THE SHAPE OF AGENCY

control, action, skill, knowledge

JOSHUA SHEPHERD

OXFORD

The Shape of Agency

The Shape of Agency

Control, Action, Skill, Knowledge

JOSHUA SHEPHERD

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

First Edition published in 2021

Impression: 1

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data

Data available

Library of Congress Control Number: 2020937525

ISBN 978–0–19–886641–1

DOI: 10.1093/oso/9780198866411.001.0001

Printed and bound in Great Britain by
Clays Ltd, Elcograf S.p.A.

This book is dedicated to Kelly, Zooey, and Finn.

Table of Contents

Acknowledgements

I had a lot of help in writing this book. For comments, conversation, and inspiration along the way, I want to thank Al Mele, Myrto Mylopoulos, Ellen Fridland, Matt Parrott, Denis Beuhler, Will Davies, Carissa Véliz, Uriah Kriegel, Tim Bayne, Nick Shea, Wayne Wu, and Elisabeth Pacherie. Many thanks to the students in my seminar at Carleton in 2018 for reading an earlier version of this. For listening to earlier versions of this, and making it better, I want to thank Neil Roughley and his group at Duisberg-Essen, people at the summer mind workshop at Columbia, including John Morrison, Katia Samoilova, and Antonia Peacocke, Thor Grünbaum and many at the University of Copenhagen, Chiara Brozzo, Hong Yu Wong, and many at the University of Tübingen, the whole workshop in the mountains crew—Balaguer, Buckareff, Downes, Grzankowski, Jacobson, Pasnau, Roskies, Strevens, and even McKenna—for conversations and encouragement regarding an early version of chapters 7 and 8, Felipe de Brigard, Santiago Amaya, Manuel Vargas, Walter Sinnott-Armstrong, and the audience at that Duke workshop, my colleagues at Carleton and also my colleagues at Universität de Barcelona, so many philosophers in the United Kingdom and Ireland, so many philosophers in Oxford and at the Uehiro Centre, and also the muses atop the Clarendon building.

For providing funding at various stages of this thing's development, I want to thank the European Research Council (Horizon 2020 grant 757698, for the project Rethinking Conscious Agency), and the Canadian Institute for Advanced Research's (CIFAR) program in Mind, Brain, and Consciousness, and the CIFAR Azrieli Global Scholar program.

For providing space to write, and music, and drinks of various sorts, I want to thank Doña Rosa in Barcelona, the many pubs of Oxford, and The Third in Ottawa.

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.

2
Control's Possession

2.1 Introduction

Anything that displays activity, that is a doing of something rather than a mere happening, will involve the exercise of control.[1]

A great sushi chef prepares a plate of food. One sees a series of precise movements. The chef makes accurate, contextually appropriate cuts into a piece of fish. They roll just the right amount of rice into the right shape. They pinch and then spread a finely chosen amount of wasabi. A great sushi chef can do this again and again and again and again.

As she practices, a professional tennis player takes balls and places them exactly where she wants across the court. Over and over she hits a small window on the court. Or, on the run, she lays down a volley with a delicate touch that slows the speed of the ball and that cuts a difficult angle. In the middle of a match this can seem completely improvised. But most successful shots are the product of years of practice.

It is said that if you want to make a certain percentage of three-point shots in basketball—in games, 40 percent is a great number—you should make double that in practice. That means day after day, in a gym alone, you should make eighty out of 100 shots from 24 feet out. And you should definitely shoot more than 100 per day if you want to get numbers that consistent.

The moral applies to any number of action-types—woodworking, painting, playing an instrument, dancing. What practitioners of these activities hone through practice, and what they display, is in part a tremendous amount of control over behavior. They can direct, or guide, their bodies and minds at a fineness of grain in order to consistently produce movements and thoughts that fall within a small space of variability. They do so because they can, and because they want to, and intend to, because these activities

[1] The ideas in the next three chapters are expansions upon, and in many cases refinements of, ideas first floated in Shepherd (2014a). I'm happier with the present versions of those ideas.

The Shape of Agency: Control, Action, Skill, Knowledge. Joshua Shepherd, Oxford University Press (2021).
 DOI: 10.1093/oso/9780198866411.003.0002

set out standards for success that they have learned to achieve again and again and again and again.

I have said that control is necessary for activity—for action, for agency. I have not said that only agents possess control. Engineers and biologists find the language of control useful, and their usage is similar to my own. A system or sub-system with control is a system or sub-system whose behavior can be modeled in terms of goal-states and processes that drive the system or sub-system into states that constitute or cause the satisfaction of these goal-states (cf. Dennett 1984). Such a system may not qualify as an agent.

The trick, for an engineer or biologist, is to understand the joints and levers of the system—to understand how the control is exercised. The trick for the philosopher of mind or agency is to elucidate the philosophically interesting components of controlled behavior and their relations set in the broader context of philosophical reflection on the nature of agents. I want to know what it is for the agent, as opposed to some non-agential system, or some sub-system within the agent—her early visual system, or her circulatory system, or whatever—to exercise control.

2.2 Control's Exercise

When an agent exercises control, they deploy behavior in the service of a certain class of mental states. The class I have in mind may be as narrow as the class of intentions. Or it may be broad. Perhaps desires, urges, various emotional states, states with imperatival content (arguably: pain states), or even certain perceptual states could qualify. That will depend on one's account of the contents and functions of such states. Perhaps packages of states could together qualify. Some think, for example, that an intention is really just a package of a desire and a certain kind of belief (Davis 1984; Sinhababu 2017). Or we could think of control with respect to a package of intentions, sub-intentions, associated beliefs, and so on. I'm neutral on all this.

My requirements: in order to be served by controlled behavior, a mental state or package of states should (a) represent (series of) events, states of affairs, or whatever, as to be done, or eventuated (that is, should set out a goal) (b) play a causal role in the production of (or, at minimum, attempts to produce) the thing to be done (that is, M should move the agent towards the goal) (c) qualify as a state or states of the agent, as opposed to some sub-system of the agent. Notice: (b) requires that the state move the

agent in the right direction, towards the goal. I require that this not be accidental. The state (or package) that moves the agent towards the goal should, then, do so at least in part because the state's content sets out a way to proceed towards the goal.

The third requirement is shifty, invoking as it seems to either a distinction between personal and sub-personal levels (Dennett 1969), or something like a distinction between doxastic and sub-doxastic states (Stich 1978). I rely on an intuitive understanding of states of the agent for now. (More detailed psychological architectures for particular agents would bring into play more detailed criteria for marking the distinction.) On the intuitive understanding, the agent's intentions, beliefs, fears, and so on are at the level of the agent. But states of the agent's early visual system, or states that regulate the agent's updating of long-term memory—states like these do not qualify. So, while the processing in early visual cortex is plausibly controlled processing, it does not qualify as control that the agent exercises. I discuss this issue further at chapter 5.5.3.

I need a term of convenience to refer to the relevant class of mental states. For reasons that will become apparent, I will call them plan-states. When agents deploy behavior in service of plan-states, they aim at success. Success involves a match between behavior and (certain aspects of) the representational content of the plan-state. That is what it is for behavior to be in service of a plan-state. Such behavior is directed towards matching aspects of the content of such a state.

A basketball player intends to make a shot. We can stipulate that the representational content of the intention includes the following plan: "square to the basket, locate the rim, aim just beyond the front of the rim, follow through, make the shot." When the agent is successful, she executes her intention as planned—she squares to the basket, locates the rim, aims just beyond the front of the rim, follows through, and makes the shot.

Talk of a content match between (aspects of) a plan-state and behavior raises the following question: what is the representational content of a plan-state, and what aspects of it are relevant? I will call the relevant aspects means-end aspects. I'm building on Myles Brand's work on the content of intentions, on which the content of an intention is a plan. Here is how Brand introduces the idea:

> An intentional action can be a momentous occasion in one's life, such as marrying, or it can be a mundane occurrence, such as showering in the morning; but in all cases, the agent is following his plan when acting. He

> has before his mind, as it were, a pattern of activity to which he brings his actions into conformity. (Brand 1986: 213)

As Brand notes, it is not immediately obvious what plans are. We sometimes talk of them in ways that suggest that plans are psychological states. But we sometimes talk of them in ways that suggest plans are abstract objects (e.g., "There is a plan for world peace but no one has thought of it and no one will" (218)). So perhaps plan-types are abstract objects and agents, via psychological states, token some of these types. Settling the ontology of plans here is not necessary.

We need some sense of the structure plans take, as well as of the kinds of content plans can embed.

Regarding the structure plans take, Brand notes that very simple plans need be little more than a linearly ordered series of steps. Plausibly a plan could involve only one step: move left, wiggle finger, scream, or whatever. But complex plans might involve conditional structures, sets of sub-goals, embedded contingency plans specifying what results count as second or third best, and so on. Brand (1986: 219) suggests we model plans as ordered triples, like so:

$$\text{Phi} = \langle A, h, g \rangle$$

Here A is a set of action-types (although one could say behavior-types instead, with action-types as a subset of these), h is a function on A that orders its members in terms of dependency relationships, and g specifies which results (events, states of affairs, or whatever) are the actual goals or subgoals embedded in the plan. Brand argues that g is necessary to capture plan structure because two agents could share A and h while having different goals. "You and I might follow the same recipe in baking a cake, yet act on different plans. The goal of your plan might be to produce a finished cake, whereas the goal of my plan might be to test the recipe; nevertheless, we both performed the same types of actions in the same order" (219).

I agree with Brand that a specification of the goal or goals is important. This specification sets a standard for success. I might be aiming to get the taste of butter to mix with the cocoa just right. You might be looking to give something to your dog for its birthday. We might produce very similar cakes, with my effort largely a failure, and yours a smashing success. The difference is in the goals. Perhaps my performance suffered. Or perhaps I

had a bad plan—that is, perhaps the behaviors in A or the dependency relationships specified by h were poorly chosen or poorly constructed. This suggests a distinction between the quality of performance of the behavior-types a plan specifies, and the quality of satisfaction of the goal(s) a plan specifies.[2]

As important as goals are the dependency relationships between behavior-types. A plan for shaving involves the application of shaving cream and the passing of a razor over the skin. It is important that one happen before the other. Other plans could embed contingency structures, back-up strategies, specification of next-best outcomes, and so on. Many behavior-types could be represented by a plan, with some to be performed only on certain branches of a tree, and only in certain orders, owing to the way the plan orders the importance of the goals to be achieved.

The dependency relationships in a plan thus set up a means-end structure as internal to, constitutive of, the nature of plans. Behaviors are indexed to goals as means, and they are weighted against other possible behaviors given various contingencies.

My claim is that controlled behavior is behavior that, with additional constraints added below, approximates means-end aspects of a plan-state. Brand speaks of the agent bringing actions into conformity with intentions. Lilian O'Brien (2012) speaks of a matching between movements and contents of intentions. The idea here is similar. The aspects at issue are the means an agent takes, the ends in view, and the dependency relationships between various means and various ends. One could pull these apart and speak of an agent's fluency at various aspects in isolation:

Agent A performs the behaviors perfectly, but gets them out of order, and fails to achieve the end.

Agent B performs the behaviors poorly, but gets the order right, and fails to achieve the end.

Agent C performs the behaviors perfectly, and gets the order wrong, and achieves the end.

[2] The distinction between behavior-types and goals is useful, and in some cases necessary to capture the structure and content of a plan. But in many cases it is plausible to think that the goal and the behavior-type share a referent. The goal will simply be to perform the behavior-type as specified.

In one sense Agent C got lucky. Usually, achieving ends requires some level of proficiency at behavioral execution, and at following the steps of a plan. In some cases, however, the end is achieved anyway. And of course in other cases, behavioral execution is perfect, as are the steps of the plan, and the agent fails. Perhaps the plan was risky. Perhaps it contained a fatal flaw. My son recently intended to eat a delicious piece of cherry candy by surreptitiously swiping the candy I held and shoving it mouthward. A flawed plan in one respect. For I held a disgusting cola-flavored piece of candy. So my son failed to achieve his aim.

The relevance of these distinctions is that when we speak of behavior conforming to aspects of a plan, we may have one of many aspects in mind. I will tend to gloss over this, speaking of behavior conforming to or approximating a plan. If it is important, however, we could always be more specific, and speak of plan quality, or of behavior conforming to a particular goal, or a particular means.

In general, then, controlled behavior involves a match or an approximation between behavior and aspects of the agent's plan. In particular, it involves a match or approximation between behavior and the ends (or goals) embedded in the plan, or between behavior and the means as indexed to specific ends. So we can speak of control with respect to a specific end, or a specific means, or with respect to the plan taken as a whole. But events not represented as contributing to the furtherance of the plan are not events under the control of the agent.

Regarding the kinds of content a plan can embed: this will depend upon the agent in question. Regarding humans the question is largely empirical. I say largely because there is a limit on what kinds of content could feature in a plan given the structure plans are supposed to take. Consider an iconically structured visual representation of a scene. Following Green and Quilty-Dunn (2017), this is a representation that meets the following principles. First, "Every part of the representation represents some part of the scene represented by the whole representation." Second, "Each part of the representation represents multiple properties at once, so that the representation does not have separate vehicles corresponding to separate properties and individuals" (11). A representation that meets these principles has very little internal structure, and given the way such a representation compresses information, it is difficult to abstract away from its parts. One lesson to draw from this is that some states, such as a simple iconic visual representation of a scene, may not be able to encode any kind of goal, and may have trouble encoding any structured sequence of behavior-types. (An icon, could,

however, serve as the specification of a goal. But a goal alone does not make a plan.) The psychological states that direct behavior need more than this.

How much more is open for some debate. Philosophers tend to talk of intentions as propositional attitudes. If the content of intentions is propositionally structured, then it is well suited for expressing plans. For propositions are systematic and recombinable, easily capable of representing sequences of behavior-types and of embedding goals. But it is arguable that we should not think of intentions as (exclusively) propositional attitudes (Coffman 2018). And anyway there are probably ways of tokening plans and goal-states without recourse to fully propositional structure. Philosophers have argued that we engage in practical reasoning via non-propositional representational states: map-like representations (Camp 2007), or analogue magnitude representations (Beck 2014), or mental imagery (Gauker 2011), or combinations of these (Shepherd 2018a). As I say, that is an empirical question, and is not my chief focus here.

Plausibly, then, plans can take a variety of representational forms (see Jeannerod 2006). The present point is that in order to exercise control over behavior, an agent needs a capacity to represent a plan for behavior, however simple or complex.

So one's representational capacities are one source of restriction on possessing a plan-state. Are there any others? Some philosophers have debated whether one could try, or intend, to do what one believes impossible (Thalberg 1962; Mele 1989; Ludwig 1992). These are not quite my questions here. I am talking about plan-states generally, and intentions are only one kind of plan-state. Intention possession may have additional restrictions that plan-states like urges, or sensory motivational states, or motor representations, do not. I am asking what is required for a system to possess a plan-state.

I am not asking what is required for a system to possess a plan-state with a specific content—a plan-state to A, for example, where A is an action variable. I have not offered an account of action yet. We are at a preliminary stage.

Even so, a worry similar to the one regarding intending the impossible arises here. Is it possible for a system to possess a plan-state, if the system cannot—lacks the capacity to—execute the plan? This kind of worry does not really arise with respect to simpler systems. If the states a simpler system tokens do not have the function of bringing about behavior that resembles the content of the state, there is little reason to consider the state a plan-state. Of course a system can find itself in unfavorable circumstances,

and token a plan-state that normally leads to success. Such a system could have a plan-state that is, in those circumstances or at that time, impossible for it to execute. We need to ask a more general version of the question. Is it possible for a system to possess a plan-state, if the system lacks the capacity to execute the plan in any circumstance in which the system could be placed?

In more complex systems, systems capable of some degree of delusion or self-deception, it becomes possible to envision a case in which the system has a plan for doing something, where the something is a thing the system cannot, in any case, do. Is that really a plan-state?

I suppose philosophers could disagree about this. Here is what I want to say. It is too strong to require that the system have the capacity to perfectly execute the plan. We should allow that a system can token plan-states that systematically aim too high, for example. What we should require is that the system have the capacity to cause (execute) some part of the plan. That is, in order to token plan-states, a system should have some causal potency. It is a minimal requirement, but a requirement nonetheless.

Causal potency can be understood as those causal powers (or dispositions) by which an agent behaves—causes things. To a rough approximation, an agent's exercise of causal potency can be measured in degrees, by indexing the exercise to a specific plan, or to a part of a plan. We can, for example, define approximation-level and perfect-level potency.

Approximation-level Potency. An agent J possesses approximation-level potency with respect to (means-end aspects of) plan-state P in circumstances C to degree D if and only if for J, P in C can play a causal role in the production of behavior that approximates (means-end aspects of) P's content to degree D in C.

Perfect-level potency. An agent J possesses perfect-level potency with respect to (means-end aspects of) plan-state P in circumstances C to degree D if and only if for J, P in C can play a causal role in the production of behavior that perfectly matches (means-end aspects of) P's content to degree D in C.

The possession of these levels of causal potency is not sufficient for the possession of corresponding forms of control. Consider Frankie:

Batter. Frankie stands in the batter's box, trembling. Frankie tends to strike out, and he has never hit a home run before. Part of the problem is his swing: an ugly, ungainly motion that rarely approaches the ball. In batting

practice, Frankie's coach will put a ball on a tee in front of him. Frankie hits the tee more often than the ball. Even so, Frankie recently saw a film that convinced him one simply needs to believe in oneself. Thus convinced, Frankie eyes the pitcher and whispers to himself, "Just believe, Frankie!" He then shuts his eyes and intends the following: "Swing hard, and hit a home run!" Here comes the pitch. With eyes still closed, Frankie swings hard and connects, producing a long fly ball deep to left field that lands just beyond the fence.

In his specific circumstances, Frankie possesses perfect-level causal potency regarding his intention to hit a home run in the given circumstances. Even so, the home run does not constitute an exercise of control (even if the eyes-closed swing of the bat does, to some degree).

What else does Frankie need? It is tempting to say that Frankie, or Frankie's intention, needs to bring about the home run in the right way. Frankie's swing, which by stipulation was an ugly, ungainly thing, is analogous to a case Al Mele (1992) introduced regarding a philosopher. This philosopher wanted to distract someone, and so intended to knock over a glass. But this intention upset him such that his hand began to shake uncontrollably, thereby knocking the glass over. The philosopher seems to have even less control than Frankie—in both cases the result accorded with the intention, but deviantly.

Consider the following as an account of control's exercise:

EC*. An agent J exercises control in service of a plan-state P to degree D if and only if J's non-deviantly caused behavior approximates (means-end aspects of) the representational content of P to degree D.

There is something right about EC*. First, it rules out cases like Batter as exercises of (high degrees of) control. Second, it is a very plausible idea that the degree of control an agent exercises has to do with the degree of approximation between behavior and plan content. An intention sometimes causes behavior that fails to perfectly follow the plan, and thus fails to perfectly match the content of the intention. Becky intends to make a shot that is all net—that goes in without hitting the rim or backboard. But the ball hits the front of the rim, then the backboard, and drops in. Clearly Becky exercised a degree of control—the shot was very close to all net, so close that it went in. But her behavior failed to perfectly match her intention. (If Becky bet

money on making it all net, this failure will be important.) Assuming that the plan is exactly the same, it seems Becky exercises less control regarding her intention if the shot is an inch shorter, hits the front rim and does not drop in, and even less if she shoots an airball. Third, EC* seems to capture a core truth about control's exercise: the exercise of control essentially includes an agent's bringing behavior to match the content of a relevant plan.

But EC*'s appeal to non-deviant causation is problematic. If there is no non-circular account of non-deviant causation in the offing, then we will rightly suspect that the account on offer is superficial. In effect, EC* will tell us that the exercise of control is essentially a matter of an agent's bringing behavior to match the representational content of a relevant intention in a controlled way.

I think there is a solution to this problem. It stems from reflection on control's possession.

2.3 Control's Possession

The agent exercises control when she behaves in a certain way, driven and guided by a plan and her commitment to it. In order to exercise control, agents must have control.

When somebody does something that seems lucky, we wonder if they could do it again. If they can do it again and again and again, we no longer believe it lucky. We think they have some control over what's going on. Agents that possess control are agents that can repeatedly execute a plan for behavior.

It's one thing to repeatedly execute a plan in very similar circumstances. But the world is capricious. We might want to see if the agent is poised to handle extenuating circumstances as she brings behavior in line with aspects of a plan. If so, the agent possesses flexible repeatability.

In general, an agent in possession of control with respect to some plan-state is an agent poised to repeatedly execute that plan, even in the face of extenuating circumstances.

To illustrate: hold fixed Frankie's intention and suppose a number of things. Maybe the ball comes in 1 mph faster or slower, or an inch higher or lower, or Frankie's muscles are slightly more fatigued, or Frankie produces a slightly different arc of swing. We can vary Frankie's circumstances any way we like and ask: across this set of circumstances, how frequently does Frankie evince the potency he evinced when he hit the home run? The

answer to this question will give us a measure of the control Frankie possesses regarding his intention.

In order to make sense of flexibility and repeatability, we have to specify a certain set of circumstances. This is not necessarily to say that the possession of control is composed (even in part) of extrinsic properties. In discussing her view of causal powers, Rae Langton distinguishes between extrinsic properties and relational properties, as follows: "whether a property is extrinsic or intrinsic is primarily a metaphysical matter . . . whether a property is relational or non-relational is primarily a conceptual matter: it is relational just in case it can be represented only by a relational concept" (2006: 173). As Langton notes, it is natural to view causal powers as both intrinsic and relational: intrinsic because such powers are "compatible with loneliness" and relational because "we need to talk about other things when describing it" (173). This view is available regarding the control an agent possesses.

Many agents are plastic—we lose limbs, muscle tissue, brain cells. Our control is therefore plastic across circumstances. We learn novel ways of performing tasks, and become adept with various tools. Andy Clark claims that our brains are "open-ended opportunistic controllers"—our brains "compute, pretty much on a moment-to-moment basis, what problem-solving resources are readily available and recruit them into temporary problem-solving wholes" (2007: 101). I think he's right. It follows that circumstances impact the amount of control we possess regarding our plans. So the specification of a set of circumstances requires care.

