

AN INTRODUCTION TO THE PHILOSOPHY OF AGENCY

Luca Ferrero

1 A basic distinction

Much of what happens, from the motions of the galaxies to the interactions of subatomic particles, is the mere product of the causal forces that operate in the universe—seemingly, these events have no point and they aren't anybody's doing: they just happen. But once in a while, especially closer home, some of the things that happen appear to have a distinctive character: they seem to unfold 'with a view' to some destination, to have a point to them, and be the doings of someone (or at least of something): a hummingbird drinking nectar from a flower, a dog chasing a ball, you reading this introduction, the clown making some children laugh, and your national soccer team celebrating their victory at the World Cup—just to give some examples.

A similar distinction is also especially noticeable about yourself. Some things just happen to you—being knocked to the ground by a gust of wind, spontaneously bleeding from your nose, or falling asleep—but many other things you *do*: you get up (or at least try to) when you are knocked down, jump in celebration when your team scores, play the Hammerklavier Sonata, converse with your friends, build a house, and turn your head to admire the starry heavens above you—just to give some examples (not to mention the things that that you might be able to do, even if you are never going to do, such as, hopefully, knocking someone to the ground or punching someone so hard to make their nose bleed).

For the most part, it seems that we can easily classify the things that happen, both in the universe at large and around us, between those that *merely* happen, with no point or purpose, and those that are *done by* someone or something, with some kind of point or purpose. Sure, there are going to be some ambiguous or borderline cases—when the children laugh at the clown, when you sneeze, or when you are in the grip of some compulsion, for instance. But the reason why we find these cases somewhat difficult to classify is that they seem to partake of some of the special features of the straightforward and uncontroversial examples of 'doings'—or what we also refer to as acts, actions, or activities.

The distinction between mere happenings versus doings is something that matters a lot to us. Whenever you are dealing with some item (an object, a situation, or an occurrence)—whether by interacting with it, trying to explain it, or appraising it—you first need to determine where that item falls in the basic distinction between mere happenings and doings. Are

you dealing with something that is just happening, or has no capacity to do anything, or is just suffering the effects of someone else's doing? Or are you dealing with something that is a doing, or has the capacity to do things and possibly actively exercising this capacity? The very character and nature of your interactions, explanations, and appraisals—the category to which they belong—is affected by the preliminary classification of their objects into the two basic categories: mere happenings versus doings.

Simply put, you would not try to make sense of a mere happening as 'intelligible' in its unfolding in terms of some of its own goals or purposes; you would not hold a mere happening or its source to any standards of success or appropriateness, and certainly not to any kind of responsibility. It would be both a category and a practical mistake to try avoiding a dog chasing you as if it were a flying ball, and vice versa. These differences become even more marked and significant when they bear on our relationship with fellow human beings (and with ourselves). This is because the explanation and assessment of, and interaction with the doings of human beings can be framed not just in purposive and teleological terms—that is, in terms of their goals or aims—but also, and much more importantly, in terms of such notions as 'reasons,' 'rationality,' 'justification,' 'responsibility,' and 'morality.'

A dog chasing you might have some reasons to do so but it would not be cognizant of them as such. Hence, it would in principle be impossible to ask the dog for a justification of its action, and inappropriate to hold the dog accountable for it. But a human being chasing you is, in principle, expected to 'respond' for this action of theirs: they are in principle under a demand to offer some plausible reason in support of their behavior. If this subject were unable to provide such reasons, they *should* stop engaging in it or, if it is too late, they *should* be held accountable for that performance and its consequences (and the use of the modal 'should' is a major element of what is ultimately at stake here).

This is not to say that everything we do is appraised and 'held' to such standards and expectations. But when it is not, it is because our conduct should somehow be exempted, exonerated, or excused—that is, a special dispensation is to be made against the default expectation that we are dealing with ordinary *purposive* and *intentional* conduct, with its full range of normative, legal, and moral implications.

2 Agency and philosophy

The distinction between the different kinds of occurrences that I have just sketched is a fairly intuitive and familiar one. But matters get more complicated and controversial once we look at what philosophy has to tell us about this distinction.

The simplest way for a philosopher to explain this intuitive distinction is in terms of the concept of 'agency.' The philosopher would say: some entities—agents—are capable of agency, whereas non-agents aren't. For this to be informative, however, we would need to know more about the nature of the capacity of agency. If we just say that agency is the capacity to act or to do things, we have hardly explained anything. What we are looking for, after all, is an account of those distinctive modes of behaving that we call 'acting' or 'doing.'

Things might not look much better if we consider some of the standard ways in which the concept of agency is introduced by philosophers, such as agency as the capacity to 'make things happen,' 'make a difference,' or 'cause some kind of change.' The trouble is that these characterizations can be used in ways that appear too broad. These expressions might refer, after all, to such things as the sun causing sunburn or the vinegar curdling the milk in the paneer.

So, any quick characterization of the capacity of agency seems either uninformative or too broad. Can we do better? I suspect not, at least not at this stage. But we should not despair. As Pamela Hieronymi correctly says at the opening of ‘Agency and responsibility’ (Chapter 25), “‘Agency’ is a term of art. Its meaning and use might be discovered by reading and studying this volume.” This is indeed the spirit in which we should approach our philosophical investigation: by reading and reflecting on the topics discussed in this handbook, we should be able to better *articulate* the various features, dimensions, and implications of the capacity of agency. In doing so, we could temporarily rely on a rough-and-ready definition—agency as the capacity to do things or to make things happen, say—but any such definition should only be used as a stand-in for the capacity that will eventually be shown to account for the intuitive distinction between purposive (and possibly intentional) conduct and mere happenings.

This is why this introduction is not going to offer a definition of agency (and why you won’t find any chapter in this handbook with such title as ‘What is agency?’ or ‘The nature of agency’—but only chapters with titles such as ‘Agency and x’ or some qualification of agency, say ‘Animal agency’ or ‘Aesthetic agency’). So the best I can do in this introduction is to point to a cluster of questions and issues that arise when discussing the differences between doing and mere happening, between making things happen rather than having things happen to you, between being active rather than passive, and so forth—with an understanding that these are just intuitive distinctions in need of further refinement as we proceed with our philosophical investigation. I will continue to use ‘agency’ unapologetically but always with the implicit understanding that it refers to what we are trying to explicate in raising these various philosophical questions.

