1
Introduction

The agent is the one that does things. There somehow in the midst of all the things that cause the agent to move, you find an agent in turn causing things. You find action. The agent displays activity.

Those things that are not agents do nothing. There in the midst of all the things that cause them to move and that they in turn cause, you find . . . mere happenings, nothing else. No action; all passivity.

—Really?

Hogweed is not an agent—not in the sense above intended. And yet hogweed will give you a nasty rash. Hogweed will render your skin extremely sensitive to the sun. You might end up with third degree burns, and scars. Hogweed does things.

Come to think of it, what doesn't do things? Numbers probably. Absences maybe. But lots of things do things. Trees fall. Stars burn. When passing through heavy water, neutrinos leave a kind of trail.

When I say the agent is the one that does things—when I engage in this philosopher's way of talking about agents—I must have a special notion in mind.

At its most metaphorical, the notion is of two planes of existence.

On one plane are mere happenings, and the things that partake in them. On this plane festers the hogweed, falls the tree, slowly cools the dead star, bombs quietly across space and time the neutrino.

On the higher plane are agents. Doing things. But for real.

This picture is gnomic. It frustrates. And yet it allures. The history of philosophical reflection on action gives the distinction between activity and passivity different names, and attempts to explain the distinction in different ways. But philosophers circle the distinction repeatedly (for a nice recent discussion, see Hyman 2015, both chapter 1 and the appendix). Aristotle wants to know the difference between being cut and cutting. Hobbes wants to know the difference between vital motions, like the motion of the blood, and voluntary motion, as in bodily action. Wittgenstein wants to know the difference between my arm going up and my raising it.

The Shape of Agency: Control, Action, Skill, Knowledge. Joshua Shepherd, Oxford University Press (2021).
 DOI: 10.1093/oso/9780198866411.003.0001

I'm hooked. Agents do seem to be importantly distinct from non-agents. Agents seem to be a special kind of thing, possessed of unique capacities and thereby capable of special kinds of achievements.

In this book I give voice to this thought. I offer a perspective on agency—on its minimal conditions and some of its exemplary instances.

The view of agency built in this book is not exactly reductionist. But it is stripped down. It is individualistic. And it is in large measure, at least in exposition, ahistorical. This is not to say it is not a product of its time. One could trace a lineage that draws significant inspiration from Aristotle, endorses some ideas found in the modern period (in, e.g., Hobbes), then begins to pick up steam with thinkers like William James, and past him diverse mid-twentieth-century sources like Gilbert Ryle, or Maurice Merleau-Ponty, and from there moves quickly towards the present, adding and pruning layers like some kind of self-critical Fibonacci sequence, by way of Hector-Neri Castañeda, Alvin Goldman, Marc Jeannerod, Myles Brand, Daniel Dennett, Michael Bratman, Alfred Mele, Elisabeth Pacherie.

From 2020, we can look back on the development of theories of agency and action over time, and see that a lot of what passed for philosophical reflection on action in the history of philosophy appears now as speculative psychology. Progress in the sciences of the mind has been slow, and full of fits and starts, but it continues. And from here it seems that earlier accounts of agency, which leaned heavily on ideas about, and relations between, faculties or capacities called reason, or the passions, or the intellect, or the will, are under pressure to accommodate different, mechanistic taxonomies that make reference to notions like associative learning, task set construction, sensorimotor adaptation, motor schema, representational format, metacognition, cognitive control, and so on. These mechanistic taxonomies and these neuropsychological concepts do not render philosophical reflection on agency irrelevant, of course—if anything, the science of agency raises as many philosophical questions as it answers. The point is simply that philosophical accounts of agency and agentive phenomena must now be developed with an awareness that the parts that compose agents are being spliced into fine levels of grain by a range of intersecting disciplines—neurobiology, cognitive psychology, cognitive ethology, motor physiology, cybernetics, and more. What this awareness has done to the book you are reading is that I have written a book full of concerns that are somewhat abstract and thoroughly architectural.