We get viable and interesting measures of control only when the set of circumstances is well selected. A set of circumstances is well selected when we follow principles for set selection that roughly mirror principles for building an accurate causal model of the agent as embedded in a broader causal system that comprises the kinds of circumstances in which we are interested.

So, for example, the set should be sufficiently large. Think of a set of circumstances with only two members: the case in which Frankie hits a home run, and a case in which he misses the ball. This set is not informative: we need a large number of cases before we get any useful information regarding just how lucky Frankie's home run was. A set is sufficiently large when adding members does not substantively impact the resulting measure of control.

Further, the circumstance selector should accurately specify the parameters that are fixed, and the parameters that vary. In some cases the selector

should specify the degree to which the parameters are permitted to vary. The selector might go beyond this, depending upon the purposes of the exercise. One might attach a probability distribution to the operation of certain parameters, for example. How all of this will go will depend to some extent on the selector's theoretical interests, and to some extent on the control-relevant features of the case—the agent's plan, the agent's constitution.[3]

There are many ways to build a theoretically fruitful set of circumstances. This follows plainly from the facts that there are many agent-types and many circumstance-types, and so the ways that these can come together in contexts of plan execution will be several. At an abstract level, given some agent-type (e.g., an adult human agent) as a target, some strategies for building sets of circumstances will be more theoretically fruitful than others. I elucidate this via discussion below.

As we have seen when discussing causal potency, we can specify a level of content approximation regarding a plan-state, or some aspect of a plan-state. Doing this gives us an in principle way to measure an agent's degree of repeatability with respect to any given level of content approximation. An agent J's degree of repeatability DR with respect to some level of content approximation L in a sufficiently large set of circumstances C is given by J's success-rate at reaching L across C, where successes are exercises of causal potency to the relevant level of content approximation or higher. An agent's degree of repeatability (with respect to some level of content approximation) gives us the degree of control she possesses (with respect to that level of content approximation):

PC. J possesses control to degree DR with respect to some level of content approximation L for a plan-state P in circumstances C if and only if J's success-rate at reaching L across C is DR, C is well selected, and P plays a causal role in J's behavior in every member of C.

Bill is throwing darts. Across a set of 100 circumstances, Bill possesses the intention to hit a bullseye. Suppose, now, that Bill hits the bullseye eleven times: his success-rate at this "perfect" level of content-approximation is

[3] Sets can be more or less comprehensive, in covering the range of factors likely to impact success-rates. Agents can possess control even if the set to which we index the measure is not comprehensive. Consider the agent who succeeds at high levels but only in rare circumstances. But comprehensiveness is important to account for. We can say that control has multiple dimensions of gradability, and one of these is comprehensiveness.

0.11. We might focus on other levels of content-approximation as well—the quality of Bill's release, of his stance, of his direction of attention, of his proximity to his target. It is informative to know that Bill places the dart within 1 inch of the bullseye forty-six times, and within 5 inches of the bullseye eighty-two times. We might even change the set of circumstances—adding in various challenging contingencies—in order to measure Bill's control in various ways (I discuss such contingencies below). With each well-selected set of circumstances, and each level of content approximation, we learn a bit more about Bill's dart throwing capacities.

2.3.1 Control with Respect to a Goal

Depending on our purposes, we may wish to hold fixed various aspects of the agent's plan. To see why we might want to hold fixed a goal (or end) embedded in a plan, consider two agents playing darts: one a recreational player (Torrey) and the other a professional (Bill). Torrey thinks he is good, but is he really? No. Especially not compared to Bill. Both possess intentions with this goal embedded: "Hit a triple-twenty." Knowing only this much, we can predict that Bill and Torrey will differ mentally. As they step up to throw, they will likely possess different plan-states. Due to the time spent honing his skills, Bill's repertoire of plans for executing his intention is larger, more complex and subtle, and so on. The details need not concern us—perhaps they include Bill's stance, the feel of Bill's throwing arm, the grip position of his fingers, the direction of Bill's attention. What is important for our purposes is that the type and availability of plans plausibly contributes to Bill's comparatively higher degree of control across a wide range of circumstances.

We can accommodate the difference between Bill and Torrey by specifying a shared aspect of their different plan-states—something they are both capable of aiming for, such as the goal to hit a triple-twenty. This allows us to capture the way an agent's representational capacities contribute to her control, and it allows a more flexible measure of control:

GC. An agent J possesses control to degree GC with respect to a goal G in circumstances C if and only if J's success-rate at achieving G across C is GC, C is well selected, and a plan-state that embeds G plays a causal role in J's behavior in every member of C.

Note that what we have done in isolating a goal we could do for a wide range of aspects: particular behavior-types, particular sequences of behavior, particular kinds of circumstance, and more. In other work I articulate a notion of weak motivation and argue that weak motivation impairs an agent's possession of control (Shepherd 2017a). One might just as well consider how agents fare in conditions of strong motivation, or in conditions of normal motivation. Or one could isolate other factors. Circumstances that involve distracted attention are interesting for some purposes. So too are circumstances that involve heightened or diminished perceptual acuity. Sport psychologists are very good at studying behavioral profiles across interesting variations of just these sorts.

Suppose the amateur is pretty good at his pub, with his darts, but that other pubs and other darts distract him, significantly lowering his success-rates. The pro, however, has a high degree of repeatability no matter the venue, and with most brands of darts. This type of example illuminates a potentially useful measure of control's possession, related to the wideness of the set of scenarios under consideration. In general if J can exercise control with a certain degree of repeatability across a wider range of scenarios than K, J is plausibly better (at least in one sense)[4] at executing the kind of plan-state in question than K.

2.3.2 Basic Levels of Control and Multiple Sets of Circumstances

As I have explicated things so far, an agent's possession of (some degree of) control is relative to a plan-state or some of its means-end aspects, as well as to a well-selected set of circumstances. There are a number of theoretically fruitful ways to select sets of circumstances. What is required is not that any particular set of circumstances be selected, but rather that the selection be fruitful.

This perspective on control leaves room for the following kind of situation. The agent is in some particular circumstance—either hypothetically, or actually—with some particular goal, and we want to know how much control she possesses with respect to that goal. It turns out that the

[4] Agents will sometimes train for, or display talent at, performance in specific circumstances. Some agents may be better than most or all others at some action-type in some restricted range of circumstances.

circumstance she is in contains a range of parameters we might hold fixed. The agent is feeling anxious, and her anxiety levels may increase. The agent is feeling fatigue, and she may begin to feel a strong sense of effort as she acts in one way or the other. The agent's environment contains certain obstacles that might impact her performance levels, and it is unclear which obstacles she may encounter. It will be possible, when assessing the agent's level of control with respect to the goal, to select multiple sets of circumstances. And the agent's level of control with respect, say, to goal achievement—here I'm adding overly precise numbers for effect—may be 0.8 across one set, 0.7 across another, 0.26 across still another.

What are we to say about such cases? One option is to claim that some particular set of circumstances—perhaps the union of all the well-selected sets—is more fundamental than others. On this option, there is a basic level of control an agent possesses. In many cases we require a massive set of circumstances to figure it out.

I resist this option. My rationale is as follows. The point of selecting a set of circumstances is to get (or to fix on) good information regarding how the agent is composed, and how she is disposed to behave, given important factors—stress, fatigue, working memory capacity, availability of tools, etc. Combining well-selected sets may provide useful information in some cases. This will be true, for example, when movement amongst distinct factors produces an interaction effect. But in other cases combining sets of circumstances will degrade information by muddling explanatory factors better kept distinct. So in the case in which we combine all well-selected sets of circumstances, we may end up with something like a single number, but at the expense of any detailed explanation for why the number is what it is. The kinds of systems that engage in controlled behavior are often complex, and there are multiple factors that influence performance, and multiple dimensions along which behavior will differ. It is ultimately unnecessary to average over all the sets of circumstances. We can, instead, view the agent's possession of control as a set of data points, indexed to multiple sets of circumstances. The possession of control looks, in the end, more like a multi-dimensional histogram than a single number; it looks more like a frequency distribution of markers of performance across the range of cases, with as many dimensions as are relevant to capture the behavioral profile.

One implication is that we should not expect to have available, in every case, a clean verdict regarding who possesses more control with respect to some goal or plan. Of course we sometimes speak in this way. We say things

like, "As a dart player, Bill has much more control over his placement than Torrey." Or we say, "Bill has more control over his placement tonight than he did last Tuesday." In my view, such talk should not be taken to refer to anything precise. Often such talk builds in assumptions about normal circumstances, or prototypical performances. Sometimes such talk involves useful oversimplification. Just as often I suspect such talk is unhelpfully vague. If we wish to be precise about the control an agent possesses (or exercises), I think the analyses offered above are the way forward. Our normal talk of control should be seen as a loose, sometimes helpful, sometimes confused, way of referring to the more precise understanding of control developed here.

2.4 Underlying Metaphysics

I have spelled out what it is for an agent to possess control. The metaphysician may feel as though I have offered a framework for control while leaving the underlying metaphysics untreated. They may think that control cannot simply be a distribution of success-rates, or behavioral profiles. Something must explain the success-rates and the profiles.

It seems to me enough to say that control is, in any given case, a package of causal properties sufficient to produce the specified performance profile across the specified circumstances. But one may wonder whether more should be said (certainly, more can be said). For example: might the package of causal properties chiefly or exclusively involve or be best identified with abilities, or perhaps dispositions? There are lively literatures discussing both the nature of abilities, as well as the nature of dispositions.[5] Perhaps the underlying metaphysics of control are best thought of along lines taken by some or another view. Or perhaps the notions of abilities or dispositions as most metaphysicians speak of them are altogether inappropriate as candidates for further explication of control.

In what follows I first consider the relation between control and ability, and then the relation between control and disposition.

[5] One might propose powers rather than dispositions or abilities. Many use the term powers as synonymous with dispositions. Some think of powers in terms of dispositions. But there may be usages that carve out a special role for powers. I do not consider the matter here.

2.4.1 Ability

I do not think control can be understood in terms of ability—at least not as typically discussed. This is because as typically discussed, abilities involve action. John Maier (2018a), in the *Stanford Encyclopedia* entry on "Abilities," suggests that abilities are special kinds of powers. Powers are possessed only by agents, and are expressed by the modal auxiliary "can." What makes abilities special is that unlike powers, they involve action. One might have the power to understand Catalan. But if one can speak Catalan, this is an ability, because speaking Catalan is an action.

Control over behavior is prior to, and necessary for, action. So the relation between control and ability will be one on which the direction of explanation is from control to ability. I do not offer a full version of such an explanation here. The main reason is that there are several candidate accounts of ability in the literature, and this is not the place to decide which ones are better. But it may be useful, for purposes of illustration, to suggest how such an account could go.

Consider a new dispositionalist's analysis of ability:[6]

> S has the ability at time t to do X iff, for some intrinsic property or set of properties B that S has at t, for some time t' after t, if S chose (decided, intended, or tried) at t to do X, and S were to retain B until t', S's choosing (deciding, intending, or trying) to do X and S's having of B would jointly be an S-complete cause of S's doing X. (Vihvelin 2004: 438)

Simplifying, the idea is that an agent possesses an ability to A iff the agent A-s if she tries, decides, or intends to A. This account has been criticized by a number of philosophers (e.g., Clarke 2009, Steward 2012); I am not trying to defend it. Observe however that the account as stated can be seen in terms of control. The account claims that an agent has an ability to A iff the agent possesses and retains certain intrinsic properties—certain dispositions—the manifestation of which constitute a successful execution of an intention (or a trying). As many have pointed out (Austin 1956), that seems far too strong. Agents have abilities to A that fall short of successful performance in every instance. A revised version

[6] See also Fara (2008).

of this account of abilities may do better by seeing abilities as graded in the manner of control. Alternatively, one may think of the possession of an ability as the possession of a sufficient degree of control in some well-selected set of circumstances, or in a sufficiently broad set of circumstances (cf. Manley and Wasserman's (2008) view of dispositions). This may not help with the problem that motivates new dispositionalists—the problem of free will—but it may help with the problem of getting the nature of abilities right.

John Maier offers a view of ability that draws on the metaphysics of modality as well as the semantics of generic statements. Maier wants an account that can explain why an unskilled golfer who luckily sinks a hole-in-one does not have the ability to do so, while a skilled golfer has the ability to sink a putt even though she sometimes misses. On Maier's account, "S has an ability to A just in case A is generally an option for S" (2018b: 425). Option is a technical term, meant to capture an intuitive sense in which some action is actually available for an agent at a time. Generally is a technical term too—"an agent generally has an option" should be understood according to the best semantics for generally. However, Maier admits that we may yet lack an adequate semantics, rendering his account "a promissory note" (421, fn. 7). Maier further suggests that the account may be seen as reducing problems regarding ability to a problem regarding the semantics of the "generally" operator.

It is not my aim here to criticize this as an account of ability. I wish only to illustrate how control may help illuminate it. We can, for example, understand options at least partially in terms of control. I find it plausible that control is at least a necessary condition on the presence of an option: an agent has an option to A in some circumstance T only if she has some sufficient degree of success-level control regarding a plan that includes A-ing as a means or an end across a well-selected set of circumstances C of which T is a member.

I also find it plausible that control is a sufficient condition on the presence of an option: an agent has an option to A in some circumstance T if she has some sufficient degree of success-level control regarding a plan that includes A-ing as a means or an end across a well-selected set of circumstances C of which T is a member.

If this is right, then we need not take options to be metaphysically primitive, as Maier does. We can instead understand the presence of options in

terms of the possession of control at a time. In this way control becomes crucial to an account of ability.

Might control's possession be important as well for an understanding of when an agent generally has an option? I suspect it could be, but this is not crucial to the point I wish to make here, which is that ability is the wrong candidate for further metaphysical elucidation of control.

2.4.2 Dispositions and Explanation

Dispositions are more promising as candidates for the metaphysical basis of control.

There are many accounts of dispositions, however. And there is little consensus regarding which of them are likely to be the truth. Some doubt dispositions can be analyzed, and happily deploy the term unanalyzed—e.g., Stanley and Williamson (2017) do this in a dispositional account of skill. Others deny that dispositions are causally efficacious (Armstrong 1968; Mackie 1977). Given the complexity and sophistication of the literature on dispositions, my goal here cannot be to cover the ground (see Choi and Fara 2018 for an introduction). I wish only to illustrate the suitability of the notion of dispositions in this context.

One view of dispositions has it that dispositions are individuated by a stimulus condition and a manifestation condition. Dispositions are fundamentally dispositions to M (to manifest) when S (when characteristically stimulated). So a vase is fragile if it is disposed to break when struck, or dropped, or…fill in the characteristic stimuli. This view can be formalized into the Simple Conditional Analysis: An object is disposed to M when S iff it would M if it were the case that S. Thus formulated, this view of dispositions is subject to numerous counterexamples (Martin 1994; Bird 1998).

Problems with the simple conditional analysis give rise to more sophisticated accounts (see, e.g., Choi 2006, 2008). Two problems for more sophisticated accounts are salient here. First, might an object or system have a disposition even if that disposition does not manifest in every relevant case? And second, how should we specify what it is for some system to be more disposed to M than another?

Manley and Wasserman (2008) note that in spite of problems with conditional analyses, there seems to be some connection between dispositions and circumstances. In some cases, at least, the relevant manifestation conditions seem built in:

> It is hard to believe that there is no interesting connection between conditionals and ordinary dispositional ascriptions. (This would be especially surprising if we accept that some explicit dispositional ascriptions—the highly specific ones—are actually equivalent to conditionals.) (2008: 73)

Manley and Wasserman propose an account that indexes dispositions to C-cases, where these are precise specifications of values for relevant parameters. So, as they say, in the case of fragility, we might wish to specify the height from which an object is dropped, the density of the medium the object is in, and other intuitively important parameters. Once we specify the relevant parameters, we understand the possession of a disposition as follows: "N is disposed to M when C if and only if N would M in some suitable proportion of C-cases" (76).

This reference to a suitable proportion allows them to truthfully ascribe dispositions even in cases of failure to manifest. But what of the gradability of dispositions? Here Manley and Wasserman recognize trouble. The trouble is due to their liberality with respect to the size of the sets delineated in their analysis. Here is what they say:

> Because the C-cases in our domain need not be restricted in any way, absent stimulus conditions are not a problem for [our analysis]. We can simply allow that N is loquacious just in case N would talk in a suitable proportion of situations—any situations at all. (2008: 77)

They wish to understand dispositions in terms of a suitable proportion of manifestation across extremely large sets of circumstances. Now, I wouldn't deny the usefulness of the set of all situations for some purposes. But in many cases, we have to do better. My framework for control counsels us to pay attention to the behavioral system in question. Depending on the system, and the behavior, some circumstances are relevant, and some are not. It is no relevant part of a decision-making system's control profile whether it can decide what to do if transported in time past the existence of the planet, or the death of the last black hole. So, instead of offering an account of dispositions in terms of proportion of manifestation across all cases, I think it better to index manifestations to more explicitly regimented sets of cases.

This allows one to say that in some circumstances, a system S is more disposed to M than another system Y, even if Y is more disposed to M in different circumstances, and even if there is no fact of the matter regarding

whether S is more disposed to M than Y across all cases. Consider this case, due to Manley and Wasserman:

> Take two objects. Suppose that humans, with our paltry strength, are incapable of breaking the first, but that it would take little or no effort for a giant to break it. The second object, however, is strange—it will break under very light pressure, but it is impervious to the destructive efforts of giants who are incapable of applying light pressure. Now, it may well be that, given a suitable notion of proportion, these two objects break in an equal proportion of cases involving the application of stress. The trouble is that we are likely to call the second object 'fragile' but not the first, while the giants will have opposite inclinations. Clearly, we are not only concerned with the proportion of stress-inducing situations in which an object would break, because some of those situations matter more to us than others. And these situations are different from the ones that matter to giants. (2008: 78–9)

How to deal with this? Manley and Wasserman propose a context-dependent weighting on cases. On this weighting, for example, Object 2 is more fragile when humans are talking about it, and Object 1 is more fragile when Giants are talking about it. It is not clear to me whether Manley and Wasserman would also want to say that there is no illuminating answer regarding which thing is more fragile, full stop. But this seems reasonable to me. So I think some dispositions may be best understood as a disposition-in-C, where C is a well-selected set of circumstances.

But we need not endorse this view in order to elucidate the connection between control and disposition. Even if dispositions cannot be understood as indexed specifically to certain sets of circumstances, control can be understood as constituted by circumscribed dispositions: dispositions-in-C.

This allows us to explain control in terms of something more fundamental than success-rates, while remaining neutral on the nature of dispositions themselves.

It may be useful to illustrate that this lesson applies even if dispositions are very different than Manley and Wasserman think. Barbara Vetter (2014) rejects the view that dispositions are individuated by a stimulus condition and a manifestation condition. Instead, Vetter proposes that dispositions are individuated by manifestation without characteristic stimuli. "A disposition is individuated by its manifestation alone: it is a disposition to M, full-stop" (2014: 134).

Vetter's motivation here is to replace analyses of dispositions in terms of counterfactuals with an analysis of dispositions in terms of potentialities—a modal notion she develops at some length in her (2015). One can see how this might allow an explanation of the possession of a disposition that is compatible with a failure to manifest on occasion. Potentialities may work like this. But how are we to understand what it is for a system to be more disposed to M than another?

What I find congenial in Vetter is that, once she has rejected the intuitive delineation of manifestation conditions that characterizes many counterfactualist discussions of dispositions, she is free to recognize just how context-relative ascriptions of dispositions may be. It looks like, even if a disposition can be understood independently of some set of circumstances or stimulus conditions, a disposition's gradability cannot be understood independently of some set of circumstances or stimulus conditions.

Vetter characterizes the gradability of a disposition in terms of proportions of manifestation across cases. A case involves a world, a time, and an object:

> What should determine the proportion that makes a person count as irascible, or a disease as transmissible, is not the number of worlds in which the person becomes angry at least once or the disease is transmitted at least once, but rather the number of individual instances of anger or transmission: in other words, of cases. (Vetter 2014: 140)

This is in the neighborhood of my own method of selecting sets of circumstances for measuring control. So, too, is Vetter's line of thought for getting the right cases. She emphasizes variety:

> Plausibly, an object which breaks in a greater variety of circumstances is more fragile than one which breaks only in one very precisely circumscribed set of circumstances, and likewise for other dispositions.
>
> (Vetter 2014: 142)

Of course, deciding which types of circumstances qualify as producing greater variety requires one to be more specific about the way one selects sets of circumstances—something I have stressed regarding control. As before, regarding control, the specificity we need will depend on the system in question, and on the behavior in question. To restate: paying attention to these details may give one a way to offer an account of control's

possession in terms of packages of dispositions-in-C—that is, not in terms of dispositions full-stop, whatever they may be, but in terms of how a system is disposed in specific circumstances. Such an understanding can explain control in terms of dispositions, while remaining neutral on the specific nature of dispositions (and on whether dispositions can be reduced to something else).

Readers may have a complaint with this dispositional understanding of control. For notice that this dispositional understanding claims to further illuminate the nature of control, by linking behavioral profiles to something slightly deeper—a system's dispositions. But are these dispositions really deeper? It may sound trivial to say, as I effectively have said, that when a system J has X-amount of control across some set of circumstances this is because J is disposed to behave in a way that confirms the X-measure across the set of circumstances.

When we are considering one set of circumstances alone, I think this complaint is fair. But the way I have articulated the relation between control and dispositions allows one to offer deeper explanations of behavioral profiles by appealing to dispositional commonalities across multiple sets of circumstances. As I have already noted, in many cases what one needs if one wants a more comprehensive understanding of a system's control, or behavioral profile, is to explore that system's behavior across circumstances that hold fixed different important parameters. So we could, for example, explain J's behavioral profile across one set of circumstances C not simply by appeal to J's dispositions-in-C, but by appeal to J's dispositions across multiple sets of circumstances, including C. C will fit into a broader set that more comprehensively illuminates J's behavioral structure, and by linking J's dispositions-in-C to J's dispositions more broadly, J's control-in-C is given an explanation.

2.5 Conclusion

An agent's possession of control is relative to a plan-state or some of its means-end aspects, as well as to a well-selected set of circumstances, which is essentially a well-constructed causal model that embeds the agent into a particular contingency space. Since there are many ways agents become embedded in these spaces, and since different spaces will test different

aspects of the agent—different dispositions and disposition-clusters—there are many theoretically fruitful ways to select sets of circumstances.