3 The exercises of agency

3.1 Agency as a capacity

Agency is a *capacity*: it can be possessed without being exercised or manifested (e.g., when we are temporarily unconscious, we possess the capacity but we are not exercising it). Those who possess this capacity are ‘agents’ (as opposed to non-agents), but they are not playing the *role* of agents when they are not exercising it (at those times, agents are rather playing the role of patients). For instance, if my arm moves because of a spasm or because someone has grabbed my arm and moved it, I am not *the* agent (as role) of that movement, even if I am still *an* agent (as a kind of being). (The distinction is in principle straightforward but there are cases where the exercise of one’s agency is partly under the control of someone else’s agency, as it might happen when someone is under the spell or the authority of another agent. In these cases, which role is one playing? Agent, patient, or a mix of both? The problem arises because, in these cases, one’s agency does not appear to be entirely bypassed but only severely constrained in its exercise.)

To begin articulating the notion of ‘agency,’ one might start by looking at some paradigmatic exercises of this capacity. A common suggestion is to look at intentional bodily movements, such as raising one’s arm or crossing the street. Here are some typical features of these cases: (a) the movements are goal-directed; (b) the goals cannot be immediately achieved, (c) one *directly* engages in these activities (i.e., one does not make these movements by doing something else that results in them), and (d) the movements are ‘productive,’ that is, they make a difference to the world (both in their unfolding—the rising of the arm as I am raising it—and when they reach their completion—the arm being raised once I have raised it). Let’s briefly consider these features in turn. What do they tell us about agency? Should they be considered paradigmatic of it? If not, why not?

3.2 Goal-directedness

A goal-directed movement has a certain ‘direction’ or ‘orientation’ toward the goal in question, which means that only some of the ways in which the agent might behave would make her attain the goal or at least move her closer to attaining it. When the agent is acting, she is behaving in ways that take her closer to reaching her goal (or at least that she believes would take her closer to it). Additionally, she is ready to attempt some compensations or adjustments if she encounters obstacles, interferences, or set-backs in her progression toward her goal. The pursuit of a goal seems to require some minimal persistence: the agent does not necessarily and immediately give up the pursuit in the face of any interference or perturbation. When she encounters some obstacles, she is normally expected to try to get around them. Relatedly, once the agent has attained the goal, she is supposed to stop the movement that was conducive to it.

The insistence on the goal-directed or purposive character of paradigmatic examples of agency seems correct. If there is an issue here, it is with the proper characterization of the expected adjustments and compensations, since we might want to exclude from purposiveness the simple behavior of physical systems that are naturally tending toward some equilibrium (and thus might seem to ‘persist’ in their behavior as if they were genuinely aiming toward that condition). This raises the question of whether goal-directedness could be characterized exclusively in ‘extrovert’ terms, that is, just in terms of patterns of observable external behavior, or it should appeal to some internal representation of the goal. Either way, some kind of purposiveness seems indeed a necessary and distinctive feature of agency.

3.3 Agency as a productive power

Let’s now consider the suggestion that agency is paradigmatically *productive*, that is, that it is a matter of interventions that make a difference in the world. This suggestion has some intuitive appeal but it seems to exclude witting or intentional omissions and refraining—an exclusion that might seem unwarranted. (See ‘Negative agency,’ Chapter 4.) Regardless of how this question is ultimately resolved, notice that even in standard exercises of agency of the productive sort, there might often be moments of ‘negative’ agency, where an agent takes advantage, at least temporarily, of already favorable conditions. That is, an agent might just let the events unfold naturally without any intervention for as long as this unfolding is getting her closer to her goal. One might just go with the flow, roll downhill, or let the wind blow in one’s sails. At these times, agency is not exercised by antagonistic and effortful interventions (including bodily movements) but simply by monitoring and supervising the unfolding of the events, while being able and ready to intervene in case the natural course of events might need to be nudged or corrected. So it seems plausible that agential guidance and control need not necessarily (and maybe not even paradigmatically) take the form of continuous, sustained, antagonistic, and productive interventions. At times, all that one must do to exercise one’s agency is to coast or drift.

Relatedly, one should be careful about picturing the exercise of agency as being always or paradigmatically about ‘initiating’ some motion or change, as if acting were modeled on such scenarios as getting our inert and slouching bodies off the couch. By contrast, many (if not most) exercises of agency seem to take place against the background of the directionless restlessness of the world (a restlessness that can also take place inside our bodies). The basic job of agency might thus be better characterized as giving some direction and shape to these underlying motions—something that can be done in a variety of forms, from pugnacious antagonistic interventions to hands-off supervision.

In addition, an agent might often be aiming at the *maintenance* of some state or condition, including one's own health and existence, rather than at the creation of some novel items or conditions. In such a case, the agent's busyness with antagonistic interventions is not constitutive of agency as such but it is only forced on the agent by the need to counter the (possibly constant) perturbations to her desired stasis or equilibrium.

3.4 *Telic and atelic goals*

The distinction between production and maintenance also relates to the paradigmatic structure of the agent's goals. A tempting picture of standard action is one in which we reach for or move toward some object, condition, or state of affairs that is at some spatial and temporal distance from us—so that there is necessarily an interval in which one is pursuing a distal goal which has not yet been achieved. In these cases, the action concludes only when the goal has been reached. However, there are many pursuits in which the end is achieved *at the same time as* one successfully engages in them. For instance, when one is leisurely strolling—walking 'aimlessly'—one succeeds at strolling at the very same time as one is taking the appropriate steps. As one strolls one has thereby strolled. And this success does not necessarily call for the termination of that activity, though it might propel its continuation.

Some maintenance activities might have a similar temporal character. The agent might make repeated and possibly continuous adjustments that make her succeed at maintaining a given condition right at the time when she makes the adjustment. But her stasis or equilibrium might be immediately and continually perturbed, hence calling for continuous and potentially indefinite maintenance (the activity of keeping oneself biological alive or healthy, for instance, might have this temporal profile, which is very different from the profile of raising one's arm or crossing the street). (See 'Agency, events, and processes' Chapter 3, and 'Diachronic agency' Chapter 30.)

3.5 *Agency and embodiment*

Let's now consider the role of the body. Does all agency necessarily or paradigmatically involve the body? As physical beings, it is unsurprising that much of what we do involves our bodies. But this should not entail that there couldn't be any mental agency. (See 'Mental agency,' Chapter 16.) A lot of our mental life, after all, appears to be active in recognizable ways: there are times where we form beliefs and intentions, engage in reasoning, direct our attention to certain matters, and so forth. These mental episodes seem to be correctly described as intentional, deliberate, or voluntary, and we might be held accountable and responsible for their occurrences and consequences (although it is still an open question whether there is a distinctive kind of *epistemic* agency, see 'Epistemic agency,' Chapter 20.)