In fact my book is architectural in two senses. In one way I am concerned with broad structures. I am less concerned with the material that composes the skeleton, than with the shape of the skeleton. I am concerned with the basic building blocks of agency in chapters 2 through 5. In chapters 6 through 8 I am concerned with the abstract form of agency, and with how agents, qua agent, might display excellence of form.

The second sense in which my book is architectural in that, rather than try to capture the essence of pre-existing agentive notions, I am trying to build something new. My approach is not conceptual analysis, but more like Carnapian explication (Carnap 1950; Justus 2012; Shepherd and Justus 2015), or what lately people have been calling conceptual engineering. Some revision of pre-existing notions is involved. But the aim is to actually capture the reality underneath, or at least to develop accounts of phenomena that might, even if flawed in some respects, promote understanding of the nature of agents. I would ask readers to bear this in mind when reading the accounts I offer of control, voluntary control, intentional action, and even skill. I am aware that usage of these words varies, and that alternative accounts are available. My claim is that the accounts I develop accurately capture phenomena of importance, and that promote fruitful theorizing, even if some departure from intuitions or common usage is required.

The shape of agency that I trace in this book comes primarily in the form of accounts of five agentive phenomena: control, non-deviance, intentional action, skill, and knowledgeable action. These accounts are interlinked. Control is closely related to non-deviance. Both are important for intentional action. Control, non-deviance, and intentional action undergird an account of skill. And everything that goes before helps elucidate knowledgeable action.

The aim is not to make good on the metaphor of two planes so much as explain its allure by explaining the ways in which agents, as agents, are special. Agents are special things, in that they are unique amalgamations of properties, of causal powers. They have a unique kind of structure. This is not to say that they do not fit perfectly within the natural order, whatever that is.

The Blueprint

Chapters 2 through 5 concern basic building blocks of agency. In chapter 2 I develop an account of control's possession. Key notions are the agent's plans

(or plan-states), the agent's behavioral patterns, and the circumstances in which plans (or plan-states) help to cause an agent's behavioral patterns. An agent possesses control over her behavior when she is constituted in a way such that in circumstances we must carefully specify, her behavioral patterns repeatedly and flexibly match her plans for behavior.

In chapter 3 I develop an account of non-deviant causation.

In chapter 4 I leverage the earlier discussion to offer an account of control's exercise. Roughly, control's exercise essentially involves non-deviant causation, and non-deviant causation is what happens when agents that possess control behave in normal ways in implementing plans in certain circumstances. I also apply this account, along with additional considerations, to offer an explication of voluntary control, and to illuminate voluntary control's relationship to nearby notions of direct control, and indirect control. I also extend the explication of voluntary control to the notion of what is "up to" an agent.

In chapter 5 I develop an account of intentional action. It transpires that intentional action is the exercise of a sufficient degree of control in bringing behavior to approximate a good plan. Laying out this view of intentional action takes some work, and I anticipate complaints. So I go on to consider a number of ancillary issues and potential objections. I also consider this account in relation to frequent complaints levied against causalism about intentional action.

Chapters 2 through 5 might be thought of as the book's first part. Chapters 6 through 8 are a second part, with chapter 6 as a kind of hinge. The main aim in this second part is to work towards an understanding of agentive excellence.

In chapter 6 I discuss the nature of agency. I do not lay out a specific account, but I try to render vivid the thought that agency is essentially a matter of a system structured so as to make appropriate the application of behavioral standards—frequently, rational standards—to the system, at least some of which the system is able to satisfy. This discussion foregrounds the accounts I offer in chapters 7 and 8. These are accounts of modes of agentive excellence.

In chapter 7, it is skill at issue. I develop thoughts about the targets of skill—especially about what I call an action domain. I also offer a novel account of skill, and of skill's gradability. I then consider the role of knowledge in an account of skill, and argue that although knowledge is frequently critical for skill, it is not necessary.

In chapter 8, knowledgeable action—action that in turn involves knowledge of what I am doing and how—is at issue. Many have found knowledge of action particularly interesting, and epistemically unique. I develop an account of the epistemic credentials of knowledge of action, I discuss competitors, and I illuminate how action that involves knowledge of action qualifies as a mode of agentive excellence.

Let's get it.