In chapter 3 I build on this understanding of control, and apply some of the ideas to the problem of deviant causation. For it seems that one cannot have control without non-deviant causation. More than one writer has made this connection. As Jesús Aguilar has put it:

> What all cases of deviance exhibit is precisely the undermining of the possibility for action by the undermining of agential control. In turn, this explains why any plausible strategy to deal with the problem of deviance boils down to an effort to restore in one way or another agential control.
>
> (Aguilar 2012: 9)

Aguilar speaks of the undermining of action by the undermining of control. I have said very little about action thus far. Just now I am talking about control. In chapter 3 I demonstrate how an account of non-deviant causation falls out of reflection on control.

But I am ultimately headed for action, and I want my account of control to be of use there. Indeed, I will argue that my account of control plays a crucial role in a satisfying account of intentional action. That happens in chapter 5.

3
Non-deviant Causation

3.1 Introduction

The last game of my high school basketball career, we almost completed an unexpected, improbable comeback against the defending state champions. After clawing back for two quarters, we had finally taken a one-point lead. We had the ball. There was little time left. A turnover occurred, and their forward dribbled madly towards an open lay-up. I, along with a teammate (I can't remember who), chased him down. He, surprised and nervous, jumped in the air with his back to the basket. I think that's what happened. Memory sometimes operates deviantly. But I recall seeing him in between us, ball down near his thigh, and breathing a sigh of relief. He had no chance of making the shot. I doubted he would even get a shot up. And yet. Perhaps in spite of his frantic state he managed an intention to make a shot—to win the game. An audacious intention, but winning a state championship, maybe, imbues you with audacity. What I know is that with a little flipping motion, he sent the ball up. It crawled onto the rim with a strange swirling spin that pulled it over and in.

Deviant causation ruins careers.

It also undermines intentional action. In cases of basic deviance, the problem occurs in between mental state and behavior. A philosopher wants to distract a commentator, and intends to knock over a glass. But this intention upsets him such that his hand shakes uncontrollably and accidentally knocks the glass over (Mele 1992: 182). Since the intention does not cause behavior in the right way, the philosopher does not intentionally knock the glass over.

Consequential deviance occurs after the behavior has begun, and before it is complete. Jones intends to kill K by shooting him. Jones misses by a mile, but the noise from his rifle scares a nearby herd of pigs that trample K (Davidson 1980: 78). The killing is unintentional because, again, it was deviantly caused.[1]

[1] A third type of deviance, tertiary deviance, concerns the history of the relevant mental state (see Mele 1987: 56). My account of non-deviant causation is not intended to address tertiary deviance.

The Shape of Agency: Control, Action, Skill, Knowledge. Joshua Shepherd, Oxford University Press (2021).
 DOI: 10.1093/oso/9780198866411.003.0003

I like to imagine the end of my high school basketball career as a case of double deviance. My opponent's intention was so audacious and unnerving that it led to the ugly-looking flip shot. And the spin was so wild, so out of control, that it corrected for the awful angle at which the ball hit the rim. The made shot was no intentional action.

Many have proposed solutions to causal deviance (e.g., Bishop 1989; Enç 2003; Schlosser 2007). None enjoy consensus. Is this a serious problem for causalists about intentional action? Some seem to think not. According to these theorists, deviant causation is a minor worry. According to others, the problem does not need to be solved (Tännsjö 2009).[2]

But for many, the problem is deep enough to motivate rejection of causalism. Ezio Di Nucci, for example, makes a characteristic inference in the following passage:

> Deviant causal chains are symptomatic of a fundamental problem with causalism . . . we should give up the fight to accommodate deviant cases and focus, rather, on developing an alternative to causal views of action . . .
> (Di Nucci 2013: 45)

Or, to take another example, Timothy Williamson writes of causalism's treatment of deviant causation in the past tense, as though the failure were complete:

> [The causal theory] succumbed to the problem of deviant causal chains: the causation from intention to success may be of the wrong kind for action . . . Attempts are still made to solve the problems for attempted analyses of action by adding epicycles, but none has worked.
> (Williamson 2017: 171–2)

In this chapter I offer a solution that works.

[2] John Hyman (2014) argues that deviant causation cases present a problem, not to causalist accounts of action, but to the Humean theory of causation. Once we reject the Humean theory in favor of a powers and dispositions-based ontology, the problem of deviant causation dissolves. I am not committing here to a particular view of causation, but Hyman's view, also expressed in the cited paper, that action explanation can be both causal and teleological, is congenial in some ways to remarks I make below.

3.2 Non-deviant Causation Explained

What we want is to explain the difference between the causal pathway(s) operative in two kinds of case.[3] The account of non-deviant causation I offer here draws on the account of control's possession developed in chapter two.[4]

Recall the shape of that account. The possession of some amount of control with respect to a plan-state is the possession of a capacity (or package of dispositions) to repeatedly and flexibly bring behavior to match the content of that plan-state, to some degree, across a well-selected set of circumstances. What we need in order to move from the possession to the exercise of this capacity is a sense of the causal routes taken in uncontroversially successful instances of the capacity's exercise.

Regarding control's possession, we were somewhat relaxed about the selection of sets of circumstances. Non-deviant causation happens in a particular circumstance. So we need a way to say, of some particular happening, that it was non-deviant. We need to move from a particular happening to the modal space surrounding that happening, against which the particular happening can be revealed as falling into a certain pattern.

We need to deploy care. Consider the following kind of issue.

Tormund is anxious at t. Fix his current level of anxiety as part of a well-selected set of similar circumstances, and he does poorly, via particular causal routes that involve scattered attention. Tormund is also well-rested at t. Fix his current level of energetic resources as part of a well-selected set of

[3] When I first developed this account, I used the notion of causal pathways without much understanding of how some scientists deploy it. Then I read Winning's (2019) excellent discussion of control, and Ross's (2018) very interesting work on differences in scientific usage of mechanisms and pathways as constructs for causal explanation. For those interested, the account of non-deviance I offer could be taken to apply to systems conceived as mechanisms, or as (sometimes ad hoc) packages of mechanisms, or to pathways. The notions of deviance and non-deviance seem to me to apply more abstractly, both to the operations of well-specified mechanisms, and to the modelling of various pathways in whatever causal space one wishes to model.

[4] My proposal is similar in spirit to Aguilar's (2012). Aguilar appeals to the notion of reliability in developing a sufficient condition for non-deviant causation: according to Aguilar, the mechanism that produces an intentional action must be reliable (2012: 10). In my view this is on the right track. One might thus see what follows here as a way to flesh out the nature of the reliability Aguilar discusses. I do not know if Aguilar would agree with this. The view I offer is, after all, very closely tied to the account of control developed above. Further, Aguilar applies his discussion of reliability only to overt basic actions—the account developed here is developed independently of an account of intentional action, but is intended to apply to all intentional action-types (mental and overt). Finally, in my view Aguilar's discussion does not place sufficient weight on the importance of circumstance to a proper understanding of reliability. Even so, I suspect that Aguilar and I are in substantial agreement.

similar circumstances, and he does well, via particular causal routes that involve more focused attention. Two well-selected sets give conflicting verdicts about how much control Tormund might be thought to possess in the actual circumstance he is in at t.

Say, now, that Tormund behaves at t in a certain way—this way is unusual given his anxiety, but fairly normal given his level of rest. Is this a deviant success, as the set that fixes his anxiety level would suggest? Or is this a non-deviant success, as the set fixing his energetic resources would suggest?

Neither. The problem is that if we wish to understand control in the actual happening, we need to give preference to the actual happening in how we understand the modal space surrounding it. Not just any old well-selected set of circumstances will do. We have to treat the actual happening as baptismal regarding the modal space surrounding it. We need to start from Tormund's actual circumstance and specify all of the relevant causal parameters—the anxiety and the energetic resources, and much else besides—into the relevant set of circumstances. The actual circumstances, then, play a constraining role on the system one is modelling by selecting a set. They do so in at least two ways.

First, they force greater specificity and detail regarding the dispositional condition of the agent in the circumstances. Agents vary over time. Here we need to get a snapshot of the agent in order to proceed.

Second, the actual circumstances have to be typed in some way. We need to understand how the circumstances may vary. This will likely come in the form of probability distributions regarding certain ways things might unfold. This will present epistemic difficulties in many cases. These problems are familiar from attempts to build causal models of complex systems. These models require judgments about causal variables and causal patterns in the flux of events. When we make such judgments, we identify the type of circumstances that fit into our model—we specify parameters for what makes some causal variable or causal pathway (which could be construed as a pathway amongst variables) count as sufficiently similar to another. Of course these models will often be rough approximations. But we can envision how this is supposed to go. And we can allow that things often change over the course of behavior, so that variations in circumstances come into play.

Call a set of circumstances that begins in this way, specifies the relevant parameters, and types the circumstances into a (probabilistic) model of possible permutations, a comprehensive set of circumstances. This need not be an infinite number of circumstances. It depends upon correctly gauging

the circumstance-type the agent is in in the baptismal circumstance. Sometimes only a few parameters are relevant. Sometimes there will be more.[5]

Here, then, is the account:

Non-deviant causation. Given an agent J, a plan-state P, a comprehensive set of circumstances C, and a token circumstance T which is a member of C, J's behavior in service of P is non-deviantly caused in T, relative to C, if and only if (a) J's behavior in T reaches a level of content-approximation L (to the content of P, or to some part of P) that is within a sufficiently normal range for J, relative to J's behavior across C, (b) J's behavior in T is produced by way of causal pathways that, when taken, lead, with sufficient frequency, to J's reaching L across C, and (c) P is among the causal influences of J's behavior.

We want to exclude deviantly caused successes. Non-deviant causation does so by way of two requirements. The relevant level of content-approximation—the level our agent actually achieves—must be within a sufficiently normal range for a comprehensive set of circumstances. So this will be something the agent can do repeatedly, in circumstances like this. And the causal pathways operative in the actual case must be such that taking those pathways leads to something close to the behavior in question, in circumstances like this. There is no room for a deviant causal pathway to operate surreptitiously.

There is nothing magic at work here. It takes some work to get it straight. But in a way, this is simply a plausible precisification of what is implied by the term non-deviance.

3.3 An Alternative: Wu on Control and Non-deviance

Before considering objections to this account, I wish to hold it up against an alternative: Wayne Wu's account. Doing so will help clarify the substance of my own.

Wu agrees with the general view that control is essential for understanding action and agency:

[5] Compare this to Manley and Wasserman's comments about "C-cases" regarding dispositions (2008: 75).

> Agentive control yields phenomena of central philosophical interest: moral, rational, reason-based, skilled, conscious, epistemic, and free agency. To understand these specific forms of agency, however, we must understand the core phenomenon, agentive control. (Wu 2016: 101)

How does Wu attempt to account for agentive control? For Wu the operation and interaction of both intention and attention hold the keys: "If you like, talk of 'cause in the right way' or 'appropriate causal role' is spelled out by delineating more completely the psychological basis of agency, namely the intention-attention nexus" (114).

Wu claims that the intention-attention nexus is the psychological basis of agency, and that understanding this nexus will help us understand non-deviant causation. How so?

According to Wu, in normal cases of action agents face a "many-many problem" (see Wu 2011). Agents face many possible targets for action, and agents face many ways to implement action related to these targets. In order to solve the many-many problem, agents must find a one-one mapping from target to action. (They need not do so through deliberation—often agents face a non-deliberative many-many problem.)

For Wu, solving the Many-Many Problem is a necessary condition on action. Entities that never solve the many-many problem would be in a position such that only one target for action, and only one movement-type, is available to them. According to Wu:

> What is by hypothesis not available to them is the possibility that that very stimulus can be mapped to different response at that time such that the creature could react to the stimulus in this alternative way rather than the pre-set way. Their behaviour does not count as action, for at each time they are driven by what is essentially a reflex, and I take it that this never exemplifies agency. (Wu 2011: 54)

The important question now, for Wu, is how human agents solve the Many-Many Problem, and thus act intentionally. Here is where the "intention-attention nexus" is critical. Wu (2016) reviews evidence that both intention and attention are crucial for establishing one-one mappings in paradigm cases:

> The idea is that in setting intention, one sets the weights that biases which selections are made in action (psychologists speak of task sets). So, if one intends to act on X, then X is selected for action; if on Y, then Y is selected.

> This is part of intention's causal role, one that is driven by the content of intention. The point is that by registering attention in the causal theory, we can expand our conception of intention's causal role. (Wu 2016: 110)

How does the biasing role of intention, and the selective function of attention, deliver an account of non-deviant causation? Wu leverages these biasing and selective functions to offer a sufficient condition for non-deviant causation:

> S exerts agentive control in F-ing at t if S's F-ing is a result of an intention to F where this intention solves the Many-Many Problem faced by S at t by directing appropriate attention-mediated coupling. (Wu 2016: 114)

In his (2016), Wu suggests we take this condition as necessary as well. This is consistent with his (2011), where he argues for the following analysis of (bodily) intentional action.

Necessarily, for all agents, S, bodily action types B, motivational states, M, towards B-ing and action contexts C:

> S's motivational state M (paradigm: an intention to B) structurally causes the selection required in implementing a solution to the Non-deliberative Many-Many Problem appropriate to action B for S in C if and only if S intentionally B's in C in the sense that she B's because she is in motivational state M. (Wu 2011: 68)

So, for Wu, agents never exercise agentive control unless a relevant intention structures attention-mediated coupling. As Wu would have it:

> It turns out then that deviance is quickly addressed in a way that we expected. For causal deviance revealed that the causal theory did not adequately explicate agentive control. Such control reflects a complex psychological capacity of the agent, so that defects in control must be defects in the implementation of that capacity. By highlighting the intention-attention nexus, we illuminate agentive control. (Wu 2016: 115)

Let us take a step back before considering this proposal. Proponents of causalism about intentional action often forward causalism as a way of naturalizing action theory—of showing how actions fit within the causal order described by science. Now, if the goal is to naturalize action theory, one

might think that the best way to fill out a causalist theory of action is with reference to cognitive psychology and neuroscience. And one might think that worries about deviant causation are ultimately worries born of ignorance regarding the science. Once we specify the relevant causes, our work is done. Alvin Goldman once endorsed an approach like this (1970: 62). According to Goldman, deviant causation is not philosophically interesting, or philosophically challenging. Discovering the nature of non-deviant causation is simply a matter of elucidating in some detail the causal processes "characteristic of intentional action" (62). Wayne Wu's work on control and non-deviance comes from a perspective similar to Goldman's. Accordingly, he directs attention to empirical details. Wu thinks that if we see how (human) agents actually exercise control, we will understand what goes wrong in cases of deviant causation.

In my view, there is something right in this approach. I have argued that in order to explain non-deviant causation in a particular case, we have to go some way towards modelling the causal system the case concerns. We have to specify causally relevant parameters, and we must have some understanding of how these parameters vary across similar circumstances. Only then does the causal pattern under investigation stand out as non-deviant. One way to think of what Wu is doing is this. Wu is suggesting that in human agents, the causal parameters will necessarily involve certain links between attention and intention.

That is partially an empirical claim, although the claims involving necessity seem to go beyond what science could be expected to show. And ultimately, regarding the full range of action-types and agent-types, I think Wu's conditions cannot be expected to hold. This is just a comment about my prior expectations. There are a wide range of agent-types and action-types. It is doubtful that a particular link between attention and intention, as implemented in human actions—and arguably only in some human actions—would set the standard for non-deviant causation in every case.

Further, in requiring that the relevant attention-mediated coupling be appropriate, both of Wu's conditions appear to assume some account of non-deviant causation. Notice that this troubles Wu's (2011) analysis of intentional action as well: that analysis appears to depend on the assumption that any implementation of a solution to the Many-Many Problem will be non-deviant. But that fails to address the possibility that deviant causation could afflict this part of the process as well. Does Wu's account of agentive control have the resources to specify "appropriate" coupling in a satisfactory way?

It looks like the answer is no. Consider the following two possibilities. On the first possibility, attention-mediated coupling happens in two ways: deviant and non-deviant. If so, then Wu's account fails to account for the non-deviant cases, and so fails to account for non-deviant causation. On the second possibility, attention-mediated coupling is by default appropriate (or non-deviant). So long as a path through the behavioral space is charted, such that the agent's behavior approximates the content of her intention, then the coupling has been appropriate. If that is right, then the word "appropriate" here does no additional work. But that cannot be right. Intention and attention do not work flawlessly—deviant causation cases are precisely cases in which the causal role of intention for behavior is deeply flawed. We have been given no reason to think that the same kinds of flaws will fail to afflict attention.[6] Wu's account fails to explain non-deviant causation.

The problem here is not Wu's interesting work in connecting intention, attention, and typical instances of human action control.[7] The problem is rather that Wu's account, while specifying mechanisms he takes to be important for the implementation of control, cannot guarantee the non-deviant working of these mechanisms. Wu is looking in the wrong direction.

In order to say what it would be for any particular mechanism or causal process to run non-deviantly, we have to move up a level of abstraction. This requires rejecting the thought that there is anything magic or special

[6] I would lodge a similar complaint against Gwen Bradford's (2015) recent, very interesting, discussion of what she calls competent causation. This may not be the same thing as non-deviant causation, but it is close. Bradford deploys the notion in an account of achievement—achievements necessarily involve competent causation. (According to Bradford, an achievement is composed of a process and a product. The process must be difficult, and the product must be competently produced by the process.) What is competent causation? Bradford offers the following:

> Given that A causes some outcome E, A competently causes E via his activity, phi, when A, while phi-ing, has the requisite amount of [Justified True Beliefs] about his phi-ing causing E. The requisite amount is a percentage from the total possible beliefs about the activity, and [Justified True Beliefs] that are about the overall structure of the activity are worth more than discrete beliefs. (Bradford 2015: 79)

This is enough to see that Bradford's account of competent causation is not an account of non-deviant causation. They can deviantly cause behavior as well as intentions. Just tell Davidson's climber case in a way that what causes the trembling hand is not the climber's intention to let go of the rope, but his belief that letting go of the rope is how he will effect the murder. Bradford's competent causation says nothing about this kind of case, but presumably deviantly caused effects are not competently caused effects. Thus Bradford's competent causation requires an ancillary account of non-deviant causation.

[7] This work is of course important. In this connection, see Buehler (2019) for criticisms of Wu and of earlier work of mine on control. As Buehler notes, however, my concern with control here is different from Wu's (and Buehler's). I do not here seek to explain how human action control works. I am concerned with the nature of control and non-deviance.

about the mechanisms that do implement non-deviant causation. And this is just as well. Action is a many-splendored phenomenon, and the set of all action-types is incredibly diverse. So the possible causal pathways from plan-state to successful behavior are likely to be diverse as well.

3.4 Questions, Objections, and Replies

In discussing this view with others, a number of questions and objections arise. I will consider several.

3.4.1 Cases of Easy Success

Davidson's (1980) climber intends to kill another person by letting go of a rope. The intention's acquisition unnerves him, leading him to deviantly let go of the rope. Suppose this person is constructed such that literally anything he does will lead to letting go of the rope. Say his grip is exceedingly loose. So if he breathes in too heavily, or has a slight motor tremor, or looks up towards the top of the cliff, or down towards the bottom, the rope is gone, the other climber plummeting. Doesn't my account implausibly imply that, in such circumstances, whatever the killer does is non-deviantly caused?

Here is an initial reply. The case is a bit under-described. Either the killer conforms to his plan or he doesn't. If the plan was to let go of the rope in a certain way, then his behavior may fail to conform. His success will occur in spite of the low degree of control exercised.

But we can fill out a better version of an easy success case. Say an angel is balanced very precariously on the head of a pin. Say also that this angel has, in most circumstances, little control over their behavior. (Maybe the devil has afflicted this angel with a condition that makes their intentions generate absolutely random movements.) In this case, however, the angel is well situated. Given how precarious their position is on the pin, any movement at all will cause a fall. This angel now intends to fall. And fall they will, although the proximate cause of the falling will be some movement generated by the intention. Say, in addition, that their plan is very simple. The angel intends to fall by causing itself to fall (perhaps by way of some basic action). They do not have a detailed conception of how they will cause themselves to fall. But they are rightly confident that they can do it. (The angel is perhaps analogous to a human who intends to do something simple, via some

opaque causal route that runs through the motor cortex.) And every time they form the intention to fall, they successfully execute their simple plan. They cause themselves to fall. Doesn't my account implausibly imply that, in such circumstances, whatever the angel does is an instance of non-deviant causation?

Yes, that is what my account implies. Although it is not implausible. It is rather a unique situation. Sometimes success is easy to come by. Sometimes plans are easy to execute. What the account requires is that, in addition to the agent's behavior conforming to the agent's plan, the agent's causal pathway is reliably productive of conformity. That is true in every token circumstance for our angel. They bring about success, perhaps, in a million different ways. That's okay. If there are a million ways to get it, choose one.

On to the next one.

3.4.2 Cases of Rare Success

Say that an agent engages in planned behavior that succeeds, across a very big set of circumstances, 4 percent of the time. Can my account explain the difference between hits and misses? In the 4 percent of cases that involve success, the agent is by definition behaving in an abnormal way. How can these be cases of non-deviant causation?

My reply begins by observing that the case is under-described. Is there a relevant difference in the causal pathways taken between hits and misses? If so, we can split the sets to get more information. Remember that the set of circumstances at issue is comprehensive. It will not miss a difference between hits and misses, if the difference is relevant.

So suppose that there is no difference between the causal pathways. The success is thus abnormal, but the causal pathway that leads to it is not. Sometimes the best an agent can do is put herself in a position to succeed, even when success is rare. In such a case the successes and the failures would be examples of non-deviant causation, even though the success qua success would be due to a certain amount of luck. This is a feature of the world, not a problem for my account.

3.4.3 Cases of Strange Success

Here is the basic idea. We imagine an agent's behavioral profile across a well-selected set of circumstances. The profile is consistent, but the causal

routes to success are consistently strange. Say, for example, that every time Kawhi takes a shot, it bounces multiple times on the rim. The bounces are unexpected. But again and again—at least much of the time—the strange bounces lead to a made shot. There are probably better examples. The worry is that my account implies that these unexpected bounces are non-deviant, and this is purportedly implausible.