In addition, an important part of our mental and affective life is a matter of our being (defeasibly) responsive to reasons. It could be plausibly argued that this responsiveness is also a manifestation of agency (maybe even the primary manifestation of our distinctive kind of agency). If that is correct, the 'at will' character of voluntary bodily movements might offer a misleading model of reason-responsiveness, with its insistence on discretionary choice rather than on the conformity of our conduct to reasons. (See 'Rational agency,' Chapter 10, and 'Aesthetic agency,' Chapter 42.)

Regardless of where one stands on the issue of mental agency, there are still questions about the role of the body in paradigmatic modes of agency. In particular, one might wonder whether agency might not be *directly* realized and exercised in other physical but

non-biological forms, such as various technologies, and what implications this might have for the individuation of the realizations of agency, including its social and political dimensions. (See ‘Material agency,’ Chapter 19.)

3.6 Agency and intentionality

The last element of the allegedly paradigmatic model of action is the ‘intentional’ qualification of productive bodily movements. The term ‘intentional’ can be read in at least two ways. According to a weak reading, it just stands for a generic kind of purposiveness or goal-directedness (possibly the one that requires at least a mental representation of the goal). As such, it can be applied to a variety of agents, including nonhuman animals. According to a stronger reading, it refers to the much more demanding form of agency that is characteristic of human beings—a form which includes such things as the understanding of goals as goals, a distinctive kind of self-knowledge about one’s goals and actions, planning abilities, and a willingness to offer and ask for reasons for action (for more on this, see the discussion of full-blooded agency in this introduction below).

Both uses of ‘intentional’ and cognate expressions are legitimate but, when using this term in the philosophical investigation of agency, one needs to make clear whether one is trying to characterize a notion of agency that has broader application across a variety of entities, including various organisms and possibly some artifacts, or one is primarily focused on the distinctive properties of human agency.

4 Kinds of agency

4.1 The varieties of agency

Unless one is skeptical about the existence of any genuine agency at all, there is a general philosophical agreement that normally developed human beings possess agency. But what about other animals, plants, and unicellular organisms? What about artifacts such as robots, or supernatural beings such as angels, demons, and gods? There are also important philosophical questions about the presence of genuine (and non-derivative) agency in sub-personal systems or functional components (say, the digestive system, the heart, a single cell in a multicellular organism) and in super-personal entities, such as social groups and institutions.

Part of what is at stake in addressing these questions is whether there are different levels or degrees of agency. It might be that some dimensions of agency could be attributed only to more complex organisms but not to simpler ones. If so, what are these dimensions? How are they related to each other? What are the normative implications of these attributions? (See Part 2, Chapters 7–10.)

Even when there is some agreement on the different levels or degrees of agency, there is still one additional methodological issue: what is the relation between the various kinds of agency? Are they set along an additive sequence, so that one can go from one level to another by simply adding some powers or capacities to the simpler forms (thereby retaining the basic operations of the lower/simpler levels)? Or are the various kinds different species of the same genus, so that possessing some different capacities has a ‘transformative’ rather than a merely ‘additive’ effect on the various capacities and powers? (See ‘Rational agency,’ Chapter 10.)

Relatedly, there is a question whether some of the more complex forms of agency might be taken to be *better* along some evaluative dimension, and whether agents might be under

a demand to acquire and exercise better forms of agency, when they have the option to do so. One might also wonder whether there are any normative pressures to relate differently toward simpler forms of agency. Concerns of this sort become particularly poignant in the case of human agency. Human agency exhibits various dimensions and layers of complexity, which are acquired as we grow up, and that we might lose or fail to develop and cultivate. How should we both conceptualize and relate to those instances of human agency that do not exhibit the standard combination of powers and capacities? (See Part 2, Chapters 7–10; ‘Agency and disability,’ Chapter 14; and ‘Pathologies of agency,’ Chapter 15.)

4.2 Full-blooded agency

Much of the philosophical reflection on agency is centered on human agency, which is often referred to as ‘full-blooded agency.’ This is also the term that I will use throughout this handbook to indicate that this agency exhibits a distinctive combination of features. In addition to purposiveness, full-blooded human agency exhibits the following features:

- 1 There is structural complexity both in our pursuits and in our ends. Much of this complexity stems from the extended temporal character of our agency in the mode of planning—we have psychological, conceptual, and reasoning capacities to engage in integrated long-term plans, often directed at ends that we could not pursue and conceive in the absence of these very capacities (see ‘Planning agency,’ Chapter 31, and ‘Agency, time, and rationality,’ Chapter 32).
- 2 We occupy a privileged epistemic position with respect to our own intentional conduct: we have a distinctive kind of self-knowledge about what we are doing and why we are doing it (a kind of self-knowledge that might extend to or explain other forms of self-knowledge). (See ‘Agency and practical knowledge,’ Chapter 21, and ‘Agency and self-knowledge,’ Chapter 23.)
- 3 In explaining, assessing, and interacting with the intentional conduct of fellow human beings, we expect them to have *reasons* that justify and make sense of their conduct (reasons that are in principle expected to be explicitly offered by the agent when one is asked for them). In this way, we make explicit what a constitutive feature of intentional conduct—being the object of a distinctive kind of explanation: explanation in terms of reasons, rationality, and intelligibility. The offering and the asking for reasons, moreover, is not an idle kind of external appraisal. We have a basic disposition to modify and adjust our behavior in light of our appreciation of the reasons that purport to justify it. Hence, we operate with the expectation that our conduct conforms to the reasons that purportedly justify it. The same disposition and expectation operate in response to the explicit requests of reasons. To the extent that our conduct does not conform to these reasons, we are expected to modify it (if still possible) or to be held accountable for the failure to do so. (See ‘Agency, reasons and rationality,’ Chapter 36, and ‘Agency and practical reasoning,’ Chapter 37.)
- 4 The accountability to rational justification is at the basis of what might be called ‘ethical responsibility.’ We hold each other responsible for our intentional conduct, its consequences, and its results (something that takes place in various settings, including interpersonal relationships, institutional and legal contexts, and morality). It is because of its intimate relation to the practice of holding agents responsible that full-blooded agency is the natural home for the application of such qualifications and concepts as voluntariness, willingness, duress, compulsion, coercion, reluctance, consent, innocence, and guilt. (See ‘Agency and responsibility,’ Chapter 25.)

