes of this book because it does not appear to impact my interpretation of monadic agency.

[4] God is different from finite substances in this respect: he does not possess any passive force, but only active force. See, for instance, a letter to Remond, February 11, 1715 (G 3:636/L 659).

[5] Leibniz does not actually use the term 'monad' in "New System," but he is clearly describing the entities that he elsewhere calls 'monads.' The first mention of the term 'monad' appears to be in an unfinished letter to the Marquis de l'Hôpital, dated July 22, 1695 (GM 2:295); see Rutherford (1995b: 166n24), Garber (2015: 165). The first published text in which the term is used is "On Nature Itself" (1698).