My reply is, first, that the violation of expectation could be a fault on our end, not on the end of the behavioral profile. But let's assume that the causal routes taken really are strange. What do we mean by this? Perhaps we mean: in nearby sets of circumstances, these causal routes are not reliably productive of success. But why think that is a problem for success in these circumstances? Sometimes a situation calls for causal pathways that are normally not great. When I was a kid they used to say that the way to a man's heart is through his stomach. Maybe the way to a man's heart is not generally through his brain. But in some circumstances, maybe that's just the way. Strange success might look strange, but when the set is well selected, this too is non-deviance.

3.4.4 Cases of Partial Deviance

In such cases the agent displays good control with most of the plan. Perhaps an obstacle is encountered, deviant causation occurs, and the agent moves on without a hitch. In such a case success depends upon a deviant joint in the causal architecture.

This is not an objection against the account. I mention it only to note that it is not an infrequent feature of our world. Intuitions will differ regarding how much deviant causation is consistent with intentional action. I wager that some is—that is, we should not rule out all deviant causation as necessary for some behavior's qualifying as intentional action. What an account of non-deviant causation allows us to do is to say what parts of a plan's execution are controlled, and what parts are due to deviance.

3.5 A Further Criticism

Erasmus Mayr (2011) offers counterexamples to what he calls sensitivity accounts of non-deviant causation. These are accounts that lean on the insight that the causal processes at issue in intentional action should be in some sense sensitive to the content of a relevant intention. My account has

affinity with this basic idea. So it should be informative to see how my account fares against Mayr's counterexamples. After discussing Mayr's counterexamples, I consider one of my own.

Mayr notes that sensitivity accounts come in two varieties. First, we might understand sensitivity as a kind of counterfactual dependence between content and behavior. Second, we might understand sensitivity as involving an explanatory connection between content and behavior (as in, e.g., Schlosser 2007). Mayr focuses on the latter kind of sensitivity account, which makes sense given Christopher Peacocke's (1979) subtle version of it. In the present context, however, this is unfortunate. For my account is closer in spirit to a counterfactual dependence version of a sensitivity view. I maintain that non-deviant causation is not a matter of the explanatory connection between content and behavior, but rather of a level of reliability displayed by the agent in executing some plan.

Even so, in discussion Mayr touches on a view close to my own. Mayr's initial target is Peacocke's view, which requires a one-to-one function between explanandum (such as behavior) and explananans (such as intention). Given this target, his use of the following case (due to Scott Sehon) makes good sense:

> A baseball pitcher tries to move his arm at a speed of 70 mph, and succeeds. Though the fast movement of his arm is an intentional action, we need not assume that the movement would have had a correspondingly greater or smaller velocity had the agent's intention been to move his arm at 69 or 71 mph. Instead, we may well believe that even if the pitcher had intended to move his arm at 69 or 71 mph, his actual speed would have remained the same. So we need not assume that there is a one-to-one function correlating the behavior with the content of the intention in order to classify the pitcher's behavior as an (intentional) action.
>
> (Mayr 2011: 119–20)

Mayr argues that cases of course-grained abilities such as this one undermine any one-to-one correlation between intentions and behavior. For intentions may be much more precise than behavior allows. This strikes me as a problem for a view that depends on one-to-one correlations. But it is no problem for the account I offer. In some cases noise makes relevant differences to success and failure; regarding the exercise of control, and regarding the measurements that underly non-deviant causation, a range representation is thus often more appropriate than a specific number. As a result, course-grained abilities are easily accommodated.

Interestingly, in commenting on this case, Mayr considers and rejects a crude version of the kind of account I offer:

> A defender of the sensitivity strategy might react to such cases by claiming that no precise one-to-one fit between the intention's content and the ensuing behaviour is required for agency, but only a 'sufficiently tight' correlation—a correlation of certain ranges of the intention's content to ranges of behaviour. However, such a move seems to give up the central idea of strongly differential explanation as the key to agency, because this kind of explanation requires one-to-one-correlations. Besides, the move will not help the causalist to escape the problem of antecedential deviance, as limiting cases of 'coarse-grained' abilities to act—'on–off' abilities—make clear.
>
> (Mayr 2011: 120)

Mayr rejects an account that requires "only a sufficiently tight correlation" between intention and behavior for two reasons. First, moving to this kind of account leaves behind strongly differential explanation—that is, explanation of B by A that requires a one-to-one function mapping A to B. Since my account of non-deviant causation does not appeal to strongly differential explanations, this is no problem for me. Second, this kind of account cannot account for limiting cases of course-grained abilities to act.

Why not? Mayr offers this case as an example. We have two agents, A and B. A can wiggle his ears, but "wiggling is an 'on-off' affair with him…he can only 'simply wiggle', but cannot wiggle in specific ways" (120). Agent B cannot wiggle, but in this case intends to anyway and deviantly does so. Mayr comments:

> With regard to the sensitivity of the ensuing bodily motion to the content of the intention, both cases are identical. Both agents' ears would not have wiggled had they failed to have the intention, and no relevant further dependences between bodily motions and intentions hold in either case. So, an adherent of the sensitivity approach must either rule both as cases where the differential explanation criterion is satisfied, and therefore both as cases of actions, or neither. But both solutions are wrong. (2011: 120)

Perhaps this case works against the proponent of strongly differential explanation. But once we have moved away from this requirement it is less convincing. The proponent of my account will say in response that Mayr has failed to consider a large enough set of circumstances. Presumably Agent B's deviantly successful attempt to wiggle his ears is not replicable to

the same degree as Agent A's (otherwise why would Mayr say Agent B "cannot wiggle"?). So, once we consider a wider range of cases, it should become clear that Agent A has some degree of control over ear wiggling, while Agent B has a very low degree of control.

3.6 A Still Further Criticism

Scott Sehon has been a leading critic of causal accounts for some time. Sehon sometimes leans on the problem of deviant causation. In one place, he criticizes an account that is in some ways similar to mine—Jesus Aguilar's. Aguilar stresses the importance of the reliability of whatever mechanisms mediate between mental state and behavior (2012: 10). Sehon is not impressed. His chief problem with this proposal is that, in his view, Aguilar is unable "to show that 'reliable' is being used as a causal term and not as a way of sneaking in . . . teleological concepts" (2016: 104).

How would this usage of reliability sneak in teleological concepts? Sehon's discussion here is odd. He quotes Aguilar's elucidation of reliability, on which it is "a measurable capacity of a process or system to satisfy a given goal," and on which levels of reliability can be assessed in terms of probability of satisfying a goal (Aguilar 2012: 7). So far, so good. He then offers an explication of Aguilar on which it is the agent—not the disjunct "process or system"—that must be reliable. He then complains that the reliability in question "does not specifically concern the causal chain between the reason state and A's B-ing; the reliability condition applies to what the agent would have done in different circumstances and even with different mental states" (2016: 107). And he uses this claim about a difference in mental states to treat Aguilar's notion of reliability as requiring an appeal to sensitivity to reasons, rather than as concerning causation; Aguilar's reliability "amounts to saying that the behavior was goal-directed if it was caused by an appropriate mental state, and, in a variety of other circumstances not necessarily even involving that mental state, the agent's behavior would have been sensitive to her reasons" (107).

This interpretation of Aguilar strikes me as wrong. But I am concerned here to see how this criticism might apply to my own account. I can resist the interpretation of reliability as involving rationalizable behavior—behavior that involves sensitivity to reasons across circumstances that do not include the motivating mental state. For the notions of repeatability and normalcy at

work in my account hold fixed the relevant plan-state. So whether Sehon's criticism of Aguilar is misguided or not, it does not apply to my account.

3.7 Conclusion

According to opponents of causalism about action, the account is bankrupt. Causalists need some specification of control, but without a solution to problems of causal deviance they write empty checks. But the opponents are wrong. Quoth Megan Thee Stallion: money good. Non-deviant causation occurs all the time, and it requires no magic to explain what happens when it does. It is the normal operation of causal pathways at a time and place, where normal is explicated by reference to the background of a well-delineated causal model that captures the time and place.

4
Varietals of Control's Exercise

4.1 Control's Exercise

We are now in position to clean up the formulation of control's exercise offered in chapter 2. Recall that I had to rely on an unanalyzed notion of non-deviance to account for the exercise of control. I no longer have to do so. Indeed, it turns out that in exercising the degree of control that she does, the agent engages in non-deviant causation:

EC. An agent J exercises control in service of (aspects or parts of) a plan-state P to degree D in some token circumstance T, where T is a member of a comprehensive set of circumstances C, if and only if (a) J's behavior in T approximates the representational content of (aspects or parts of) P to degree D, (b) J's behavior in T is within a sufficiently normal range for J, relative to J's behavior across C, (c) J's behavior in T is produced by way of those causal pathways that, when taken, lead, with sufficient frequency, to J's reaching D across C, and (d) P is among the causal influences of J's behavior.

4.1.1 A Difference with Bishop

Notice the difference between this claim and John Bishop's understanding of control:

> There is exercise of control if and only if the causal link from basic intention to matching behaviour is sensitive, in the sense that over a sufficiently wide range of differences, had the agent's intention differed in content, the resulting behaviour would have differed correspondingly.
>
> (Bishop 1989: 150)

This counterfactual understanding of control's exercise does seem to build in an intuitive understanding of control's possession, by reference to ways

The Shape of Agency: Control, Action, Skill, Knowledge. Joshua Shepherd, Oxford University Press (2021).
 DOI: 10.1093/oso/9780198866411.003.0004

the agent would behave when intending different things. My own account of exercise indexes the exercise to a finer set of circumstances, involving the same plan-state. I would agree with a claim that the agent does not possess much control over her behavior if she could only execute her plan as presently constructed—tweaks to the plan are par for the course.

But control over behavior and control as indexed to a plan-state are subtly different. So, imagine a case in which an agent has a plan, and is well equipped to execute it well across many permutations of the circumstance they find themselves in. But there is a catch—the agent could not execute any of many slight differences in the plan. Certainly this agent is precariously poised in one sense. They should not change their plan. But the exercise of control seems possible for them. Like Dom Toretto reminds his team in *Fast and Furious 6*—stick to the plan.

But doesn't Bishop's account capture something important about control? Certainly. Bishop's account requires the possession of control over multiple plan-states (or if you like, multiple permutations of the same plan). We cannot fully understand that notion without the account of control I have provided. But agents would be in trouble if they did not often possess the kind of sensitivity Bishop targets. This kind of sensitivity is what agents develop as they develop skill at various activities—something I discuss in chapter 7.

4.1.2 Exercising Control over Omissions?

One might wonder how this account of control's exercise squares with omissions—things agents do not (omit) to do. Robinson (2014) has argued that "agents can have at least as much control over their omissions as they can over their actions," (439) although it seems he would not extend this to unintentional omissions. What about intentional omissions?

In an earlier paper I offered two conditions on (intended) intentional omissions (Shepherd 2014b). (I remained neutral on whether side-effects could qualify as unintended intentional omissions.) First, the omission should match the representational content of a relevant intention. Second, the match should be explained in part by what the intention (or its acquisition) non-deviantly causes (2014b: 20). I explained the causal work of the intention as follows. In cases of intentional omission, the relevant intention (or its acquisition) non-deviantly causes in the agent "a disposition not to A (where non-deviant causation here involves the intention's making changes

to [the agent's] cognitive and motivational systems that are coherently related to the intention)" (23).

I am not going to assess this account here, although I would probably add a control condition. (Something like: the agent should, in the circumstances at hand, possess control over possibilities relevant to the action's non-occurrence.) The question is whether it makes sense to talk of the agent *exercising* control over her intentional omission. In my view, this kind of talk does not make good sense. Those who view omissions as events, or who view absences as part of the causal order, may disagree. I view omissions as absences, and absences as no part of the causal order. Concordantly, I view talk of dispositions to omit to A as elliptical for dispositions the manifestation of which is explanatorily relevant to A's non-occurrence. And, strictly speaking, I do not require these dispositions to manifest—the causal work of the intention is simply to dispose the agent.

So I would rather say that agents *possess* control over events relevant to the intentional omission (this much is consistent with Robinson's above-quoted claim), but that in intentionally omitting to do something, agents need not actually exercise control.

4.2 Voluntary Control

In the remainder of this chapter, I wish to use this account of control to explicate related notions of philosophical significance.

I begin with the notion of voluntary control. Apparently, this notion is central to more than one philosophical debate. At the intersection of action theory and epistemology, philosophers debate whether we have voluntary control over our beliefs (Chuard and Southwood 2009; McHugh 2014; Helton forthcoming). At the intersection of action theory and ethics, philosophers debate whether moral responsibility requires voluntary control (Adams 1985; Smith 2005; Fritz 2018). And the notion makes guest appearances in areas like the philosophy of psychiatry, where theorists have debated whether, for example, addiction undermines voluntary control (Hyman 2007).

The term is often used as though its content is transparent. Sometimes theorists offer a gloss—usually explicating it further in terms of notions like choice or intention. More often, they offer no gloss at all.

Robert Adams is an exception. Here is what Robert Adams writes about voluntary control:

> To say that something is (directly) within my voluntary control is to say that I would do it (right away) if and only if I (fully) tried or chose or meant to do so, and hence that if I did it I would do it because I tried or chose or meant to do it, and in that sense voluntarily. (Adams 1985: 8)

Preliminarily: the term "directly" is in parentheses because Adams is working with a distinction between direct and indirect voluntary control, where the former involves only basic action, and the latter non-basic. I discuss direct control more below, so I leave the distinction aside for now.

This account has an intuitive ring to it. Adams is clearly tracking some features of voluntary control. But as it stands, this account does not work.

First, this account allows for compelled voluntary control. An on-the-nose alleyway criminal shows you a gun and directs you to hand over the money. Assume you have voluntary control in Adams's sense—if you fully try, you'll hand over the money. This is not, I submit, an instance of voluntary control. The notion of the voluntary is at odds with compulsion. This is easily fixed by adding a non-compulsion condition. I add one to my account.

The second problem is not as easily fixed. It has been recently noticed by Kyle Fritz, who writes that "Not even a professional basketball player can sink a free throw on every attempt, even though she might try, choose, or intend to do so" (2018: 839). That's true. The best free throw shooters in the world miss between 5 and 10 percent of the time, at least in gametime conditions. Given the high degree of control they possess, I find it plausible that they nonetheless exercise voluntary control over their shots, and in good cases, over their makes.

The more general point is that human agents are imperfect. With respect to even the easiest actions—walking, articulating a familiar word, etc.—we sometimes fail. So the best candidates for actions over which we might have voluntary control are not actions over which we have failproof control. If we require failproof control for voluntary control, no one will have it.

If we wish to fix this problem, we need to be clearer about control itself, and how control relates to voluntary control. Voluntary control, I submit, is a sub-type of control.

A question: what is the target of voluntary control? What is voluntary control over? As with control, voluntary control may be over anything that is within the agent's causal remit (cf. Robinson 2014: 439). The notion of behavior in play is broad. The agent may have control over her movements, or mental events, or actions, or over specific components of plans or actions that extend well out into the broader world.

Notice, incidentally, that this account places no temporal restriction on what the agent may voluntarily control. Actions take time, and some take more time than others. Why deny that an agent lacks voluntary control simply because the action she is engaged in takes more than a millisecond? Suppose a nearly omnipotent angel forms a plan. Due to her near-omnipotence, there is vanishingly little chance that she will fail to carry out the plan perfectly. Suppose, for example, she intends to perform some sophisticated mental action that is difficult for minds like ours to understand—some feat of imagination. Suppose the action takes a day to complete. Why claim she lacks voluntary control over this action? I see no good reason. I do, however, see a reason to deny that human agents typically have voluntary control over plans that take long periods to execute. Our powers are limited. The future is unknown. Plans that extend far into the future contain joints and steps that range over very uncertain circumstances. So it sounds odd to claim that a human agent has voluntary control over, for example, acquiring a PhD, when the agent is still an undergraduate.

Since the targets of voluntary control are many, I refer to them with the variable X. Voluntary control's possession over X is the possession of control in circumstances that contain two riders.

First, there should be no impediments to the acquisition of X-relevant plan-states—plan-states that include bringing about X as an end or means of the plan. This addition is due to cases in which agents could exercise control in bringing X about if only they could come to possess plan-states (e.g., intentions) to do so, but in which agents cannot do the latter because of a phobia or some other blocker.

Second, agents should not be in circumstances that have them being compelled to bring about X. Voluntary control's possession requires the absence of compulsion:[1]

Voluntary control's possession

An agent J possesses voluntary control over X across a set of circumstances C if and only if [a] C is well selected, [b] C contains no impediments to the acquisition of X-relevant plan-states, [c] compulsion to acquire or execute X-relevant plan-states is absent across C, [d] J possesses a degree of control with respect to the execution of X-relevant plans, such that J brings X about

[1] What is compulsion? That could take a long time to spell out. So the explication I offer is in this way incomplete. For here I am working with an intuitive, and rough, understanding of compulsion.

a sufficient number of times across instances in which J acquires an X-relevant plan.

If we wish to move from voluntary control's possession to its exercise, we place the agent in an actual circumstance, and we set a comprehensive set of circumstances. That is, we build a causal model of the situation, specifying the circumstance-type, and causal parameters of the agent and environment, against the background of how these parameters vary in the circumstance-type:

Voluntary control's exercise

An agent J exercises voluntary control over X in a token instance T if and only if [a] J possesses voluntary control over X across a comprehensive set of circumstances of which T is a member, [b] J is executing an X-relevant plan in T, [c] J exercises a sufficient degree of control in bringing about X.

A few features of the account deserve special mention.

First, the account is sensitive to the possibility of an agent exercising voluntary control while failing to bring X about. This would occur if the agent fails to exercise a sufficient degree of control, thus failing. In that case we would say the agent exercised voluntary control over a different target—trying to bring X about. The agent could also fail to exercise a sufficient degree of control, and get lucky in bringing X about. That might be a case of deviant causation, and again we would say that the agent exercised voluntary control in trying to bring X about, but not in actually bringing X about.

Second, this account is consistent with the fact that circumstances can be variable, and that this can impact an agent's control. One might have lots of control in some circumstances, and not in others. This may be due to the environment, in different ways. It is more difficult to play basketball outside when it is windy. But difficulty is not the only relevant feature. Some tennis players are much worse on clay than on hardcourt. But clay is not a more difficult surface, just a different one. Control variances may also be due to features of the agent. Lack of sleep, high levels of anxiety, and many other features may influence the agent's control. It is perfectly legitimate to select a set of circumstances that holds fixed some of these features. So the same agent may have a high level of control in one well-selected set of circumstances, and a low level in another.

Third, this account indexes control not only to sets of circumstances, but to plan-states. It makes no comment about how plan-states come about,

beyond the rider that they come about in an uncompelled way. But the acquisition of plan-states could be placed directly under the microscope, by making such an acquisition the potential target of voluntary control. Something like this has been done by philosophers who debate whether decisions—events of intention formation—are ever intentional actions (Mele 2003b; Shepherd 2015a).

4.3 Direct and Indirect Control

Voluntary control is not the same thing as direct control. But the notions can be fruitfully brought into contact.

The most illuminating discussion of direct control is due to Mele (2017a), who notes that "nothing approaching a full account of it exists" (278). Nor does Mele attempt to provide a full account. But he offers significant guidance. First, he remains neutral between two ways of understanding the target of direct control. The target may be an agent's action, or it may be events or states of affairs that are intrinsic to an agent's action. The disjunction need not, to my mind, be exclusive—agents may exercise direct control over both.

Second, Mele offers a plausible condition on direct control: "If S exercises direct control over X, then S does not exercise control over X only by exercising control over something else (or, more precisely, something that does not include X)" (280). This is similar to Adams's idea when discussing direct voluntary control. The notion of directness may be understood at least partially in terms of basic action, where a basic action is an action an agent can perform without performing any other action.

I resist this understanding, however, because in my view control is prior to action. So direct control over behavior is prior to—and would play a role in explaining—basic action. Instead of speaking of basic action, then, I speak only of direct control, understood in terms of Mele's above condition.

Third, Mele distinguishes direct control from complete control (which is, it seems, the same as what I earlier called failproof control). The distinction is useful. Consider the following case, which is a paraphrase of Mele's.

Sol is instructed to press only one of two keys on a keyboard (Q or P). He is to decide which one. Each key has a genuinely indeterministic randomizer such that no matter how hard Sol presses, in some cases the key will fail to fully depress. Mele comments: "So Sol never has complete control over whether he fully depresses the key he has selected and never has

complete control over whether he fully presses the Q key or the P key" (283). I agree. I also agree that in spite of lacking complete control, Sol can exercise direct control in pressing a key.

With that understood, we can distinguish between direct and indirect voluntary control:

> Direct voluntary control's exercise
>
> The agent J exercises direct voluntary control over X in a token instance T if and only if [a] J possesses voluntary control over X across a comprehensive set of circumstances of which T is a member, [b] J exercises sufficient voluntary control in bringing X about, and [c] J does not bring X about by exercising voluntary control over something that does not include X.
>
> Indirect voluntary control's exercise
>
> The agent J exercises indirect voluntary control over X in a token instance T if and only if [a] J possesses voluntary control over X across a comprehensive set of circumstances of which T is a member, [b] J exercises sufficient voluntary control in bringing X about, and [c] J's brings X about by exercising control over something else, something that does not include X.

One interesting upshot of thinking of things in this way is that the distinction between direct and indirect voluntary control may come to seem less interesting.

Consider an agent who desperately wants to come to believe something. This is usually discussed in terms of whether the agent can form a belief "at will," where this is plausibly a stand-in for forming a belief by way of a basic action, or an exercise of direct control. But with an account of voluntary control more fully in view, one may reasonably think that the more important issue is simply how much voluntary control the agent has over the item in question. Why worry if the process takes a few steps as opposed to one, if the control is the same?

Of course there is one reason to worry. For human agents, multi-step processes invite more opportunity for failure. Such processes may thereby correlate with less control. But not always. So agents may well have voluntary control over certain items, such as formations of belief, even if they lack direct voluntary control over them. The issue, in the case of belief, is of course partially empirical. It does not seem to me that we have that much control over our beliefs, although we may in some circumstances (Shepherd 2018b). But there is no guarantee, absent empirical details, that

we will have more control over an item simply because we have direct control over it.