- 5 Full-blooded agency is the ground of the possibility of attributing normative or deontic statuses: only full-blooded agents can be regarded as genuine and full bearers of duties, obligations, rights, and responsibilities. Full-blooded agency is thus a precondition for both legal and moral personhood. (See ‘Agency and personal identity,’ Chapter 34.)
- 6 Traditional questions about freedom and free will seem to be appropriate only once full-blooded agency is on the scene. (See ‘Agency, will, and freedom,’ Chapter 24.)
- 7 Questions about identification, self-governance, authenticity, and autonomy only arise in the presence of full-blooded agency. (See ‘Agency and identification,’ Chapter 26, and ‘Agency and autonomy,’ Chapter 27.)
- 8 Full-blooded agency comes in a distinctive social form. This sociality is not limited to the capacity to engage in social activities, but it is also a constitutive aspect of the practices of explanation, justification, and accountability.
- 9 Finally, all these dimensions, features, and implications of full-blooded agency either go together with, or are partly made possible, constituted, or enhanced by our critical and reflective capacities. Minimally, we have the linguistic, conceptual, and reasoning resources to articulate our goals and the reasons in their support, and to understand the very notion of goals and reasons. This is our basic critical capacity: we can represent and understand the basic rational structure of our intentional agency and its normative implications. An important consequence of this critical capacity is that rational scrutiny and justification are not limited to our instrumental conduct but also extend to our ends. Higher-order reflective abilities can then be turned toward all features and implications of our full-blooded agency, in its psychological, conceptual, rational, normative, and moral dimensions—as evinced by the very existence of the *philosophy of agency*. (See Part 8, ‘Agency, reasoning, and normativity,’ Chapters 36–42.)

Much of the work in the philosophy of agency is centered on full-blooded agency, given that we are supposed to be agents of this kind. Unsurprisingly, many of the chapters in this volume have the same focus. The appearance of the word ‘agency’ in them (including in some of their titles) might often be better interpreted as ‘full-blooded agency,’ which is something that should be apparent to the attentive reader.

5 Four pictures of agency

This volume does not try to advocate for a specific positive account of agency. Its aim is rather to introduce the reader to several central issues and some different approaches in the philosophy of agency. The reader is invited to explore various aspects of agency, their connections, and their implications. There is no privileged place of entry or re-entry, although several chapters are loosely organized around some larger set of questions, as described in the introduction to the sections of this handbook.

To provide some initial direction, however, I’d like to offer four general *pictures* of agency that can seem to inform, often implicitly, various specific philosophical views about agency. By ‘picture’ I mean an approach to the philosophy of agency that tends to provide some basic orientation in selecting the questions that are considered worth asking and in setting the desiderata and constraints on their answers, including judgments about salience and irrelevance, the style of argumentations, and the examples that tend to be used (including the choice of paradigmatic illustrations of agency). This orientation might not be self-conscious or explicit, and it might originate in a pre-theoretical view on the nature of agency, which

might then be elevated into a more systematic philosophical account. A picture is not the same as a full account or theory that aspires to consistency and coherence. Rather, a picture tends to motivate certain lines of inquiry and to frame the questions to be raised. It might not exclude other approaches. One might find strands of different pictures within more elaborate philosophical accounts of agency, some of which might actually and explicitly try to reconcile different pictures. This should not be surprising. After all, each of the pictures that I am about to present appears to highlight some intuitively plausible aspects or dimensions of agency. Thus, as we start to reflect on the nature of agency, we should try to become aware of the role that any of these pictures might implicitly play in framing our investigation. And we should not be too cavalier in dismissing (or privileging) any of them.

Given my understanding of what a ‘picture’ of agency is, it should not be surprising that what I am about to offer are only some impressionistic sketches, where I hint at some variations on what I take the distinctive theme of each of these pictures (which also means that it is unlikely that one could find a philosopher that fits neatly into a single picture and that exhibits all of the traits that I am about to present). Notice that I am not claiming that these are the only four possible pictures of agency, but this is my best attempt to begin to articulate some of the approaches that can be encountered in the philosophy of agency literature.

5.1 Agency as creation

According to this picture, agency is the capacity to create or produce, to bring about something new, such as the initiation of an action. A down-to-earth example of this creative power is the ‘at will’ raising of one’s arm from a position of rest, which is prompted by nothing other than one’s choice or decision to do so. A similar example is the case of the arbitrary selection between two or more open paths, especially in the case in which they appear to be equally desirable (think about Buridan’s ass scenario, for instance).

This picture emphasizes the role of the agent as the *source* or *origin* of action, where the action is added as something *new* to the world. This is why I call this approach ‘agency as creation’ (rather than agency as the mere power of initiation or selection). I also suspect that, for some proponents of this picture, the ideal or model of agential power might be something like a divine ‘fiat’—a divine *ex-nihilo* creative act. This is not to say that raising one’s arm or selecting from among open paths is without constraints. But within those restrictions, for this picture, agency operates unfettered, hence its ‘discretionary,’ ‘at will,’ or ‘arbitrary’ character. According to this picture, exercises of agency are ultimately manifestations of a radical kind of freedom (possibly as radical as the ‘liberty of indifference,’ which is unfettered even by the constraints imposed by responsiveness to reasons and thus open in principle to the possibility of perverse action). The insistence in this picture on the (locally) unconstrained creative or productive powers of the agent might ultimately reflect and expose a deep concern with our possible ‘captivity,’ with the worry that we might ultimately lack freedom.

Relatedly, in this picture, the attribution of some conduct to the agent is primarily a matter of *tracing the source* of that conduct back to the agent as the place where one cannot go any further. The investigation of the internal structure of the agent appears to be a less urgent task, possibly because of the expectation that this investigation could not reveal much. This creative power might thus be a distinctive but *irreducible* feature of agency. This irreducibility, however, raises the worry that this picture might be difficult to reconcile with any naturalistic view of the world—especially if one thinks of the basic exercises of agency as ‘inserting’ themselves *ex nihilo*, so to say, into the natural causal order.

5.2 Agency as self-constitution

According to the second picture, agency is ultimately a matter of self-constitution or self-maintenance. In this picture, the paradigmatic example of the exercise of agency is the conduct of an organism, which is ultimately directed at the organism's self-maintenance, that is, at securing its continuous survival in response to the ultimate existential threat: that of dissolution and death. For this picture, agency is ultimately the same as the capacity of life: agents are first of all organisms (self-constituting and self-maintaining entities), and different kinds of agency reflect different kinds of life-form.

The sense of life (and organism) in this picture need not be restricted to the 'biological,' to material organisms with a metabolism. In principle, it seems possible to extend the idea of self-constitution and self-maintenance to *rational* life, to the life of a rational subject *as a* rational subject, where the existential threat arises within the rational order rather than within the causal one. Inconsistency and incoherence might be to rational life what material disintegration is to physical life. In both cases, the threat is to the maintenance of the kind of integrity and unity required to sustain continuous existence either as a rational or a biological agent (or both). (Notice that this description of rational life concerns the characterization of the kind of unity and existence at stake in the power of agency—it is a separate issue whether a rational life-form could exist in a disembodied form or as embodied in a non-biological form, say in a material artifact.)

This picture of agency gives expression to the idea of agency as self-motion, as self-originating conduct with a necessary teleological orientation. Within this picture, there is no question about the source of the agent's basic *business*, since self-maintenance comes with the built-in preoccupation to fend off any potential fatal threats—which the agent might constantly encounter in its path to survival. Unlike agency-as-creation, the basic concern here is not captivity, but mortality.

Agency-as-self-constitution offers a straightforward account of the attributability of the action to the agent. As an exercise of self-motion, any piece of conduct is necessarily of the 'self,' that is, of the entity that is, by its nature, in the business of constituting itself—of making its 'self.' This invites an investigation in the conditions that make possible for the entity to have both adequate internal integrity and sufficient separation from the external world to constitute a distinct and viable *unit* of agency. Likewise, in this account, there is a straightforward path to the attribution of agential conduct to the agent 'as a whole.' Agency is to be primarily attributed to the self-maintaining entity rather than to any of its functional components, such as an organ or a subsystem. The operations of these components might have their own teleological orientation, but if these components are not self-sustaining in relative isolation from the larger entity, they might be said to have agency only in a derivative sense, if at all. Likewise, larger aggregates of self-maintaining agents would not exhibit genuine agency unless they reach a level of integration such that, at that level of organization, they would count as self-maintaining units in their own right. (This is not to say that self-maintenance is maintenance in isolation from the external world or from other agents. The picture of agency-as-self-maintenance can actually be well-positioned to account for the mutual interdependence of agents, including various forms and levels of sociality and the nesting of different levels of agents within each other.)

In agency-as-creation, the creative power is primarily manifested in each individual exercise of agency, as directed at bringing into existence the particular object of that individual act of creation. In agency-as-self-constitution, agency still has a creative aspect, but what is ultimately (and constantly) brought into existence is the agent itself. The agency of self-constitution is the agency of continuous self-(re)creation.

This does not mean that self-constituting agents must explicitly represent their own maintenance as the object of any of their particular activities. Self-maintenance might often be performed by the explicit pursuit of instrumental means, without necessarily representing or appreciating them as means to self-maintenance. An organism might never seek self-maintenance as the direct object of any of its goal-directed pursuits, whose proximal object is always some specific (and usually telic) way of contributing to its survival, such as getting food, avoiding predators, and seeking shelter. It is only for a reflective self-maintaining agent, who becomes cognizant of the fundamental structure of their agency, that (atelic) self-maintenance might become the primary and explicit object of concern (for the telic/atelic distinction see above and ‘Diachronic agency,’ Chapter 30).

Here are some of the attractive features of this picture. First, it is easy for it to emphasize the temporal dimension of agency. This is because self-constitution is a continuous process, which is terminated only by the agent’s death. The continuity of self-constitution should not be confused, however, with a relentless process of positive intervention in the world. There is nothing in the idea of a self-maintaining entity that stands opposed to taking advantage of already favorable conditions by coasting and drifting. If anything, the negative mode of agency—letting things unfold naturally—could be seen as the preferred mode of self-maintenance, if only one were not under constant threats. After all, for this picture of agency, positive interventions are ultimately only in response to present or anticipated threats. It is only the indifference or the hostility of the world that might force the self-maintaining agent into a relentless series of mostly positive interventions. In agency-as-self-constitution, there might be no special value attached to spontaneous or ‘unforced’ individual acts of positive creation or production as such.

Another interesting feature of this picture is that it might offer the basis for an account of the connection between agency and normativity—by leveraging either the built-in aim or function of self-maintaining entities or the force of the existential threat that always hangs over the heads of self-constituting entities. (See ‘Agency and normativity,’ Chapter 38, and ‘The aim of agency,’ Chapter 39.)

Finally, this picture might make it easier to account for the natural history of agency and the existence of different kinds of agents, since this history and variety might just reflect the unfolding of the tree of life on Earth. But for this same reason, this picture might run into problems when trying to account for the rational aspect of agency, especially in the form of self-conscious reason-responsiveness. Although agency-as-self-constitution might explicitly contemplate ‘rational’ life-forms, there is a legitimate worry about whether the notions of self-maintenance and self-constitution can really apply to rational agency, especially of the critical and reflective kind.

Agency as self-constitution seems especially apt at accounting for the agential character of the operations of simpler forms of life. The concern is that it might become much less plausible when applied to the distinctive features of full-blooded agency. In a related vein, one might wonder whether an excessive reliance on a picture of agency inspired by the structure of biological life might rule out the possibility of genuine agency in the absence of either metabolism (say, in robots or androids) or of any existential threats (say, in immortal beings).

5.3 Agency as psychological causality

According to a third picture, agency is fundamentally the *psychological* capacity to bring about bodily movements that intelligibly fit with the agent’s desires, cares, concerns, or commitments. A straightforward illustration of this picture is found in what is known as

the ‘standard story about action’ (see Introduction to Section I). According to this story, an action is a bodily movement caused (in the right kind of way) and rationalized (that is, made intelligible) by the agent’s desire for a certain end and her belief that moving her body in that particular way will bring about that end.

There is something intuitively appealing about the standard story, since it seems to conform to ordinary folk-psychological explanations of action, in terms of the so-called belief/desire psychology. For instance, one explains the action of reaching for a glass of water out of the combination of one’s desire to quench one’s thirst and one’s belief that one can satisfy that desire by extending one’s arm in the direction of the glass of water. This arm movement ‘makes sense’—it is made intelligible or ‘rationalized’—by that particular combination of mental states. (By contrast, it would not make sense to reach for the glass of water if, say, one believed that the glass was outside of one’s reach or one had no desire for the water.) In this account, the elements of one’s psychology play a double explanatory role: they make the action intelligible by providing reasons for its performance; and they bring about the action by causing it (in the right way—this qualification is always implicit because of the so-called problem of ‘deviant causal chains,’ on which see ‘Agency and causation,’ Chapter 1, and ‘Agency, functions, and teleology,’ Chapter 2).