4.4 What Is "up to" an Agent

I wish now to extend these thoughts on voluntary control to a further notion, one that haunts the free will debates. This is the notion of what is "up to" an agent. Many find the following claim at least intuitive: An action A is free only if it was up to the agent whether she A-ed. And yet, as Seth Shabo notes in an illuminating discussion, "the 'is up to . . . whether' locution and its cognates have largely escaped close examination" (2014: 379).

Much discussion of what is up to an agent focuses on the moment of decision (or choice)—the moment of intention formation. This introduces additional complications, for it requires some work to see how events of intention formation could be legitimately considered intentional actions, or exercises of control (see Mele 2003b; Shepherd 2015a). But it is possible to speak of what is up to an agent independently of the moment of decision. It might be up to an agent whether she acts in the way that she does, given some pre-existing intention. Or it might be up to an agent whether she succeeds in A-ing. Perhaps, for example, if an agent had exerted more effort, or paid more attention, she would have succeeded.

Here I wish to make a suggestion. What if we understood what is up to an agent in terms of voluntary control? I would not propose an identification of these notions. For an agent may have some low degree of voluntary control over X. The notion of what is up to an agent seems a bit stronger. What about this:

Up to an agent (simple reading)

X is up to an agent J in some token circumstance T if and only if J possesses a sufficiently high degree of voluntary control over X in a well-selected set of circumstances C of which T is a member.

We understand a "sufficiently high" degree of voluntary control over X as the possession of a degree of control with respect to the execution of X-relevant plans, such that J brings X about a sufficiently high number of times across instances in C in which J acquires an X-relevant plan.

I do not use the term "whether" above, but we could add it in without, it seems to me, doing any violence. So it is up to an agent whether X when the agent is in a circumstance without the compulsion to bring X about, in which there are no impediments to acquisition of an X-relevant plan, and in which the agent possesses a sufficiently high degree of control over the X-relevant aspects of the plan.

This seems at least a plausible explication of what is up to an agent. Thus explicated, however, some philosophers may wonder whether it applies to the usage of "up to" in the free will debates.

Some use the notion of what is up to an agent in a more expansive sense. Shabo (2014), for example, wants to illuminate the kinds of alternative possibilities that are relevant to moral responsibility. He thinks, like many others think, that not every alternative possibility will do. The alternative must be robust. And Shabo understands a robust alternative possibility as one on which it is true to say of the agent that it is up to her whether she realizes it.[2]

One might place an alternative possibilities reading of "up to" claims against the one I have so far advanced. One might claim, for example, that for any action option A, agent J, and time t, it is up to an agent whether A only if it is up to J whether A or B at t, where B is a second option (either an action, or an omission (perhaps an omission to A), depending on how strong one wants the alternative possibilities to be).

This seems too strong to me, at least in the sense that it does not sound incorrect to my ear to claim that X may be up to an agent even in the absence of robust alternative possibilities. Something like this may be true of Frankfurt's cases of volitional necessity (Frankfurt 1988), in which an agent's motivation in some instance is so strong and so clear and so central to how she sees the world, that she can do no other than X-ing at that time.

[2] Shabo also claims that the "up to…whether" locution introduces an intensional context, and that this places a kind of epistemic requirement upon the agent. Consider Tom, who does not know which ticket is the winner, but who wishes to pick the winner. Shabo writes:

> Consider the inference from 'It's up to Tom whether or not he picks the ticket on the left' and 'The ticket on the left is the winning ticket' to 'It's up to Tom whether or not he picks the winning ticket'. This inference is invalid; it follows only that it's up to Tom whether or not he picks what is in fact the winning ticket. (2014: 386)

I agree with the point about intensional contexts. This part is covered by my requirement that the agent's voluntary control be indexed to an aspect of a plan-state. I do not here take on board Shabo's additional epistemic requirements. For discussion, see Shabo (2014) as well as Kittle (2017).

I could be wrong about this. But for now let us admit two readings of "up to" claims. The simple reading does not require robust alternative possibilities. The robust reading does require them. We can understand both in terms of voluntary control:

Up to an agent (robust reading)

X is up to an agent J in some token circumstance T, and at some time t in T, only if the simple reading applies to X at t, and if in addition, the simple reading applies regarding a second option Y for J at t.

Neither reading says anything about determinism. Both incompatibilists and compatibilists can help themselves. One upshot of the discussion, then, is the ability to distinguish between these readings, both of which may be relevant to free will, depending of course on ancillary considerations.

A second upshot is the ability to discern what is being claimed when we link "up to" claims with free will. Consider the following options.

[1] An agent J A-ed freely (at time t) only if the simple reading applied to J's A-ing at t.
[2] An agent J A-ed freely (at time t) only if the robust reading applied to J's A-ing at t.

Both [1] and [2] are to be understood in terms of voluntary control. Both could be strengthened into claims of necessity and sufficiency. The advantage is that we now know what we are assessing when we assess such claims.

I sense that my friends who focus more heavily on free will might find my explication of voluntary control, or of what is "up to" an agent, not quite to the point. I would be interested to hear why. But until then, I offer these explications to them roughly in the spirit in which Salinger offered *Franny and Zooey* (1961) to his editor—that is, the spirit of a one-year-old "urging his luncheon companion to accept a cool lima bean."

4.5 Conclusion

I have devoted this chapter and chapters 2 and 3 to development of notions that constitute key building blocks of agency. These notions are also critical for an understanding of intentional action. Chapter 5 is about intentional action. I offer an account.

5
Intentional Action

5.1 Introduction

A causalist account of intentional action gives causation a necessary role to play. The caricature has it that an agent acts intentionally only if her behavior is caused in the right way by a relevant mental state (or that state's) acquisition. What more is required is a matter of much debate—there are many varieties of causalist account.[1] A shared central flaw is the lingering worry regarding deviant causation.

We should be feeling okay. We solved that problem.

In this chapter, I leverage accounts of control and non-deviance to offer a new account of intentional action.

I say new. But the account has much in common with, and is inspired by, earlier causalist accounts (e.g., Goldman 1970; Brand 1984). Some of this will become apparent in the next section. I begin to explicate my own account in conversation with the account offered by Mele and Moser (1994).

5.2 Intentional Action: Mele and Moser

Mele and Moser's account is still, twenty-five years on, the most detailed and to my mind convincing analysis of intentional action. The paper that sets out the analysis has been influential for my own thinking on intentional action. Further, some of the merits of my account are more easily seen via engagement with this account. For an account of control helps to undergird some of the points Mele and Moser make. Finally, the differences that emerge may prove instructive for those invested in the project of offering a causalist account (or explication) of intentional action. I do not doubt many

[1] There are also many discussions of the many varieties of causalist account. For good introductions, see O'Brien (2015), or essays in Aguilar and Buckareff (2010), or in D'Oro and Sandis (2013).

The Shape of Agency: Control, Action, Skill, Knowledge. Joshua Shepherd, Oxford University Press (2021).
 DOI: 10.1093/oso/9780198866411.003.0005

will disagree with some of the judgments my account embeds. The discussion will serve at least to make my judgments plain and to offer readers the chance to explicitly diverge at various choice points.

Here is the final analysis Mele and Moser (1994: 63) offer:

> Necessarily, an agent, S, intentionally performs an action, A, at a time, t, if and only if:
> (i) at t, S A-s and her A-ing is an action;
> (ii) at t, S suitably follows—hence, is suitably guided by—an intention-embedded plan, P, of hers in A-ing;
> (iii) at the time of S's actual involvement in A-ing at t, the process indicated with significantly preponderant probability by S's on balance evidence at t as being at least partly constitutive of her A-ing at t does not diverge significantly from the process that is in fact constitutive of her A-ing at t; or (b) S's A-ing at t manifests a suitably reliable skill of S's in A-ing in the way S A-s at t; and
> (iv) the route to A-ing that S follows in executing her action plan, P, at t is, under S's current circumstances, a suitably predictively reliable means of S's A-ing at t, and the predictive reliability of that means depends appropriately on S's having suitably reliable control over whether, given that she acts with A-ing as a goal, she succeeds in A-ing at t.

Mele and Moser arrive at this analysis via consideration of a wide range of puzzling and difficult cases. In order to illuminate the zones of agreement and disagreement, I will offer commentary on aspects of each of their four conditions, with discussion of some relevant cases.

First, this is an analysis of what it is to intentionally perform an action. As condition (i) makes clear, this leaves unanalyzed the notion of action. That is something I would rather avoid. I do not speak, then, of actions that are intentional. Instead I speak of behavior that amounts to intentional action. The notion of behavior is, in my view, exceedingly broad. To call some event a bit of an agent's behavior is not to say the agent has done anything in any significant sense. Anything within the scope of an agent's causal potency could be construed as her behavior. The interesting thing is that some bits of an agent's behavior qualify as intentional action.

Condition (ii) speaks of an agent's suitable following of a plan. Mele and Moser do not offer an account of following. But they draw a connection between following and guidance by a plan, and between guidance and deviant causation cases. This is plausibly the identical, or at least a similar, connection

to the one between control and non-deviance. It is thus not implausible to offer an account of following, and hence guidance, in terms of control:

Plan Following. An agent J suitably follows—hence, is suitably guided by—a plan P (or certain aspects of P) when J exercises control in bringing her behavior to approximate (to some sufficient degree) the representational content of P (or certain aspects of P).

Condition (iii) is disjunctive. The first part mentions an agent's evidence. The second part mentions an agent's suitably reliable skill. I will discuss the second part first. Mele and Moser do not offer a detailed discussion of skill, but do offer the following useful characterization:

> At time t, an agent, S, is skilled to some non-zero degree at A-ing in manner M if and only if S, at t, has a propensity to A in manner M, given that S has a corresponding intention in a situation suitably accommodating for S's A-ing in manner M, and that propensity exceeds any provided by mere chance given such an intention and situation. (Mele and Moser 1994: 57)

This characterization appears congenial to my account of control. In some respects, my account seems simply a more involved working out of various details—what it is for a situation to be suitably accommodating, for example, and how best to think of degrees of control beyond zero. But one might pause over their language of a propensity that exceeds any provided by mere chance. Mele and Moser comment: "To have a high propensity to roll non-doubles with a pair of fair dice (simply by rolling the dice) when one has a corresponding intention is not to be skilled at rolling non-doubles. Such a propensity is predicted by mere chance" (57).

This claim has the potential to create a problem for my account of control. To see why, consider the following case from Mele and Moser, who credit Michael Costa:

> Mike, a normal person, is playing a game with a pair of fair dice. He will win $20 on his next roll if and only if he throws something other than "boxcars" (two sixes). Mike, wanting to win, has a simple plan: he will throw a non-boxcar roll and win the money. Mike realizes that there is a slight chance that he will roll boxcars, but this does not threaten his plan. As it happens, he throws a seven. (Mele and Moser 1994: 62)

Given that the dice are fair, one can imagine that Mike has very close to a 92 percent probability of rolling non-boxcars. On my account then, doesn't Mike have a high degree of control with respect to intentions to roll

non-boxcars? And isn't this an odd thing to think? After offering the case, for example, Mele and Moser make this comment: "Mike lacks a kind of control over the dice required for his intentionally throwing a non-boxcar roll" (62). Intuitions may side with Mele and Moser here, suggesting that my account does violence to the notion of "control over."

Mele and Moser's decision here is to exclude cases like these—cases of what they call statistically reliable luck—from their account of intentional action. In so doing, they explicitly contrast control with statistically reliable luck. And in their condition (iv), they claim that an agent's following a "suitably predictively reliable means" of A-ing depends upon her "having suitably reliable control over" whether she succeeds.

One might wonder, however, what reliable control amounts to if it is not simply a high propensity to succeed. Regarding Mike and the dice, Mele and Moser comment that "Mike has no control over which sides land face up. He thus has no control over whether, given his throwing the dice, he throws a non-boxcar roll (as opposed to a boxcar roll)" (62). This suggests a requirement of something like a power of influence regarding certain aspects of one's behavior. How to think of this power of influence? A power to determine every aspect of one's behavior is far too strong.

Some have thought of this power in terms of positively shifting the probability of success (Ross 1978; Costa 1986). But Joshua Gert (2004) has argued convincingly against this kind of view. For one can construct cases of obvious intentional action on which agents lack the ability to shift the probability of success—just imagine an agent triggering a device that has a 92 percent probability of killing another person.

I doubt there is a good way to finesse the intuition behind dice roll cases.[2] But let's try. Consider Nash, who is very good at shooting free throws. Nash makes 92 percent of the ones he shoots. One might think that the 8 percent are in some sense Nash's fault. In those cases, Nash failed to exercise his abilities properly. He could have made the shot, but he didn't. Ask Nash about this, however, and he might disagree. Nash might admit that some subset of the misses are his fault. He might say this:

Nash's Speech. Look, sometimes I know why I missed. And sometimes I can intervene to change the probabilities in my favor. I can intervene on the

[2] Gert suggests that dice roll cases are not cases of intentional action because the agent's plan "does not include a representation of the fact that the dice come up [non-boxcars] because she rolls them" (Gert 2004: 155). But it seems clear an agent's plan could include such a representation—indeed, in Mele and Moser's case the agent throws the dice in order to roll non-boxcars in order to win money.

processes constitutive of my behavior. So, in some cases I think I might have done better if I had taken a deep breath, focused, tried harder, removed distractions, visualized my plan more clearly, or whatever. But even having done these things, sometimes I get to a place at which I have done everything I can do, and the probabilities are what they are, and they are short of 1. Sometimes I just miss. I'm good, but I'm not perfect. I think that my missed free throws, sometimes, are not the result of any failure on my part—at least no failure I can do anything about. I think sometimes my miss is the result of a normally very reliable process misfiring. Or, possibly, the process underlying my shots has margins that occasionally fall outside what is required for success.

I'm on Nash's side. I think reliability is the most we can require. If so, I see no obvious reason to exclude processes like those involved in the dice roll. After all, if Nash takes up a dice-throwing game that requires frequent non-boxcar rolls, he might plausibly come to view his situation as similar to the free throw situation. It takes no practice for Nash to be just as good at non-boxcar rolls as free throws, of course, but that is an inessential detail.

I admit, mind you, the unintuitive nature of this response in Mike's case. The rigidity of the probabilities, the opaqueness of the details to the agent that initiated them, and the agent's inability to intervene on these processes all play with our intuitions, tempting us to think of what is happening as unusual or somehow deeply non-agentive. But I submit that most intentional action includes aspects that have rigid probabilities, that have opaque details, and that do not permit intervention. This is something we should simply accept. If we do, then we can understand "suitably reliable skill" in condition (iii) simply as sufficient control, and we need not separate the skill of condition (iii) from the control of condition (iv).

The first part of condition (iii) mentions an agent's evidence. Why think an agent's evidence is relevant to whether she has intentionally acted? Mele and Moser offer this case:

> Ann works as an admissions supervisor at an orchestra hall. She gives red admission slips to women and blue slips to men. Since the orchestral concerts are formal occasions, Ann always wears white gloves while working. Without examining the slips, she efficiently gives the men blue slips from her left pocket, and the women red slips from her right pocket. All goes as planned until Ann hands Harry an admission slip from her left pocket. Unbeknown to all concerned, Ann's admission slip for Harry is actually a white piece of litmus paper that instantaneously turns blue when touched

> by a human hand. Harry's hand turns Ann's white litmus paper blue. Consequently, Ann gives Harry an appropriate admission slip.
>
> (Mele and Moser 1994: 52)

What did Ann do wrong? Mele and Moser diagnose the source of the coincidence as stemming from a divergence between the route Ann took and her evidence regarding that route. Now, I agree that an agent's evidence can be relevant to their behavior. But I suggest that this relevance runs through the notion of control. Insofar as agents rely on their evidence, they need good evidence. Misleading evidence will tend to lead to failure, rather than success. The way it does so is by infecting an agent's plan for action. Agents acting on bad evidence are agents acting on a bad plan—a plan that embeds bad expectations, that gives improper directions. Bad plans succeed less often than good plans. That is, bad plans lower an agent's degree of control. So I suggest we assimilate the condition regarding evidence into a broader condition regarding plan quality and control.

Condition (iv) requires further attention. Mele and Moser maintain that the route the agent takes in executing her plan must be a suitably predictively reliable means to success. They link this requirement of predictive reliability to the agent's control, where the control operative here is a notion I have challenged. But there may nonetheless be something in this notion of predictive reliability. They explicate it as follows:

> Intentional action, on the intended interpretation . . . requires that—given just (a) S's suitably reliable nonmisleading evidence [that is, evidence devoid of false propositions] concerning whether she will A at t in her present circumstances, and (b) knowledge concerning what sort of reliable skill, if any, S has with respect to A-ing at t in her present circumstances—a conceiver who understands all the relevant concepts (sufficiently to wield them in any prediction involving just those concepts) could reasonably predict that (the route followed in) S's attempted execution at t of an intention incorporating the relevant action plan P will result in her A-ing at t. (Mele and Moser 1994: 60)

I have challenged the necessity of an agent's evidence for her ability to act intentionally, except as it contributes to the quality of an agent's plan, and thereby her control. Beyond this, whether an outside observer who has relevant knowledge could use the agent's route to predict whether she will succeed can be understood as a way of honing in on the agent's own success-rate

in behaving as she does. That is, we can understand the requirement of predictive reliability in terms of the possession of control.

5.3 Intentional Action

As I said above, I already found Mele and Moser's proposed analysis of intentional action compelling. It is possible to undergird elements of that analysis with the account of control I have developed, deepening our understanding of the nature of notions like plan following, and affording a response to the most long-standing complaint against causalism about intentional action, namely, that deviant causation undermines any causalist account. Further, if one sides with me on certain choice points in the analytical tree, one may be disposed to prefer a more parsimonious account of intentional action.

Before I offer that account, additional commentary is required. As I noted, I am not offering an account of action that qualifies as intentional action. I cannot help myself to an unanalyzed notion of action, nor of the content of the action at hand—that is, of an A-ing. I am accounting for behavior that qualifies as intentional action. How would some bit of behavior come to do so?

By entering into relationships of sufficient approximation, or resemblance, with other key relata. The relata are the agent's behavior, (aspects of) the agent's relevant psychological state(s), and the action-type of which the action under inspection is a token. The idea is that behavior qualifies as intentional action only if the behavior sufficiently approximates an aspect of the relevant plan, as well as the action-type of which it qualifies as a token.

I am not, note, putting much weight on the notion of an action-type. This is a way of capturing a couple of facts. First, there are many ways to describe an agent's behavior as an intentional action. Sometimes some bit of behavior can be described as multiple different intentional actions. Sometimes it doesn't much matter—one or two or three different descriptions could be equally accurate. Second, it is arguably possible to describe some bit of behavior as an intentional action even if the agent was not conceptualizing the behavior as such (see the discussion below, in section 5.5.2). Here we impose an action-type on the behavior. If we are to do so, the imposition should be roughly accurate, and I assume we have a rough way of accurately deploying action descriptions.

Consider an uncontroversial case. LeBron intends to go to Los Angeles. He has a plan for doing so. It involves catching a flight from Cleveland. Suppose that LeBron's behavior is caused by the acquisition and persistence of this intention, and that LeBron exercises sufficient control over his behavior—he tells his driver which airport to go to, he finds the right plane. LeBron has intentionally done several things—told his driver where to go, caught a flight from Cleveland, gone to Los Angeles. Why? Consider:

It is important that LeBron's behavior conformed sufficiently to his plan, or to relevant aspects of it. If LeBron told the driver the wrong airport, and found the wrong plane, but strong gusts sent the plane off course and to Los Angeles, we might well think LeBron's arriving in Los Angeles unintentional.

It is important that LeBron's behavior conformed sufficiently to the action-type(s) of which the actions under consideration are tokens. The action-type going to Los Angeles is broad, but LeBron's behavior falls under it as a token. So, too, with LeBron's catching a flight. If LeBron announced an intention to go to Los Angeles, after which insidious agents gave him spiked milk, and transported him sleeping to the city of angels, we might have a question about whether this is a valid way to intentionally go to Los Angeles.

Finally, it is important that LeBron's plan was well-constructed in the sense that if LeBron followed it—if he exercised a sufficient degree of control in bringing behavior to approximate the plan—he would indeed bring behavior to approximate the relevant action-type.

So we arrive at the following proposal:

Necessarily, an agent, J, intentionally performs an action, A, at a time, t, if and only if:

(i) at t, J's behavior B is caused by a relevant plan-state P of J's (or its acquisition and persistence), or a package of such states, P*;

(ii) B sufficiently approximates a relevant aspect of the plan;

(iii) B approximates a relevant aspect of J's plan because J exercised control over B;

(iv) B sufficiently approximates the action-type of which A is a token;

(v) J has a good plan: J's following the plan, by exercising a certain degree of control in bringing B to sufficiently approximate the plan, is itself a reliable method, in the relevant comprehensive set of circumstances, for sufficiently approximating the action-type of which A is a token.

5.4 The Account Paraphrased

That account of intentional action took most of the book until now to develop. It contains terms the full explication of which date back to chapter 2, and include parts of later chapters. So it may be useful to state the ideas behind the account in plain language.

Intentional action is, in essence, the exercise of a sufficient degree of control in bringing behavior to approximate a good plan. It is composed of the following elements. First, a plan for action. Second, that the plan be good—that following the plan be a good way to satisfy goals embedded in the plan. Third, control over behaviors required by the plan. Fourth, a causal pathway that includes the psychological state(s) that represent the plan. Fifth, a relationship of approximation between controlled behavior and whatever action-type we use to classify the controlled behavior as intentional action.

5.5 Ancillary Issues

I may be able to further illuminate this account's commitments and (de) merits by discussing cases often pressed for or against one or another account of intentional action.

5.5.1 Side-effects

Gilbert Harman (1976) offers the case of a sniper who intends to kill a soldier, and who is aware—without intending—that in firing his rifle he will alert the enemy. He fires the rifle, kills the soldier, and alerts the enemy. Does he alert the enemy intentionally?

Some have said yes; many no. Mele and Moser argues as follows:

> Since the sniper does not unknowingly, inadvertently, or accidentally alert the enemy, it is natural to insist that he does not unintentionally alert the enemy. Such insistence does not entail, however, that the sniper intentionally alerts the enemy. There is a middle ground between A-ing intentionally and A-ing unintentionally. We locate "side-effect actions" of the kind in question on that ground. Insofar as such actions are not done unknowingly, inadvertently, or accidentally, they are not unintentional. Insofar as the agent is not aiming at the performance of these actions, either as ends

> or as means to (or constituents of) ends, they are not intentional either. We shall say that they are nonintentional. The ordinary co cept of intentional action requires the agent of an intentional A-ing to be aiming at A-ing. (Mele and Moser 1994: 45)

My account would not accord intentionality to the sniper's alerting the enemy. That was no part of the sniper's plan. I agree with Mele and Moser and many others about that. Whether a notion of nonintentional action is available is a separate matter, outside the parameters of my account. One might offer an account of such a notion, making it derivative on intentional action. I have no plan to do so. I would be just as happy saying the sniper knowingly caused the alerting of the enemy, without saying anything about action. That might give us all we could need (e.g., we can blame agents for knowingly causing things as well as for intentionally doing things). I might even add that an agent capable of knowingly causing A is an agent capable of intentionally A-ing. That does not render the notions equivalent, but it does indicate an important connection. If an agent knows that something is within her causal remit, then she can include that thing in her plan as a means or an end. And if she plans to bring it about, and succeeds in the way my account indicates, then it qualifies as an intentional action.

5.5.2 The Simple View and the Single Phenomenon View

One ongoing debate in action theory concerns a question about the relation between intentional action and intention (or, more broadly, the agent's plan-state) (see, e.g., Bratman 1984; Adams 1986; McCann 1987; Nadelhoffer 2006; McCann 2011; Wasserman 2011; Amaya 2018). Here are the two views:

Simple View (SV). An agent cannot intentionally A unless she intends to A.

Single Phenomenon View (SPV). In order to intentionally A an agent must execute some relevant intention, even if not an intention to A.

While SV presents initially as intuitive, it has been undermined by a range of cases, as well as by experimental philosophy suggesting that laypeople's philosophical judgments are better explained by SPV (Nadelhoffer 2006). For cases, see Bratman (1984), Di Nucci (2009), or this case from Mele:

> Alice is mowing her lawn—intentionally, of course. In the process, she has taken many steps. It would be strange—and, I believe, mistaken—to maintain that her taking her next routine step while pushing the mower is not an intentional action. But there is no need to suppose that Alice has an intention specifically to take that step. Given that she intends to mow her lawn at the time, is a proficient mower, encounters no obstacles requiring alteration of her gait, and so on, her mowing intention can do its work without her having a series of intentions corresponding to each routine step. (Mele 2005: 150)

I side with proponents of the SPV. The present point is that my account of intentional action permits latitude on the matter. I require sufficient approximation between plan-state and behavior. Depending on how one reads the SV vs. SPV debate, one might want sufficient approximation to amount to the Simple View or the Single Phenomenon View. If one feels the pull of Mele's above case, one might think behavior need not be exactly in line with the plan. One's plan is to mow the lawn, and that plan includes a representation of pushing the mower around the yard, but not a representation of taking each specific step. One might nonetheless judge that taking a specific step sufficiently approximates the general plan of pushing the mower around the yard.

Alternatively, one might feel strict about the approximation to the plan. Santiago Amaya appeals to the role intentions play in setting evaluative standards for action, and argues that the SPV fails to adequately explain the sense in which certain failures to achieve one's goals are really mistakes. Regarding cases like Mele's, he has this to say:

> Intending to achieve a certain goal normally involves performing some subsidiary actions. Insofar as one intends some of them as implementations of one's goals, these will be intentional . . . At the same time, it is often the case that one is perfectly indifferent about many of the subsidiary actions one performs in the pursuit one one's goal, even when those actions serve the purpose of helping one achieve that goal. These actions seem better described as non-intentional. As long as one manages to act as intended, not performing one of them would not count as an executive mistake. (Amaya 2018: 1784)

It's a merit of my account that it permits further specification, in terms of the SV or the SPV, along whatever lines the debate ultimately takes.

5.5.3 Senseless Movements

Some will wonder how this account squares up to senseless movements. Brian O'Shaughnessy (1973) saw in such actions—he lists "moving one's toes as one sits reading a book" as an example—a middle ground between intentional action and non-action:

> Does nothing lie between being a corpse-like graven image and a vehicle for reason? How else but as action is one to characterize the making of these movements, and to what but the person is one to attribute them? One can hardly telescope them into mere spasms on the part of the toes. And does one not suddenly become aware of doing them? Yet it is not as if they were intentionally senseless, as it were the 'small talk' of bodily movement, for they are not chosen . . . these trifling actions can express nothing more distinctive or more mental in our inner life than vague unease; and anyhow they are an afterthought in the scheme of things. They relate to standard examples of action somewhat as do objects that are mere lumps of stuff, say rough diamonds, to objects that are both lumps of stuff and more, e.g., artefacts, natural kinds. (But whereas all matter might have been in the form of mere chunks, these could not be the only examples of action in the universe.) Such senseless 'raw' acts are not amenable to interpretation, not even the interpretation of having no interpretation, and that is why they are not intentional under any description. Excluding them from the class of all actions would be roughly akin to excluding gold nuggets from the class of material objects.
>
> (O'Shaughnessy 1973: 366–7, fn. 2)

I reject O'Shaughnessy's claim that we can consider senseless movements neither intentional nor non-action. They are one or the other, and it will depend upon the case.[3] O'Shaughnessy hits on the right question when he asks whether we should attribute them to the person. Or rather, he is close

[3] John Hyman offers a longer list: "automatic reactions, such as ducking or drawing back one's head to avoid a blow, or making an involuntary adjustment to one's posture to maintain balance; some kinds of habitual action, including verbal tics such as echolalia (the automatic repetition of words and phrases spoken by a person one is conversing with) or interspersing speech with words or phrases like 'you know'; some kinds of uncontrolled action done in abnormal or pathological states of mind, such as panic or psychosis; unconscious action such as murmuring in one's sleep; and the spontaneous expression of emotion in facial expressions, vocalizations, and gestures, such as smiling, scowling, pouting, shrugging, and laughing, or crying out with pleasure or pain" (Hyman 2015: 50–1).

to the right question. I think the issue is whether such movements stem from central psychological states that we should attribute to the agent. Some idle movements qualify—as I write I twitch quite a bit, and it is possible to find in these twitches, I would think, a connection to urges to move based in vague unease or physical discomfort or some energy associated with the sentence at hand. Such behavior is intentional action, albeit a trifling sort. Other twitches and idle movements may be due to noise in my motor cortex, to reflex arcs further down the neurological hierarchy. There is less pressure to attribute these movements to me. They are then not action at all.

In agents like humans, who possess a complex neurological architecture supporting behavior, there may be in-between cases. We know that there are levels in a processing hierarchy supporting human action control (Fridland 2014; Shepherd 2015b; Christensen et al. 2016). We know that at times the agent's intentions can be at cross purposes with the somewhat independently structured processing supporting motor execution and sensorimotor adaptation (Day et al. 2016; Mylopoulos and Pacherie 2017; Shepherd 2019). It is an empirical question just how and just whether lower-level states like motor commands should be included within the agent's broader plan-state. Perhaps some should; perhaps some should not. I do not wish to commit here; the science is ongoing. The point is simply that if some movement follows in the right way from an agent's plan-state, and approximates some aspect of the plan sufficiently, we should think of it as intentional action. If the movement follows from something else—some state or some event that is outside the scope of any plan of the agent's—then it is not action at all.

5.5.4 Belief and Knowledge

My account of intentional action says nothing explicit about knowledge, and little about belief. One might, of course, think of the agent's plan-state or package of plan-states as involving or requiring belief. I remain neutral on that. One might also think of the execution of intentional action as necessarily involving knowledge how. On an anti-intellectualist view, on

Walking through each item on the list would be tedious. But I don't see any deeply problematic cases here. Some of our plan-states (i.e., ducking) are acquired very rapidly, and some involuntarily. Some cases tempt intuitions. But I do not think murmuring in one's sleep is an intentional action—it is an action-like effect of something one may be doing intentionally in one's sleep.

which knowledge how is an ability, I have no deep problem with that. Here I wish to discuss a different sort of view, regarding the importance of belief, and arguably knowledge that, for intentional action.

Intellectualism about knowledge how is the view that 'for one to know how to φ, for some task of φ-ing, is for one to know a proposition that truly answers the practical question "How could one φ?"' (Pavese 2016a: 650). Carlotta Pavese argues that knowledge how to A is necessary for possession of the ability to intentionally A (Pavese 2016a, 2016b, 2018).

Pavese argues as follows:

1] An agent J's success at intentionally A-ing requires that J have a true belief about how to A.[4]

2] Given 1], J cannot be in a position to intentionally A unless J has a true belief about how to A.

3] J cannot have the ability to intentionally A unless J is in a position to intentionally A.

C] The ability to intentionally A requires a true belief about how to A.

Ultimately, Pavese thinks that more than true belief is required for intentional success. Later in her (2018) she offers arguments based on cases as well as claims about what is required for satisfactory explanations of intentional success that conclude that intentional success at A-ing requires not just true belief, but knowledge of how to A. It follows that the ability to intentionally A requires knowledge how to A—a conclusion that affords the analysis described above. These latter arguments are interesting. Before discussing them, I discuss a potentially more basic issue.

It is clear that without premise 1], Pavese's argument will not go through. Pavese supports this premise initially by reference to the wide endorsement this premise receives amongst action theorists.[5] It is certainly true that many have endorsed a belief requirement on intending, and by extension on intentional action.

But what kind of belief is required? One way to get a belief requirement on intentional action more or less for free would be to endorse cognitivism

[4] As Pavese puts the thought, the intentionality aspect of the ability to intentionally A "reduces to a doxastic attitude" (Pavese 2018: 4).

[5] Among others, she cites Goldman (1970); Brand (1984); Harman (1976); Thalberg (1984); Ginet (1990); Mele (1992); Mele and Moser (1994), though we will see that Mele (1992) does not endorse it, nor, on my reading, does Mele and Moser (1994).

about intention. According to cognitivists, an intention just is a sort of belief. There is something to be said for this view (see Marûsić and Schwenkler 2018). But I will remain neutral on it for present purposes. In any case, it remains a controversial view in the philosophy of action, and Pavese explicitly states she is not here assuming cognitivism.[6]

The kind of belief requirement typically endorsed in the philosophy of action is motivated by the following consideration. An agent cannot be said to be genuinely committed to A-ing in the way an intention to A would require if the agent believes that she is likely to fail. So agents should believe that they are likely to succeed. The belief Pavese offers is in line with this: "If one successfully intentionally φs at t, then at t one believes, for some way ψ of φ-ing, that one is sufficiently likely to φ by ψ-ing" (2018: 6).

The viability of this belief requirement on intentional action depends on whether such a belief is really important for explaining intentional success. I am dubious that it is. Consider the following case due to Al Mele:

> In numerous instances of intentional A-ing, the question whether we will succeed in A-ing would appear to be the furthest thing from our minds. Yet in many such cases, the claim that we intend to A is unproblematic. A few minutes ago, while I was typing, I heard a knock at my office door. As is my habit, I answered the knock. I answered it intentionally, and the suggestion that I intended to answer it would encounter little resistance. Yet I do not remember having a belief at the time to the effect that I (probably) would answer the knock. Indeed, I seem to recall that I had no such conscious belief. Moreover, given that my intention . . . can do its work without the assistance of a belief—conscious or unconscious—that I (probably) would answer the door, there is no apparent need to postulate the existence of this belief. (Mele 1992: 147)

Kieran Setiya offers a similar case and verdict:

> I need no more than the intention to clench my fist, and the disposition to do so under the guidance of that intention, in order to clench my fist intentionally as a basic action . . . What [the intellectualist] owes us . . . is an account of what it is to have the relevant intention and disposition on which they entail the belief that I could clench my fist by intending to do

[6] Could one use cognitivism to get to an interesting conclusion regarding the relationship between knowledge how to A and the ability to intentionally A? Probably so, in my view.

> so. The problem here is not simply one of omission, that there is more to say, but that it is difficult to see how any plausible account of intention or guidance would entail this belief. (Setiya 2012: 294)

Such cases motivate a negative belief criteria on intending (and thus, on intentional action): "J intends (at time t) to A (at time t') only if he does not believe (at t) that he (probably) will not A (at t')" (Mele 1992: 147). (We could expand this to plan-states generally.) This criteria would spell trouble for the first (and a key) premise in Pavese's argument. Should we accept it? Beyond the intuitive nature of the cases, the primary consideration Mele and Setiya advance is that the belief is doing no relevant work in explaining the intentional success.

Pavese, however, advances arguments in favor of the explanatory relevance of a belief requirement. I will consider two. The first is offered in response to Setiya. Pavese distinguishes between two notions of intentionality. Intentionality-minus is "a property that an action-token possesses just in case (1) the action is intended and (2) it is caused in the right non-deviant way by that intention" (Pavese 2018: 11). Intentionality-plus adds to this a belief requirement. Pavese then argues that basic actions should only qualify as intentional if they meet intentionality-plus. Her reasoning is as follows:

> Basic actions for a subject at a time must be actions a subject is, at that time, able to plan from, if only the subject were to engage in a little [bit] of reflective effort. But in a circumstance where I explicitly (that is, upon considering the question) did not believe that I can perform an action (such as clenching my fist) by merely intending to, I would not be in position to plan a complex action that has clenching my fist as its part.
>
> (Pavese 2018: 11–12)

There is something to be said for Pavese's claim about being able to plan from actions. But in any case this reasoning does not work. It is consistent with a negative belief requirement on intentional action that basic actions be available for planning. So the claim does not uniquely support a positive belief requirement.

The second argument of Pavese's I wish to consider she calls the argument from verbal feedback. The argument involves two moves. First, we acquire knowledge how through verbal feedback and through trial and error. Second, if we did not possess standing belief states about how to A

that verbal feedback manipulated, then explaining how verbal feedback works would be difficult. Pavese writes:

> Suppose my tennis teacher tells me that I am not holding the racket correctly. It seems natural that their feedback will result in revising my beliefs about how to hold the racket. More generally, revising one's belief about how to perform an action seems a natural way in which know-how can be acquired and improved through verbal feedback. If Belief requirement on know-how is correct, this aspect of know-how can be easily captured. By contrast, if know-how did not require a standing belief state from one occasion to another, the acquisition of know-how through verbal feedback and trial and error could not be a matter of revising one's belief.
>
> (Pavese 2018: 8)

In response, we can grant that the acquisition of knowledge how can be influenced by changes in beliefs. This may happen via many mechanisms, including the changing of the trial and error process that may be more critical for knowledge how acquisition. But the influence of belief on knowledge how acquisition—or on acquisition of new abilities to intentionally A—does not entail the necessity of belief for such acquisition in every case. If there are cases where know how, or the ability to intentionally do a new thing, are acquired without verbal feedback, this places the argument from verbal feedback in doubt. And it does seem like there are many such cases. When a young Björn Borg spends hours volleying a tennis ball to himself off of a garage door, he seems to be building his base of abilities to intentionally hit the ball in various ways. But there is no verbal feedback offered.

The main point of contention here concerns the relevance of beliefs about how to do something to the explanation of intentional successes. I am claiming beliefs are not necessary. A different way to test this claim may be useful for people whose intuitions are stuck in the middle.

Let us consider a case of contrastive explanation of intentional success. The case of Dan and Dave:

Dan and Dave are sprinters who have trained together since childhood. Their practices are identical, and in learning their technique their coach is careful to teach them the exact same methods—for he believes his methods are uniquely the best for successful sprinting. Over time Dan and Dave internalize these techniques so well that they can coach each other, which

they do by reminding one another of the same mantras, and by stressing the appropriate ways to conduct each training exercise. Indeed, by the time their coach is ready to allow Dan and Dave to compete, the propositions that they explicitly believe about sprinting and how to do it are exactly the same. Now, as it happens, Dan is slightly faster than Dave, although as their coach does not allow them to race each other, neither knows this. Indeed, neither knows just how fast they are—for their coach does not release their times. But Dan is able to reliably run sub-10-second 100-meter dashes, while Dave cannot break 10 seconds. When Dan and Dave run their sprints, they intend only to run fast, utilizing the techniques they have learned.

Here is a verdict on this case that I find intuitive. When Dan runs with the intention to run a fast 100-meter dash, he not only intentionally runs fast, he intentionally runs a sub-10-second 100-meter dash (9.92 or whatever). (Accepting this intuition pushes one to the SPV, but that is hardly a cost.) When Dave runs with the same type of intention, he not only intentionally runs fast, he intentionally runs a near-10-second 100-meter dash (10.06 or whatever).

Now, both Dan and Dave may or may not have beliefs about their likelihood of success. This case is not an illustration of the negative belief requirement on intending or intentional action. The point is that Dan and Dave have different abilities. Dan is able to intentionally run sub-10-second 100-meter dashes. Dave is not. But their sprinting-relevant beliefs are conceivably exactly the same. So their beliefs are doing no work in the explanation of the difference in their abilities, and hence of the difference in their intentional successes. Very plausibly, what is doing the explanatory work here are differences in physicality.

One might reply that Dan does not have the ability to intentionally run sub-10-second dashes, since Dan is not intentionally doing so in the case I described. That violates intuition—after all, Dan is in control of his behavior throughout the run, and he runs a sub-10-second dash. Resistance to this verdict must be based in some ancillary theory. Perhaps on the basis of a principle like this: "An agent intentionally A-s only if she is aware of A-ing." Dan is not aware that he is running a sub-10-second dash. But, as others have pointed out, this principle is problematic. Mele (2001) offers a case in which he intends to make his daughter laugh by sending her an e-mail. He is not aware that this is what he is doing as she opens the e-mail and laughs. But he intentionally makes her laugh all the same.

Pavese might appeal to her theory of knowledge how to reply that Dan and Dave are different in an epistemically relevant way. In other work

Pavese has argued that when an agent knows how to A, she possesses a practical concept that is a component of a proposition regarding how to A, the grasping of which enables her to follow a rule to A (Pavese 2015a, 2015b). One claim she defends along the way is that knowing a rule to A is necessary for knowing how to A (Pavese 2015b). Perhaps, then, Dan knows a rule that Dave does not. Dan knows a rule for running sub-10-second dashes. And since rules are components in crucial propositions, perhaps Dan's knowledge how to run is distinct from Dave's in a way friendly to the intellectualist about knowledge how.

But we can be more explicit about the case, telling the story in such a way that the rules Dan and Dave come to know are exactly the same, and that the only difference between them has to do with physical features such as the elasticity of their calf muscles. The intellectualist about knowledge how will not want these features to make a difference to their knowledge how. Indeed, intellectualists are explicit about ruling such cases out (see, e.g., Pavese 2017: 376). So the intellectualist will need to deny that Dan and Dave have different abilities to intentionally act. But of course Dan and Dave do have different abilities to intentionally act. Dan can intentionally run sub-10-second dashes, and Dave cannot. Once their coach lets them race on the international stage, this important difference in their abilities will be a part of the reason that Dan becomes famous, and widely acclaimed as a great sprinter, while Dave does not.

This case highlights an additional way that beliefs are superfluous to the explanation of intentional successes. In conjunction with cases like Mele's and Setiya's, it should lead us to reject such a requirement. The result is that Pavese's argument, leading to the conclusion that the ability to intentionally A requires intellectualist knowledge how to A, should be rejected.[7]

5.5.4.1 Explaining Intentional Success

I want to discuss Pavese's arguments regarding the explanation of intentional success. The arguments are intrinsically interesting. But my rationale is to highlight an important difference in approaches to action like Pavese's and those like mine.

[7] A further result follows from the conjunction of the above considerations with the principle elucidated at this chapter's beginning. Pavese accepts that if an agent can intentionally A, then that agent knows how to A. As we have seen, Dan can intentionally do something Dave cannot. So Dan knows how to do something Dave cannot. But Dan and Dave can be described in a way such that their beliefs are the same. So the intellectualist about knowledge how has to reject this principle in order to hold onto intellectualism.

Pavese first argues that there must exist a safety constraint on the belief(s) that figure in knowledge how. Weak safety holds when a belief is true "in most (or sufficiently many) of close cases." Safety holds when a belief is true in all the close (i.e., similar) cases. She links ideas thus: "only modally robust beliefs can figure in satisfactory explanations of success and . . . know-how figures essentially in explanations of successes" (Pavese 2018: 14). The reason is that an explanation of a success needs to show that the success was not a mere fluke—was not an accident, was not deviant, was not due to too much luck.

Pavese uses an example of Daniel Greco's (2016) to illustrate the belief case. Alice makes it to the Colosseum. The explanation may very well refer to true beliefs about the Colosseum's location. Pavese says that "Alice cannot intentionally have ended up at the Colosseum without true beliefs about how to get there" (16). And these beliefs should be safe—an accidentally true belief will not show that the success was not deviant.

Regarding this explanatory principle that an explanation of a success needs to show that the success was not a mere fluke, we are in agreement. However, belief does not always play a role in explanations of intentional action—sometimes, intentions or other plan-states will do just fine.

And there is an interesting difference that emerges here. (A version of this difference will re-emerge in chapter seven, when I contrast my account of skill with accounts like Pavese's.) The difference is that the aim of a plan-state (like an intention) is different from that of belief. A plan need not be perfectly veridical to be good. The account of intentional action I have offered requires that the plan be a good plan—the agent's exercising control with respect to the plan or part of the plan is a reliable way for the agent to meet with success—a success rightly deemed an intentional action. It is possible, even if unlikely, for a plan to be good—for a plan to be safe or modally robust regarding the connection to success—without embedding truths about the environment in which an agent acts.

Harriet is a high jumper with a perceptual defect. She systematically perceives, and comes to believe, that the bar is higher than it is. This leads her to form plans to jump with a certain amount of force that reliably leads to her clearing the bar. Suppose that were it not for this perceptual defect, she would be less reliable at clearing the bar. Here is it a false belief that helps explain her success.

Intentional action's proximity to success sometimes crowds out its proximity to truth. This furnishes a further reason why intentional action does not require knowledge.