The standard story is not the only version of what I am calling agency-as-psychological-causation. For instance, one could offer a much more elaborate characterization of the psychological structure (or ‘psychic economy,’ as this is sometimes called) that is explanatory of the action, including the existence of different kinds of mental states (including intentions, plan states, and higher-order mental states) and more complex ways in which they might relate to each other. (See ‘Intentional agency,’ Chapter 9, ‘Planning agency,’ Chapter 31, and ‘Agency and identification,’ Chapter 26.)

What makes these various accounts versions of the psychological-causation picture is their focus on an account of action in terms of its psychological (as rationalizing) causal antecedents. Unlike the pictures of agency-as-creation and agency-as-self-constitution, this picture makes the presence of a mind—of a psychic economy—central to the nature of agency; it does so by articulating the internal structure of the psychological structure and its contribution to bringing about genuine exercises of agency. As I remarked earlier, this articulation appears to be a problem for agency-as-creation, especially if that picture insists on the *sui generis* character of the agential powers. The agency-as-creation picture is at risk of locating agency outside of the natural causal order. A commitment to a naturalistic explanation appears to be a major motivation behind the agency-as-psychological-causation model, hence its insistence that both the internal operation and the external outputs of the psychology be accounted for in terms of the generic bond of ordinary causation. Psychological causation is ordinary causation *by* elements of one’s psychology, not some kind of supernatural, mysterious, or spooky power.

Agency-as-creation is not necessarily guilty of invoking some kind of magic, but it can lend itself to such invocation. For, unlike agency-as-psychological-causation, it does not start from the very idea that agency is a capacity to be explained in terms of the naturalistic operation of some underlying causal structure.

Notice that this is not a concern with agency-as-self-constitution. A self-maintaining entity is one with an internal structure and organization that is in principle naturalistically explainable. What differentiates agency-as-psychological-causation from agency-as-self-constitution is that the former starts from a *psychological* structure. This appears to be an advantage when one is to account for the distinctive features and implications of full-blooded agency, which is imbued with psychological and rational

attributes. The psychological and rational dimensions of agency, instead, are only a secondary feature of agency-as-self-constitution, given that self-maintenance is, first of all, an ontological rather than a psychological or rational property. By starting with life rather than with mind, agency-as-self-constitution might end up being too generous in the attribution of agency to simpler kinds of organisms while struggling to account for the distinctive rational dimension of full-blooded agency at the other end of the spectrum. A related advantage of agency-as-psychological-causation is that, in principle, it faces no difficulty accommodating the possibility of genuine agency in entities that are not self-maintaining—such as artificial, non-biological machines—as long as they are realizations of a sufficiently complex psychic economy.

Agency-as-self-constitution, however, might fare better than agency-as-psychological-causation when we come to the question of how to attribute actions to their agents. As we have seen above, in agency-as-self-constitution, the self-originating character of the exercises of agency is guaranteed by the very notion of self-constitution. The psychological causation picture, instead, might struggle on this front. According to psychological causation, where exactly is the capacity of agency located? Who or what is exercising this capacity? ‘The psychic economy!’—one might say. But what does make this economy that of a genuine *agent*—in the sense of a sufficiently integrated unit to count as the proper locus of the exercise of agency? The concern is that the elements of the psychic economy might be doing the work by themselves, without any discernible role for the agent (this is known as the problem of the ‘disappearing agent’). The risk is that the psychological-causation picture might either explain the agent away or surreptitiously and uninformatively assume some homuncular unity within the internal working of the psychic economy. True, the psychological causation story is framed in terms of the operation of a *system*—a psychic ‘economy’—but there seems to be nothing in the basic idea of psychological causation that, by itself, explains how to individuate a system with the kind of unity and integration that we expect of the idea of a genuine agent.

In addition to the worry of the disappearing agent, many objections have been raised against the standard story of action. These objections target its metaphysics of mind and of causation, its seeming inability to account for extended guidance rather than just the initiation of actions, and its troubles with the putative agential character of omissions and refraining. (See the introduction to Part I, ‘The metaphysics of agency,’ and Chapters 1–6.) However, the standard story is only one possible version, even if an intuitive one, of agency-as-psychological-causation. Its problems need not necessarily invalidate the picture in all of its possible forms. After all, the basic idea of psychological causation seems most welcoming to different characterizations of the psychic economy, including a variety of mental architectures, with very different components and structures (including different metaphysical accounts of causation and the appeal to psychological elements belonging to different metaphysical categories—such as states, events, and processes).

In my view, what makes these different accounts instances of the psychological-causation picture is not a commitment to the specific features of the standard story but the aspiration to provide a naturalistic account of how a psychology can operate causally and yet stay within the rational order, so as to give rise to full-blooded agency. This is what makes agency-as-psychological-causation especially attractive but also potentially most vulnerable. For there is always the worry that the causality invoked by this picture to explain the operations of the psychic economy might not be able to explain the truly distinctive aspects of full-blooded agency, including its normative character and its self-conscious and reflective dimensions.

5.4 Agency as reason responsiveness

According to this last general picture of agency, agency is primarily the capacity to respond to reasons. Agency is first of all exercised in making up our minds on the basis of normative or rational considerations about how we ought to make up our minds. In the paradigmatic cases, in making up our minds about what to do, we thereby also give shape to our conduct in the material world. Our agency, however, is not primarily exhibited in the ‘at will’ shaping of our physical actions but in our avowals or disavowals of reason-responsive attitudes.

It is useful to compare this picture to agency-as-creation. In agency-as-creation, agency is ultimately a matter of making a difference within the causal fabric of the world. This kind of difference can be modeled in terms of changes in the physical world, such as selecting which path to take at a junction or setting into motion an inert body. In agency-as-reason-responsiveness, instead, the change is first of all within a *normative* rather than a physical space. In agency-as-reason-responsiveness, the agent takes on a new shape because of the new status acquired by some of one’s own attitudes (paradigmatically, one’s judgments) in response to one’s sensitivity to normative considerations.

Crucial to both pictures is the idea of the agent as the direct *source* of one’s own agential conduct. But the two pictures drastically differ in the character of the proposed source. This difference can be illustrated by the distinction between ‘authorship’ and ‘authority.’ According to agency-as-creation, the agent is the source of agential conduct. It is so because the agent is the author—the creator or originator of this conduct. This authorship, which retains a voluntary character, is ultimately accounted for in causal terms.

By contrast, in agency-as-rational-responsiveness, the agent is a source in the sense of being the authority that endorses, judges, or avows something. Being the agent is, first of all, a matter of putting a stamp of approval (hence the change in normative status within the space of reasons) rather than of directing one’s conduct into some physical direction instead of another (even if one’s conduct is by default expected to take on a specific direction in physical space as a result of the change in one’s normative status).

The two pictures also differ in their interpretation of what they fear the most. Both are worried about the loss of freedom, especially in the form of being pushed around by mere causality. But agency-as-creation might respond to this concern by encouraging a radical resistance to any form of ‘constraining,’ even the one induced by responsiveness to reasons. Thus, agency-as-creation might end up encouraging an extreme view of agency as the defiance of reason, as epitomized by choices done in an arbitrary matter or in the absolute refusal to be shaped by any rational consideration.

The defiance of reason, instead, is contrary to the very spirit of agency-as-reason-responsiveness. The kind of freedom that the latter view wants to affirm is that of autonomy—of self-determination in response to one’s appreciation of the reasons for action, not that of arbitrary indifference (if not even defiance) of one’s assessment of which reasons bear on one’s practical circumstances.

What counts in favor of the agency-as-reason-responsiveness? First, this picture accounts for the seeming agential and active character of aspects of our lives that go beyond physical actions, including the vicissitudes not only of our cognitive attitudes but also of some of the conative and affective ones. This is shown by our practices of explaining, appraising, and holding agents responsible on account of the reason-responsive (rather than ‘at will’) formation, retention, and revision of these attitudes. Second, the view does not have problems with the attribution of a piece of conduct to its own agent. This is because, at least in the paradigmatic cases, the subject’s judgment that so-and-so is to be done *constitutes* the fact that

the doing of so-and-so is the doing *of* that subject. Third, this picture has no trouble accounting for the relation between agency, reasons, rationality, and, more broadly, the normative domain. This is because the connection between agency and normativity is built right into this picture of agency.

These same features, however, might also constitute the major vulnerability of agency-as-reason-responsiveness. First, by being centered on reason-responsiveness, this picture might be unable to account for the simpler kinds of agency. Second, it is unclear how this picture fits with a naturalistic outlook. It is not that this picture is in principle anti-naturalistic, but, unlike agency-as-psychological-causation, it does not seem to be bothered with securing foundations with impeccable naturalistic credentials (which is unsurprising since it primarily trades in the abstract space of the normative).

5.5 Conclusion

This concludes my initial and very rough sketch of four basic pictures of agency that might, at least implicitly, motivate and shape many philosophical investigations of agency. As I remarked at the outset, these are general pictures that might be neither fully explicit nor articulated by those who are under their influence. And it would not be surprising if many philosophers are attracted by more than one picture at the same time and possibly try to reconcile them, given that each of these pictures seems to highlight some distinctive and attractive features of our full-blooded agency. However, it is too early to say whether any such reconciliation is possible or whether one of these pictures (or possibly even a different one altogether) will eventually emerge as the correct one.

6 What's next?

6.1 'Philosophy of agency'

The questions and pictures that I have presented in this introduction are a sample of the philosophical issues raised by the notion of 'agency.' Many additional topics will be presented in the following chapters. Taken all together, they should give the reader a good (although far from comprehensive) overview of the 'Philosophy of Agency.' Notice that this label, although it is gaining prominence, has not yet replaced 'Philosophy of Action' as the standard name for this area of philosophy (whereas the old-fashioned and far too restrictive label, 'action theory,' seems to be, fortunately, almost forgotten).

Whenever it is in my power to do so (such as editing this handbook), I will insist on the use of 'philosophy of agency.' This is, I believe, for very good philosophical reasons. The primary topic of investigation is not 'actions,' which are just among the possible exercises of the capacity of agency. True, actions (together with acts and activities) appear to be the primary manifestations of agency, and, as such, they might correctly get the bulk of the attention. But this does not justify making actions, rather than the capacity that underlies them, the primary object of philosophical reflection. And even if one were to argue that agency can only be exercised in actions, this would still be a *substantive* view within the philosophy of agency, not a neutral definition of an area of philosophical inquiry.

Even those who might profess neutrality on this matter should be wary of the surreptitious effects of the use of the label 'philosophy of action.' This name might favor a potentially tendentious philosophical agenda, which could give undue prominence to actions over other manifestations of agency. Just to give an example about some hidden dangers: by starting the

investigation with talk of ‘actions’ one might privilege certain phenomena in their nominalized form rather than in a verbal one, including variations in tense and aspect. But who is to say that the focus should be primarily on actions rather than on acting, say? (Likewise, why should one privilege deeds over doings? Intentions over intending? Plans over planning?) In addition, the primary focus on actions might implicitly suggest that they could stand as separate items, which could be studied piecemeal, in relative isolation from each other and from their sources, and used as building blocks in the reconstruction of extended exercises of agency. Last but not least, the primary focus on the exercises of a capacity also risks obscuring the role played by agents, as the loci of this capacity. The philosophy of agency should be, at the very least, a philosophy of agency, agents, and actions. These are not idle worries, since important and foundational philosophical matters might lie behind these seemingly innocent linguistic choices, especially in the case of the philosophy of agency. (In further support for my terminological choice, consider how odd—and philosophical prejudicial—it would sound to refer to other philosophical areas by the analog of ‘philosophy of action’: for example, ‘philosophy of belief’ or ‘belief theory’ instead of epistemology; ‘philosophy of mental states,’ or ‘propositional attitudes theory’ instead of ‘philosophy of mind,’ ‘philosophy of moral conduct’ instead of ‘ethics’ or ‘moral philosophy,’ etc.)