5.6 A Problematic Case (Systematic Gettierization)

An interesting problem case can be drawn from the literature on Gettier cases. The relevance of Gettier cases to action runs via structural similarities: it has been suggested that, given parallels between knowledge and action, such cases should apply to the case of action (Sosa 2015; Miracchi 2017).

The kind of case I have in mind is what Miracchi (2015) calls a systematic Gettier case. The structure of the case is such that an agent engages in a process, or exercises a competence, that usually produces some product, be it knowledge-formation or action-production. But in this case the conductance of the process or the exercise of the competence is unusual. First, the agent undergoes what Miracchi calls bad luck. Something happens that, in other circumstances, would completely derail the knowledge-formation or action-production. Second, however, the agent encounters what Miracchi calls good luck. Some feature of the situation or the process corrects for the bad luck, such that the agent forms a belief, or produces behavior, that accords with the expected or normal result. What's more, it looks like the agent's performance is explained by her conductance of the process, or her exercise of her competence. And yet the belief is claimed not to amount to knowledge; the behavior is claimed not to amount to intentional action. For the luck involved renders the success a result of deviant causation.

To get a better feel for the structure of the cases, and to see the parallels from knowledge to action, consider the following two cases. The first is taken from Miracchi (2015: 39), and the second is inspired by a case Sosa (2015: 13–14) discusses:

DOUBLE TROUBLE: Creola's two friends Fred and George sometimes like to play tricks. Fred goes on a month-long trip, and communicates with George daily. Fred decides that he will play a trick on George and tell him the complete opposite of what actually happens to him on his trip. George, charged with relaying news of Fred to Creola, decides he will play a trick on her, and tells her the complete opposite of what Fred tells him. As it so happens, Fred and George's subterfuges reliably cancel each other out, so that Creola reliably receives true information about Fred's trip. She forms beliefs on the basis of George's testimony just as she would in any other standard case of belief-formation on the basis of testimony.

THE SHOT: Robin notches an arrow in his bow and prepares to fire at a far-off target. He knows where the target is, because his friend John told

him, but he does not inspect the scene in any detail. This is normal for Robin: he is an accomplished archer, and he and John enjoy this game in which John tells him of a target, and Robin hits it. Robin intends to shoot a shot he is confident will hit the target, even though the shot must curve around a bend, out of sight, in order to do so. Unbeknownst to Robin a strange wind swirls around the bend. The arrow will pass through it, thereby deviating from the target. But the deviation is such that the arrow's tail will graze a nearby tree, putting the arrow back on target.

We are told to think that Creola forms a true belief by exercising her normal competence to believe truly on the basis of testimony:

> We may suppose that, had Creola not formed beliefs using her default competence to believe truly on the basis of testimony, she would have attended to subtle signs in George's voice, eyes, etc., that would reveal his subversion. She thus would have believed he was lying, and have formed false beliefs about Fred's trip. Thus her believing truly, and not just her believing, is causally explained by her competence. (Miracchi 2015: 39)

And yet, we are supposed to think, Creola's belief does not amount to knowledge. So the exercise of a reliable competence to believe truly is not enough for knowledge.

What about the structurally similar case regarding action? Robin executes an intention as he had planned. And his behavior was, in the specific circumstances at hand, a reliable way of satisfying the intention—of hitting the target. And yet one might have the intuition that Robin did not hit the target intentionally.

Ernest Sosa shares that intuition, but pushes back nonetheless. His account of non-deviant causation is similar to my own, but differs on an important detail. Given this—and given that his view of knowledge is a primary target of the systematic Gettier cases in the first place—his response is worth considering.

5.6.1 Sosa's Response and my Own

Sosa's response is situated in a broader competence-based account of intentional action. This account is of a piece with related competence-based accounts of knowledge and perception. But the focus here is action.

According to Sosa, in intentional action the agent must not only reach her aim, she must do so by way of a competent performance. Sosa claims that "aptness—success that manifests competence—is the key to 'the right way'" (Sosa 2015: 19).

It is important for Sosa that the agent's competence possess a certain structure. Sosa calls this a SSS structure. The agent must have Skill, which for Sosa is a basic package of dispositions that enable successful performance. The agent must also have Shape, which is to be in the right circumstances. For example, if the agent is driving, she should be awake and sober and so on. To Skill and Shape, Sosa adds Situation. So, regarding driving, Sosa states that the car should actually work, the roads should be passable, and so on.

This is clearly similar to my own understanding of non-deviant causation. But there is an important difference. Although Sosa embeds skill within shape within situation in the good case, they appear to be separate. That is, an agent may possess the skill to drive even when she is not in the right circumstances, either because the Shape or the Situation are wrong. This hearkens to a distinction between general and specific abilities (Mele 2003a). I do not oppose the usefulness of this distinction for some purposes, but here it may mislead. For I claim that the agent does not possess what Sosa calls Skill in the absence of any circumstances whatsoever—some set of circumstances must be posited in order to get the dispositional structure off the ground. It is better, especially when dealing with difficult cases like the present set, to be explicit about this.

Sosa is not explicit about this, and as a result his response to the cases runs in a different direction than my own. Sosa's verdict is that in cases like THE SHOT the agent does not intentionally hit the target. Sosa requires three features for intentional action: success, competent performance, and that the success itself manifests competence (2015: 19). Robin's performance has the first two, but not, Sosa avers, the third. Here is what he says of a case similar to THE SHOT:

> Why is this shot not apt after all? A performance is apt when it succeeds because of the agent's competence. But our archer's wind-aided shot does seem to succeed because of his competence! If the agent's competence had not resulted in the right orientation and speed upon release from the bow, then the arrow would not have hit the target.
>
> Taking a leaf from Davidson and Grice, we might judge success to be apt only if it derives causally from competence in the right way. Success

> essentially aided by lucky gusts of wind would not derive in the right way from the archer's competence. (Sosa 2015: 13–14)

Sosa is aware that this seems to introduce circularity into the analysis. Non-deviant causation is causation in the right way. If all we can say is that this amounts to success that derives from competence in the right way, we seem to have said too little. But Sosa argues he can go beyond this appeal to the right way. He does so by appeal to a primitive relation of manifestation.

An agent's success derives from a competence in the right way when it manifests that competence.

Manifestation is primitive. We can say little about it explicitly. We grasp it via intuition.

Sosa offers an analogy with a case regarding dispositions. A glass is fragile because it is disposed to break when struck. But suppose the glass, while about to smash on the floor, is instead zapped by a powerful ray that would shatter even an iron dumbbell. Sosa claims that, intuitively, the fragility of the glass is not manifested in the right way. And this kind of intuition is enough to justify claims about non-deviant causation in the relevant cases. So we can say, without offering any further explanation, that in some cases Robin's success will manifest competence, and in some cases not. The differences are picked out by intuition, and these discriminate deviant from non-deviant causation.

This is a bold line to take, and one with a certain attraction. Cases of deviant causation are, after all, picked out by intuition. Why not cases of non-deviant causation?

But relatively brute appeals to intuition can easily mislead. So, what of the intuition that Robin does not intentionally hit the target in THE SHOT? Here intuitions mislead. THE SHOT, and similarly structured cases, tempt us to imagine a different set of circumstances than the one we are explicitly told to consider. This different set of circumstances involves unreliability in the process. Fred could easily err, and tell George the wrong thing. The wind in Robin's case could easily swirl a different way. When discussing these kinds of cases Miracchi uses the term luck, explicitly suggesting that there is something odd about the relevant processes.

So the oddity is just that in the circumstances we are tempted to imagine, the processes would prove unreliable. But these are not supposed to be simple deviance cases. The Gettierization is supposed to be systematic.

But who will stand surety for the modal guarantee? It is the selector of the systematically Gettiered set of circumstances.

And how then do we imagine a case in which the systematically Gettiered process is unsafe? The problem is to be found in one of two places.

In some cases, the systematic Gettierization will be complete, and the problem will be with our imagination. We will wrongly imagine a simple deviance case. And we will wrongly attribute luck to the situation. In this case, however, in the relevant circumstances, these processes are not lucky or unlucky. They are a part of the fabric of things. They are how things work.

So consider a case where things always work in the way they work for Robin the archer. True, Robin cannot intervene in the part of the process that involves the wind and the tree. But we have already considered a requirement of intervenability on control. The lesson here is the same as with the dice. We reject such a requirement. Many intentional actions have significant components that operate ballistically or opaquely.

In other cases, the problem will be with the set the selector has constructed. If things do not really consistently work that way in the type of circumstances specified, then we do have a case of luck, but the luck is no problem. The systematic Gettierization is not complete—it is rather an artefact of the subset of circumstances under specification. And if so, the luck undermines the behavior's qualification as intentional action, and not the account of intentional action under consideration.

As I said above, there is nothing magic about non-deviant causation. It does not matter how ballistic or opaque parts of causal processes operate, so long as they are sufficiently reliable. Systematic Gettier cases have such processes, and agents in these cases act intentionally.[8]

5.7 Causalism and Reductionism

I offer an account of intentional action within the causalist tradition. The non-causalist's most common complaint against the tradition is deviant causation. Underlying this complaint, however, is a potentially deeper issue. For deviant causation is usually mentioned because the non-causalist finds something malignant in what she reckons the more general causalist outlook. The worry is that causalism has reductive ambitions, and that these ambitions are out of place in an account of intentional action. (Why this is so will differ depending on the non-causalist.)

[8] There may be lessons here for those seeking parallels between intentional action and knowledge.

Perhaps some causalists have a reductive ambition. But perhaps they hold this ambition in the way that a philosophy lecturer holds an ambition to amuse her students. Being funny is a thing additional to the core aims of the philosophy lecturer (or assume it is for illustrative purposes). She may succeed in her ancillary aims or not, and her success or failure is apart from her success or failure in teaching philosophy. So too, I suggest, with causalism and reduction. Qua causalist, we need only commit to the existence of a causal condition on the nature and explanation of intentional action.[9] And it is not difficult to see the attractions of this minimal commitment. For intentional action is about things that agents do, things agents make happen, things agents bring about. Some think the causal relation should hold between psychological states and behavior. Some think it should hold between reasons and behavior. Some think it should hold between agents and behavior. Some think it should hold between facts regarding one of these things and behavior. Be all of that as it may—to claim that a causal condition is not necessary for an account of intentional action is to allow intentional actions that take place in spite of the absence of a causal relation between whatever relata one prefers. It is to allow intentional actions that neither the agent nor any relevant agent-involving process causes in any way.[10] That strains credulity.

What, then, are non-causalists upset about? That may depend. Here I discuss two ways non-causalists have pressed causalists regarding reductionism. In each case the worry is pressed as though it is essential to the causalist view. That is in part why I feel compelled to address it. I do not see the reductive ambition in the way many non-causalists seem to. Perhaps, then, there is some prospect for peace. But we will see that there

[9] Something like the following proposal, due to Mele (2017b: 30–1), should do:

> D5. Necessarily, if E is an adequate explanation of an intentional action A performed by an individual agent S, then E cites (1) a reason that was a cause of A or (2) a belief, desire, or intention that was a cause of A or (3) a neural realizer of a belief, desire, or intention, which neural realizer was a cause of A or (4) a fact about something the agent believed, desired, or intended, which fact was a cause of A.

[10] One might say instead that all the non-causalist is commited to is a claim about action explanations: action explanations need not reference any causal condition (see Wilson 1989; Sehon 2016). Or, as Sandis usefully glosses the view, "it is a sufficient condition of a reason-giving explanation of action that the reason cited renders the action intelligible" (Sandis 2006: 12)—where the reason cited need not reference causation, even obliquely, in order to do its work. But it is difficult to avoid the implication that, if causation plays no role in the explanation of intentional action, then it is possible to have actions that do not involve causation. See Mele (2003b, 2019) for discussion.

is a sense in which the perspective I endorse is reductive. So the peace can be, at best, partial.

5.7.1 Basic Action and Intentionality

Douglas Lavin has pressed a worry against causalism's reductionist ambitions by way of a discussion of basic action. Roughly, basic action is an action an agent performs without performing any other action. Lavin argues that causalists need a notion of basic action. According to Lavin the legitimacy of "the seemingly innocent idea of basic action . . . is vital to the intelligibility of the causal theory of action" (Lavin 2013: 274). Why so? Lavin explains:

> Its legitimacy is vital to the intelligibility of the causal theory of action, according to which physical action consists of a mere event and a condition of mind joined (in the right way) by the bond of causality. Left unchecked, means-end reason threatens to permeate physical action, and thus threatens the sovereignty of the sphere of material events at the center of the causal theory: such events, including the movements of one's body when one intentionally moves it, are thought to be constitutively independent of the subject's rational capacities. Basic action is a necessary countermeasure, a sort of metaphysical containment wall needed to preserve the separate jurisdictions of the mind of the acting subject and what merely happens. (Lavin 2013: 274)

Lavin is saying a lot. Obviously, he is working with a certain interpretation of the causal theory, and the metaphysics to which the causal theory is committed. I do not doubt some causalists would recognize their view in Lavin's depiction. I do not recognize my own, however.

Like some (but not all) causal theorists, I have in the past been happy to think of basic action as a useful but inessential notion. As Jennifer Hornsby puts a thought I am happy endorsing, "there must be something right in saying that no one would do anything if everything she might do was something she could only do by doing something else" (Hornsby 2013: 2). Before reading Lavin's paper, I would not have thought that my endorsement of Hornsby here was an endorsement of anything vital to the causal theory, nor of anything resembling a metaphysical containment wall separating the

mind from what merely happens. If Lavin is right, however, that is exactly what I am endorsing. Why does Lavin think this?

Lavin thinks that without basic action, causalism cannot give an account of the structure of intentional action. Consider two representative passages:

> Consider illuminating a room, building a house, or baking a cake. How do we execute complex projects such as these? The answer, says contemporary action theory, is that we perform complex actions by performing a more or less intricate sequence of basic actions, while we perform basic actions immediately, directly or 'just like that'. (Lavin 2013: 273)

> The classifications [of actions as basic and non-basic] is meant to be one we must recognize if we are to understand the very structure of intentionally doing something: whatever large-scale projects one has realized through the ordering of means to ends, one must eventually reach a fine enough resolution and come upon things that have been done without any thought about how to get them done. (Lavin 2013: 276)

Here Lavin gives two characterizations of his target. Basic action is action performed "just like that," and basic action is action performed with no thought about how to get it done. Lavin clearly thinks that the existence of such actions is necessary for a causal account of the structure of intentional action. Unfortunately, without further explication, it is not transparent what Lavin means with notions such as actions performed "just like that" and without thought about how to get them done. Fortunately, Lavin offers a more explicit explication of what he has in mind in the following passage.

> [Basic action can be described] through the concept of an end: a basic action is not the end of any other action; nothing else is done in order to do it; it is not an answer to 'Why?' when asked about any other action. And equally through the concept of a means: no means are taken in the execution of a basic action; it is not done by doing anything else; there is no answer to 'How?' when asked of it. (Lavin 2013: 275)

I will eventually suggest that all these descriptions of a basic action, which are intended to amount to the same thing, in fact do not. Before I get there, however, I want to diagnose why Lavin thinks basic action is needed as a metaphysical containment wall without which causalism would fail to achieve its central motivation. I want to do this because I agree with Lavin

that basic action is dispensable, but I disagree that this makes any problem for causalism.

Lavin thinks basic action is essential to causalism because he sees causalism in a certain way. Lavin thinks causalism is committed to the existence of actions without means-end structure. Basic actions are meant to be practical atoms out of which non-basic actions—those with a means-end structure—are built.

What happens, then, if we show this notion of basic action to be unintelligible? The key aim of causalism is revealed as a failure:

> In giving up on basic action, the means-end order is shown to be at once an order of causality (the means realize the end) and an order of reason (the end rationalizes the means). And this order, an order of practical reason, is shown to be internal to what happens, to the progress of the deed itself. (Lavin 2013: 274)

Or:

> Without basic action . . . a residue of intentionality would always remain in the representation of material processes themselves, of movements as movements. It would not be possible to realize the explanatory ambition of the causal theory, namely to fit action into a world that does not contain intrinsically intentional material processes, unless basic action is the fundamental manifestation of rational agency. (Lavin 2013: 279)

According to Lavin, if causalism cannot remove the residue of intentionality in the representation of material processes, it fails. Similarly, if causalism cannot get rid of the rationalizing of means via ends—presumably, by explaining away such rationalizing in causal terms—it fails. Causalism cannot do these things, Lavin argues, because it relies on basic action to do these things for it.

Suppose we did remove the residue of intentionality from the representation of material processes. What would we have done? On a broad reading of "removing the residue," we might have just naturalized intentionality, where to naturalize intentionality is to show that the semantic (and/or intentional) is not "permanently recalcitrant to integration in the natural order" (Fodor 1984: 232). How that should go is a topic for another place. Here we want to know: is the naturalization of intentionality something to

which causalists must be committed, qua causalists? Certainly many causalists have wanted to fit a workable notion of intentional action into the natural order. But one way to do that is to show that there is a workable notion of intentional action that—while not itself naturalizing intentionality—poses no special problem for the naturalization program.

Perhaps Lavin only means to give causalism the task of naturalizing action theory. But if so, it is not clear why it is a problem that a residue of intentionality remains. It this connection, it is instructive to consider how one of the great causalists, Myles Brand—in a book with the subtitle *Towards a Naturalized Action Theory*—understood the causal theory. In his (1984) book's first chapter, Brand explicitly contrasted the causal theory with what he called the Oldtime Volitional Theory. The latter theory is reductive, aiming to show that "human action consists in causally related nonactional events" (Brand 1984: 7). Brand rejects this theory, noting that he sees no good way it can avoid essential use of action terms in its definition of action. Instead, he proposes the causal theory, which "does not attempt to provide a reductionist account, which it seems clear cannot be sustained" (Brand 1984: 17). Brand is trying to naturalize action theory in one sense, namely by preparing an action theory that is amenable to, and that can be integrated with, a scientific study of human action. But Brand—qua causalist—is not trying to remove the residue of intentionality from his account of action.

I think this is how many causalists view one motivation for causalism. Causalism, as a theory of action, is free to be agnostic over the prospects for naturalizing intentionality.[11] And if so, the causalist need not worry over the residue of intentionality in her representation of material processes, any more than the cognitive scientist does. Whether the residue can be removed depends on progress in naturalizing intentionality much more broadly. Qua action theorist, the causalist is free to utilize intentional mental states (e.g., intentions) and processes (e.g., intentional activity).

How does causalism look without the kind of practical atomism that bothers Lavin? Recall that one way Lavin describes basic action is in terms of means: "no means are taken in the execution of a basic action; it is not done by doing anything else; there is no answer to 'How?' when asked of it" (Lavin 2013: 275). Consider two kinds of actions. For the first kind, illuminating answers to how questions are available. How do I bike from Catherine

[11] Causalism may—depending on the brand of causalism in play—remove the existence of any special action-theoretic problem for the naturalization program.

Street to the Chester Arms? I get on my bike, start pedaling, turn left on Magdalen, right on Stanley, and so on. But there are also actions for which I cannot offer illuminating answers. How do I twitch my thumb just so? I have no explicit knowledge of how I do it, other than to say I know how to form an intention to do it and I know how to direct effort towards the satisfaction of such an intention.

(Presumably Lavin would deny that my directing effort towards the satisfaction of such an intention counts as taking means, but I'm not sure why. Basic actions can be effortful, and we can deliberate about whether the effort is worth it. An agent may think: "Perhaps the means taken to achieve a different end would be preferable." Such an agent does not seem deluded.)

We should not take my own limitations with respect to thumb twitching to signal a limit on what can be said regarding how the thing is done. In principle, for any segment of activity B within some action A, there could be an agent who could provide an illuminating answer regarding how B is done. Perhaps such an agent would need to possess an infinite practical intelligence and bodily abilities to match. Supposing such an agent conceivable, this agent would always be able to break down components of her A-ing into further components. The agent could go on for an infinite amount of time, telling you how the infinitely small subcomponents of her action are executed.[12]

[12] Does this mean that there must be an infinite series of intentional actions the agent must perform in order to intentionally A? Michael Thompson (2008), imagines a case in which one pushes an object from α to ω, and in doing so pushes the object past any of an infinite number of places. Thompson thinks the following:

> It seems that we must also allow that I am doing each of these things intentionally, and moreover that I am doing each 'because' I'm doing the next one. And so, even though the imagined series of isolated positions has an obvious geometrical limit in ω, it seems that an interlocutor and I might together forge a potentially infinite sequence of perfectly legitimate questions and answers, 'Why?'
>
> (Thompson 2008: 113)

I don't deny this is possible in principle for certain kinds of agents. But as applied to human agents, here is a reason to demur. I have offered an account of intentional action on which some bit of behavior B is an intentional action only if B sufficiently approximates the representational content of some relevant plan-state. Now, suppose it is possible for there to be some agent capable of representing not only some prospective action A, but also an infinite number of segments of behavior that together comprise A-ing. If so, then it might be possible for there to be an infinite series of intentional actions an agent will perform on her way to intentionally A-ing. But such an agent will not be a human agent. Human agents do not represent their actions in such ways. So, suppose our infinitely intelligent agent acquires an intention to twitch her thumb that is coded in the same kind of formats as a human intention. In discussing her action of thumb-twitching, her descriptions of her own behavior will at some point cease to sufficiently approximate the content of her intention. As a result, she will no longer be describing anything she does intentionally.

Thompson has this to say:

The causalist need not be committed to the existence of practical atoms, nor to the existence of actions the agent performs without taking means. This means that, for all the causalist says, a residue of intentionality remains. But the causalist—qua causalist—should never have doubted this (if she ever did).[13]

Given that basic action is dispensable, why does basic action of a sort seem inevitable? And why does it look like a useful notion? Why does Hornsby's thought—"there must be something right in saying that no one would do anything if everything she might do was something she could only do by doing something else" (Hornsby 2013: 2)—look so plausible?

My answer is inspired by Hornsby's recent discussion. She outlines a kind of basic action according to which an action A is basic for an agent if she can intentionally A and if she lacks means-end knowledge of how to A (Hornsby 2013: 16–17). Depending on how one understands knowledge of how to A, this may not be quite right. What Hornsby seems to have in mind here is a lack of explicit means-end knowledge of how to A—a lack of the kind of knowledge that would allow an agent, in the course of a piece of practical reasoning, to put segments of her performance into the form of a practical syllogism.[14] It is certainly true that agents sometimes lack explicit knowledge of this form. So I think Hornsby is onto something important.