6.2 *How to read this book*

This handbook is not going to propose any positive characterization of the nature of agency but rather to represent some of the variety of issues and viewpoints to be found in the philosophy of agency (not to mention that many of the central questions in this area revolve quite directly with addressing the very question of the nature of agency). This is also why there is no specific chapter in this volume titled ‘The Nature of Agency.’ The reader is rather invited to try to discover and reflect on the various questions by sampling different topics in this handbook, with no specific point of entry or progression. At the end of each chapter, the reader will find some additional suggestions for further related topics, which one might follow up at their discretion. Some chapters are organized around a set of common questions, and the reader is invited to read the introduction to these sections to learn more about how the topics might be related. This organization, however, is not a rigid one. Many chapters could have as easily been assigned to different sections, and different groupings of questions might have been proposed altogether.

The volume aims to be comprehensive but it is by no means exhaustive. Many topics are not covered here. The absence of any particular topic is hardly an indication of its lesser philosophical importance. Many are missing simply because this volume is the product of the coordination of the limited and constrained agency of the editor, the authors, and other potential contributors. (Consider, for instance, some of the topics that went through some advanced stages of planning but had to be dropped at some point: ‘Agency, desire, and motivation,’ ‘Agency and the first person,’ ‘The logic of agency,’ ‘The sense of agency,’ ‘Agency and the virtues,’ ‘Agency and perception,’ just to name a few.) In addition, there are for sure many unwitting omissions that just reflect the ignorance and imperfection of this editor.

But there are a couple of noticeable intentional omissions that require some justification. First, except for the brief discussion in ‘Agency, will, and freedom,’ Chapter 24 (and indirectly in ‘Agency and autonomy,’ Chapter 27, and ‘Agency and responsibility,’ Chapter 25), there are no chapters that deal directly with the question of freedom and free will, despite the centrality of these topics to the philosophical reflection on agency. This ultimately reflects a major division of philosophical labor that is now well-established within contemporary

Western Anglophone philosophy. There is an entire philosophical area specifically devoted to questions of freedom, free will, and responsibility, and several introductory and more advanced resources are already available (see, for instance, Griffith, M., Levy, N., & Timpe, K. (eds.), 2017, *The Routledge Companion to Free Will*). I have thus decided that there was no point in duplicating in this volume what is already available elsewhere, and I invite the readers specifically interested in issues about freedom and free will to consult these other resources.

Likewise, there is no chapter devoted to shared, joint, collective, or social agency. This is not because these are not important aspects of agency. A reasonable claim can be made that agency, especially full-blooded agency, must necessarily come in some kind of social form. Some of these social aspects get explicitly discussed in some chapters (such as ‘Material agency,’ Chapter 19, ‘Planning agency,’ Chapter 31, ‘Agency and the emotions,’ Chapter 29, and ‘Agency and responsibility,’ Chapter 25). But a single introductory chapter on social agency would hardly add any new information which couldn’t already be available in such publications as Jankovic, Marija, and Kirk Ludwig (eds.), 2017, *The Routledge Handbook of Collective Intentionality* or the relevant entries in the *Stanford Encyclopedia of Philosophy* (<https://plato.stanford.edu>). As for the case of freedom, readers specifically interested in collective or social agency should consult these resources.

Many readers of this handbook might also benefit from reading it in conjunction with a systematic introduction to the philosophy of agency, such as the excellent volume by Sarah Paul, (2021) *The Philosophy of Action*, Routledge (see also Lilian O’Brien, 2015, *Philosophy of Action*, Palgrave and Rowland Stout, 2005, *Action*, Routledge).

Finally, this handbook does not specifically discuss any of the long history of the philosophy of agency, although many of the topics have been the object of philosophical reflection and discussion for centuries. This is, once again, a major editorial choice forced by practical constraints in putting a volume of this kind together. It does not reflect any negative assessment of the philosophical value of a historical look at this discipline. On a related note, this handbook does not want to deny or hide its historical, geographical, and cultural situatedness. The general philosophical outlook presented in this volume is the one informed by the scholarly discussion in Western Anglophone philosophy from about the 1950s, with some occasional influences from the so-called ‘continental’ tradition and some glances at selected portions of the history of Western philosophy. Much of the philosophical discussion that has taken place within this tradition in the last seven decades or so is done with the aspiration that the questions that are raised and the answers that are offered might, at least in part, transcend the confines of the particular philosophical outlook within which they have been raised and offered. Whether this aspiration succeeds (or is ultimately in vain), however, is not something that could be addressed in this volume. My hope, however, is that much more work will be done to investigate the nature of agency within other philosophical traditions and other disciplines. Although I believe that any reader should feel confident about beginning the study of the philosophy of agency by reading this handbook, the reader should never be under the impression that the philosophy of agency either begins or ends with the contributions collected in this volume.¹

Note

- 1 This handbook is the product of the joint, telic, and extended intentional activity of many agents. Whether or not they want to admit it, they all bear some responsibility for it (with the possible exception of their unwitting omissions). But I suspect that I am the ‘most agent-of’ this final product and so I will take most of the responsibility for it. The least burdensome and most welcome aspect of this responsibility is my duty (and desire) to thank all who contributed to the final attainment

by playing very different roles at different places and times. To begin with, I want to thank the editorial staff at Routledge: Rebecca Shillabeer—who first commissioned the handbook—, Gabrielle Coakeley, and Adam Johnson—who has overseen its further development and completion. This book could not have been conceivable without the existence of a thriving community of philosophers, who continue to engage in exciting conversations and reflections on the nature of agency. A fair number of them have directly participated in this volume, so I thank them all, both individually and jointly, for their chapters (and for their patience and promptness in dealing with my slow and desultory work toward the final success and closure of our joint pursuit). I also want to thank the many friends and colleagues who have not written a chapter for the handbook but have provided essential and invaluable support, advice, and above all philosophical inspiration over the years. In particular, John Fischer, David Horst, Richard Moran, Michael Nelson, Alexandra Newton, Sarah Paul, Andrews Reath, Karl Schafer, Tamar Schapiro, Sergio Tenenbaum. I also thank my colleagues and students at the University of Wisconsin-Milwaukee and, more recently, at the University of California-Riverside. A special thank goes to the organizers and participants of the many meetings of the DFG Netzwerk ‘Practisches Denken und Gutes Handeln’ and its descendant (aka ‘The Action Network’) who have done so much over the years to keep the philosophical conversation about agency both alive and lively. Last but not least, I have to thank the members of my small extended family—Marilena, Carla, and Susan—who have paid many of the direct costs of this project, by being around me during the ups and downs of the writing and editing process, but are only going to gain very indirect benefits from its eventual success: as much as it might surprise the readers of this handbook, not every full-blooded agent turns out to be also a philosopher of agency—something which, after all, we should all be thankful for.