> Of course, I might put an end to this torture at any one of the interpolated points, saying, 'Well, I'm pushing it to ϕ, you know, because I'm pushing it to ω.' But this doesn't show that any of the intervening 'because'-statements that I have thus left unframed would not have been perfectly legitimate and true. (Thompson 2008: 113)

That may be true in a sense. But there is less reason to think that, for example, my pushing the object from $n_{4238948743}$ to $n_{4238948744}$ is an intentional action. It may be no part of my plan, even if my plan entails such a movement. Might we nonetheless describe the movement as an intentional action? I suppose that we could try, although many would find this under-motivated. My current point is that there is good reason to deny that we must describe the movement as an intentional action.

[13] In fact I suspect many action theorists do not have a fundamental need for basic actions. Two of the better book-length examples of causalist theorizing—Brand (1984) and Mele (1992)—make almost no mention of basic action. And one causal theorist Lavin cites as requiring basic action actually admits that one could build an action theory without reference to basic action (Enç 2003: 47). It is true that Enç goes on to build an account of basic action and make it central to his causal theory. But in my view this is because he thinks that the existence of basic action is an empirical fact about how human action is constructed, and so, as Enç writes, "a definition of action that acknowledges them would seem to be more informative" (Enç 2003: 48). This suggests that Enç is less interested in a conceptual question about the nature of action and more interested in human action, and that he conflates the two issues.

[14] Enç (2003) seems to have something similar in mind. Compare also Enç's definition of basic action, which appeals to an agent's bringing about some result R "without using her knowledge of how to bring about any other event in order to bring about the *result*" R (Enç 2003: 75). By "use of knowledge how," Enç appears to mean use in explicit practical reasoning.

I think basic action seems inevitable to us because we are not agents of infinite practical intelligence. There are bits of behavior that we can perform, even though we have no explicit knowledge of how we do such things beyond the relatively unhelpful knowledge that we do them by trying.[15] Some mental actions seem almost essentially this way. How do I remember things? Sometimes I go through subsidiary mental actions of rehearsal. Other times, however, I simply direct effort towards remembering, and somehow I remember. The same is true of some physical actions. How do I twitch my thumb just so? As I said above, I have no explicit knowledge of how I do it, other than to say I know how to form an intention to do it and I know how to direct effort towards the satisfaction of such an intention.

This, I claim, is a useful notion of basic action even though it commits no one to anything like a metaphysical containment wall needed to preserve a jurisdiction between my mind and what merely happens. It is useful because it marks out an interesting feature of human agency—our knowledge how to do things often outruns our explicit knowledge of how they are done.

5.7.2 Action First

Recently an approach to the nature of action has been gaining some traction. Yair Levy (2013) calls this approach intentional action first, a name that is meant to echo Timothy Williamson's knowledge first approach in epistemology. Lucy O'Brien writes instead of action as prime. In both cases the motivation is the same. We are to give up on offering a reductive analysis of action. Instead, we take action as metaphysically basic. Once we have done so, we can then use the notion of action to explain other items in the philosophy of agency.

The rationale for giving up an analysis of action is thus two-fold. There is a positive side—the notion of action is explanatorily useful, and we need it. There is a negative side—attempts to explain action, or to reductively analyze it at least, have failed. As Levy writes, citing the deviant causation literature "the continued failure to vindicate [a causalist] analysis merits exploring alternative, arguably more promising, research programmes" (2013: 710). O'Brien's opening salvo is rather direct:

[15] There are actions, in my view, which we do not explicitly know how to try to do, though we are able to try to do them, and though we are able to try to do them intentionally.

> Philosophers of action very often start with the question: what happens when someone acts? . . . I am going to go on to urge that there is a certain kind of answer to that question that is often expected, but that we cannot have, and that we do not need. (O'Brien 2017: 265)

O'Brien's position is useful in that, among its other merits, it is clear that her problem is not with the involvement of causation in action, but rather with a reductive ambition in accounting for action. O'Brien lists and discusses a range of potential necessary conditions upon action. Roughly, these can be conjoined into the claim that when I act "I, myself, change in a way that is up to me" (O'Brien 2017: 268). O'Brien's problem is with any account of action that would take a list of necessary conditions—whether hers, or some further analysis of hers, or some alternate list—and claim that they are jointly sufficient in a way that provides an analysis. This would be to "attempt to analyze my action in other terms" (2017: 270). This is to be resisted. For this "would imply that the act—my raising of my arm—is not actually a single unified element in my psychological life but is psychologically molecular. It is composed of all, or some, of the 'more basic' elements we have on the list [of necessary conditions] we gave" (2017: 270).

Why resist a psychologically molecular account? O'Brien presses two worries. One centrally involves deviant causation. Leave it aside. The other is a worry regarding circularity. Here is O'Brien:

> You might think it is true that you need to want, know how to, intend, will, when you act. But what is it you need to want, know how to, intend or will to do? The answer, in our case, is 'to raise my arm'. But to raise one's arm is the action we are trying to understand, so to know what all those other conditions are, we need to know what an action is. (2017: 270)

See where this is going. It is not that the agent cannot somehow intend to do something without having a concept of action, though O'Brien's language suggests such an interpretation. The claim is that no account of action is possible that could factor away action into components—action-types must appear as the content of the psychological states driving action-tokens.

This is unconvincing. Above I offered an account of action in terms of controlled behavior driven by motivational (plan-)states. This account need not depend on the notion of action-types in any circular way. For the content of a motivational state could well be the procurement of some desired object, or the production of some stable pattern of behavior, independently of any

notion of an action-type. Once behavior has stabilized—once an agent can produce the pattern with reliability and flexibility in a given set of circumstances—it becomes useful to talk in terms of action-types and tokens. But action itself is not needed to understand the content of every possible plan-state.

We need not take action to be prime. It is composed of elements—in particular, control and a plan.

5.8 Conclusion

We have reached the end of what I think of as this book's first part. Basic elements of agency have been elucidated. We have accounts of control, nondeviance, and intentional action. In the book's second part I begin by continuing to discuss some basic building blocks of agency. But my aim is different. I use this discussion to begin to turn focus away from what is basic, and towards ways that instances of agency can be in excellent form.

6
The Shape of Agency

6.1 Introduction

I have thus far discussed various components of agency: control, non-deviant causation, intentional action. In this chapter I step back to consider the agent in broader context. The discussion helps motivate what I wish to say in the two chapters that follow this one. Those chapters regard modes of agentive excellence: skill and knowledgeable action. Those modes are easier to understand as excellences once we have a deeper understanding of certain basic features of agency: of what being excellent as an agent is excellence in being.

6.2 Standards, Coherence, and Success

Not every system qualifies as an agent. What is special about those that do?

It is an often endorsed, and plausible, idea that to qualify as an agent a system should conform to certain normative standards. Often this is put in terms of rationality. So, for example, Donald Davidson claims that "An agent cannot fail to comport most of the time with the basic norms of rationality" (Davidson 2004: 196–7). And Christian List and Phillip Pettit claim the following:

> The very idea of an agent is associated with some standards of performance or functioning, which we call "standards of rationality." These must be satisfied at some minimal level if a system is to count as an agent at all.
>
> (List and Pettit 2011: 24)

This sounds plausible. But we might want a more encompassing notion than that given by rationality. Consider a very simple system—the paramecium, a single-celled eukaryote. It has some internal structure. It has some causal powers. It moves through certain liquids by the coordinated beating of its cilia. It can navigate around obstacles or escape certain substances by

The Shape of Agency: Control, Action, Skill, Knowledge. Joshua Shepherd, Oxford University Press (2021).
 DOI: 10.1093/oso/9780198866411.003.0006

way of an "avoiding reaction"—a reversal of its direction of movement, and a slight changing of course in some random direction (Kung and Saimi 1982). This is not a very efficient way of navigating, but the thing is stuck with very short cilia. In any case it is also capable of reproducing, and its methods appear good enough for evolution to keep it employed—many a paramecia has survived long enough to reproduce.

Most philosophers would not want to apply a notion of rationality to the movements of the paramecium. But we might still apply standards of success to its behavior.

In virtue of what? In virtue of the needs or functions we might, as biological theorists constructing a good biological theory or a fruitful model of the thing, impute to it. The constitution of the system—its causal powers, its conditions for continued existence or for reproduction or for functional contribution in some larger system—set in a context of some set of chosen or typical circumstances, sets up a space of standards for that system's behavior. In the case of the paramecium, we might want to model its navigation capacities, and we might impute a goal to avoid some obstacle or some substance as a way of doing so. Whether the goal is a legitimate standard of success for the paramecium may depend upon the quality of our model—whether navigating the obstacle contributes to survival or reproduction rates.

A system might meet imputed behavioral standards in a way that does not even trend in the direction of agency. Consider a system only capable of moving in one direction along a flat surface. We might impute a function to this system, based on its survival needs: the system needs to find and fall into small gaps in a surface on which it moves. If it fits into the gap, it wins. The system does so in the only way it can, by moving blindly in one direction along the wall. There aren't so many small gaps in the wall, but every once in a long while—make it as unlikely as you like—it comes across one. It wins. This may be enough for the system. Such a system does not trend in the direction of agency.

The paramecium does, however. Indeed, Tyler Burge argues that "primitive agency" extends down to the level, at least, of single-celled eukaryotes. Burge points to the orientation behavior of such organisms:

> Taxes are directional movements with respect to stimulations in the environment. They require sensory capacities that are directional. Usually determining direction depends on there being two or more locations of sensory receptors on the body of the organism. Directional movement is

> usually achieved by some mechanism in the animal for simultaneous differentiation of intensities of stimulus registration in different bodily sensors. For example, the animal might turn toward or away from the direction perpendicular to the side of its body that receives the most intense stimulus registration. (Burge 2009: 258)

Burge judges that coordinated, functioning orientation behavior of simple organisms—e.g., "The paramecium's swimming through the beating of its cilia, in a coordinated way, and perhaps its initial reversal of direction" (259)—qualify them as agents. As Burge writes, "Such organisms are capable of steering toward or away from a stimulus source, subsequent to internal differentiations between stimulus intensities in different areas of the body" (258). The movement toward a stimulus is caused in a different way than the movement away from a stimulus, and the difference makes sense in light of the system's own activity—the transitions between states of the system that are differentially sensitive to stimulus source and intensity. That is, the system's behavior is not only reliably produced, it is coherently produced given the circumstances. And it permits something like success. The system's behavior is related to imputable goals regarding their needs (for safety, for finding energy sources, or whatever) with respect to their environment. In their typical behavioral circumstances, this orientation behavior reliably leads to successful (enough) approximation of these goals.

Many will disagree with Burge that we find agency at this level.[1] After all, reproduction is no less complicated and important a process for the

[1] Consider, however, John Hyman's impressively sparse view on (non-intentional) action and agency. For Hyman, Agency is present whenever there is action, and action is present whenever a substance causes a change. Thus, we see action "whenever we see someone walking or speaking…when we feel the sun warming our skin or ice cooling our tongue" (Hyman 2015: 29) One might worry, of course, that Hyman is here changing the subject, or considering a different one. But Hyman does go on to define notions of activity and passivity in complex agents. Activity in the life of complex agents is understood as action that proceeds from the functioning integration of the system's parts:

> The proportion of human acts that stem from an intentional movement is neither here nor there, as far as defining the distinction between activity and passivity is concerned. For the agency of complex substances with functionally differentiated parts always depends on the integrated operation of these parts, rather than on the operation of a specific part or faculty—e.g. the amygdala or the will. This applies to human beings in the same way as it applies to other animals and colonies of animals, and all other complex agents, including institutions and machines. Hence if some basic human activity that involves the integrated operation of cognitive and motor systems, such as feeding or copulating, were only conscious and controllable to the extent that breathing is, a smaller proportion of human acts would stem from an intentional movement than is actually the case, but activity and passivity in human life would be distinguished in exactly the same way. (Hyman 2015: 52)

paramecium than is locomotion. But it is less intuitive to think of a paramecium's asexual reproduction, by a process of binary fission, as an example of primitive agency. That's just the mechanics of life. And if so, why not the beating of the cilia, or the avoiding reaction (which, by the way, often occurs spontaneously)?

Whatever we think about the agency of a paramecium, Burge is right to emphasize continuity between this level and others. At this low level we find key ingredients of agency. Behavioral standards must be imputable to the system. Behavior must be coherent in light of the relevant behavioral standards. And behavior must be reliable in meeting or approximating these standards—the system must succeed, to some degree.

One could try to forge imputability, coherence, and reliability into something like necessary conditions for agency. I decline to do so here, in part because doing so would require a lot of work, in part because I am not sure of the prospects for success, and in part because I do not need to do so. These are, I think, key elements of at least many agents. And these features vary in sophistication, such that more sophisticated agents correlate with increases in the standards imputable, changes in the nature of the standards, differences in the methods by which coherence at meeting the standards is reached, and more sophisticated layering in the mechanisms that help secure reliability in behavior.

6.3 Psychological Agency

A level up from systems like the paramecium, behavioral standards that apply to the system can still be drawn from the system's needs or functions. But at this level the system begins to set certain standards for itself. For at this level the system has the capacity to represent objects, and goals for action regarding these objects—to token psychological states that link it and its behavior to the world in reliable ways. Burge here invokes the notion of a perspective:

> When perception sets an object for animal action, agency reaches a new level of sophistication. The action is suited to a goal that the animal itself

This is an interesting view. But I worry it explains too little. Functional integration is obviously important for any system to qualify as a system. And the notion of a function may give us some indication of what is really important here: the application of behavioral standards to the system. That is the notion that seems to be missing, or too nascent, in Hyman's view.

> perceptually represents. If an animal can perceive, it has some perspective on its objectives. (Burge 2009: 267)

One might think that this is the level at which agency truly emerges. This is what Sean Thomas Foran (1997) argues. According to Foran, an animal moves itself, as opposed to being passively moved as a rock is, when the animal's movements are shaped with respect to objects of that animal's perception. Foran's notion of movement being shaped seems similar to the notion I offered just above, of a system's behavior being coherently produced:

> "Movement shaped with respect to an object of perception" does not simply mean "movement caused by perception." Movement can be caused, in some quite general sense of "caused," by perception without being shaped with respect to the object of that perception. Consider this example. Suppose that when a certain kind of quadruped animal sees one of its natural predators, it immediately lowers itself to the ground and remains still. Perceiving the predator causes the animal to lower itself, but the movement that is caused is not shaped with respect to the predator. The movement is still shaped with respect to something the animal perceives, the ground, but its perception of the ground is not what led it to lower itself: this episode of movement was caused by perceiving the predator.
>
> (Foran 1997: 41–2)

At this level, perhaps, it becomes appropriate to think of coherent production of behavior in terms of practical rationality. When a system can represent behavioral targets, and can implement plans for behavior that approximate standards of success regarding these targets, that system's behavior might well be considered practically rational. And some of that system's behavior might be considered intentional action.

We are still, however, at a level of relative simplicity. At this level it is important that the system be embedded in circumstances in the right ways. For, while the system may be able to represent targets for behavior and deploy plans to hit these targets, the behavioral profiles deployed in following the plan may be inflexible. And inflexible behavioral profiles contain a flaw regarding the meeting of certain behavioral standards.

Distinguish between success according to the standard a system's particular goal or plan sets, and success according to the standards that apply to that system more broadly. If the system is at all complex, then the standards that apply to it will be broader than the standards a particular goal or

plan sets. It will have a range of needs, or perform a range of functions.[2] It may even have a range of intentions, which need to be delicately executed in order not to fail with respect to some of them. Inflexible behavioral routines lock the system into one way of behaving, making it difficult for the system to change tack, or to adjust even slightly. As a result any infelicitous circumstances, or any kinks in the plan, may throw the system off course.

Consider the digger wasp, Sphex ichneumoneus. In preparing to lay her eggs, the Sphex displays some extraordinarily intelligent-seeming behavior. It catches and drags a cricket into its burrow, lays eggs in the burrow, closes the burrow, and leaves.

Regarding this behavior, Woolridge (quoted in Dennett 1984: 11) comments (though the actual details regarding Sphex behavior may be more complicated—see Keijzer 2013):

> To the human mind, such an elaborately organized and seemingly purposeful routine conveys a convincing flavor of logic and thoughtfulness—until more details are examined. For example, the wasp's routine is to bring the paralyzed cricket to the burrow, leave it on the threshold, go inside to see that all is well, emerge, and then drag the cricket in. If the cricket is moved a few inches away while the wasp is inside making her preliminary inspection, the wasp, on emerging from the burrow, will bring the cricket back to the threshold, but not inside, and will then repeat the preparatory procedure of entering the burrow to see that everything is alright. (Woolridge 1963: 82)

Apparently, the Sphex will do this repeatedly, no matter how many times one tampers with the cricket. Commenting on the Sphex's strange behavior, Dennett writes: "Lower animals, such as Sphex, are constitutionally oblivious to many of the reasons that concern them" (Dennet 1984: 24) By reasons Dennett is referring to certain courses of action rationalized by the animal's

[2] This is certainly true of human agency. I say this with Neil Sinhababu's partial account of human agency in mind. Sinhababu advances the following claim:

> Humean Self-Constitution: Agents are constituted in part by all of their desires, and aren't constituted by any other motivational states. (Sinhababu 2017: 167)

Sinhababu continues: "Humean Self-Constitution is neutral about which other non-motivational mental states also go into constituting an agent, or whether any do" (167). My suggestion to Sinhababu would be that one cannot capture human agency without capturing the way humans meet the behavioral standards that apply. At least Sinhababuian intentions—which include means-end beliefs fused to desires—would be required, plus relevant behavioral (i.e., control) capacities.

own background needs, drives, and (if such states can be legitimately attributed to the animal) beliefs and desires. One problem with the Sphex's behavior is it appears blind to a wide range of pressing practical reasons, in the sense that the animal can be placed in circumstances that render it systematically poor at achieving its own basic goals.

This may lead one to reject the idea that such animals can be assessed according to standards of practical rationality. Running against this rejection is the fact that all animals, including humans—such is the lesson of many years of research on human reasoning and its pitfalls (see Gilovich, Griffin, and Kahneman 2002)—can be placed in circumstances that render them systematically poor at achieving behavioral goals.

Susan Hurley argues that many non-human animals can be assessed according to rational standards, but only in certain circumstances:

> An intentional agent who lacks context-free conceptual and inferential abilities and does not conceptualize her reasons can still act for reasons that are her own, reasons from her perspective. Her point of view may provide islands of practical rationality rather than a continuous space of reasons. Reasons for action can be context-bound and lack conceptual generality. (Hurley 2003: 231)

Hurley illustrates these claims by contrasting two kinds of experiments involving chimpanzees. In one experiment, a chimpanzee sees that another chimpanzee (the "indicator") has visual access to the location of a reward in a box, and learns to access the reward by following what the indicator indicates. Then the chimpanzee sees that the boxes are switched, unbeknownst to the indicator, who then indicates where the reward is. The chimpanzee should infer that the reward is in the box that the deceived indicator failed to indicate. But the chimpanzee never does. She follows the indicator's advice and receives no reward. It seems as though the chimpanzee has no access to a crucial reason for action, and that she would have access if she were tracking the relation of the mental states of the indicator to the location of the reward. The chimpanzee must be tracking something else.

And yet, in a second experiment, dominant and subordinate chimpanzees are put in the position to access food. The subordinate chimpanzees are very good at noticing when the dominant has or has not seen the food. When the dominant has not seen the food, the subordinates tend to go and safely access the food. In this case it seems as though the chimpanzee has access to a crucial reason for action, and that this access stems from her

tracking the relation of the mental states of the dominant to the location of the food. Hurley comments: "It may be natural for chimps to compete over food; their ability to act rationally in light of the mental states of others may be evolutionarily tuned to competitive practical contexts rather than cooperative ones" (Hurley 2003: 250–1).[3]

Whatever the proper psychological explanation, the case illustrates the possibility of a system that can be assessable by reference to the norms of practical rationality in certain special circumstances but not in others (cf. Morton 2017).

Now, the range of circumstances in which a system can follow or approximate various behavioral standards will probably vary by degree. In biological creatures, increasingly sophisticated psychological structures correlate with a wider range of behavioral success. The insight here was articulated by Paul Grice in a comment on his method of "creature construction"—a method of modelling psychological systems (which he, in reference to Carnap, calls pirots):

> "What are the general principles exemplified, in creature-construction, in progressing from one type of pirot to a higher type? What kinds of steps are being made?" The kinds of step with which I shall deal here are those which culminate in a licence to include, within the specification of the content of the psychological states of certain pirots, a range of expressions which would be inappropriate with respect to lower pirots; such expressions include connectives, quantifiers, temporal modifiers, mood-indicators, modal operators, and (importantly) names of psychological states like 'judge' and 'will'; expressions the availability of which leads to the structural enrichment of specifications of content. In general, these steps will be ones by which items or ideas which have, initially, a legitimate place outside the scope of psychological instantiables (or, if you will, the expressions for which occur legitimately outside the scope of psychological verbs) come to have a legitimate place within the scope of such instantiables: steps by which (one might say) such items or ideas come to be internalized.
>
> (Grice 1974: 41)

As Grice envisions it, higher and higher forms of psychological system evince psychological states with more and more sophisticated structures.

[3] But see Melis et al. 2006.

Some animals appear to follow rules that approximate or embody norms of practical rationality only in some circumstances simply because the representational systems they use to track the world and drive behavior are not able to deploy rules that perfectly mirror practical norms across all circumstances. The honeybee has evolved a richly combinatorial communicative system—the waggle dance—and a good navigational system. The properties of one honeybee's waggle dance will tell other honeybees where to go to find nectar. But consider a series of experiments in which Gould and Gould (1988) (and Tautz et al. (2004)) had honeybees discover nectar in the middle of the lake, which they then reported to their colleagues. Almost as if they didn't believe what they were seeing, the honeybees ignored the waggle dance. One interpretation of this, as Camp (2009) notes, is that the bees put the states <nectar is there> and <there is lake> together into the state <nectar in lake>, which they subsequently rejected. But an alternative interpretation is that the bees failed to make sense of what they saw because of a limit in their representational system. As Camp puts it, "Perhaps their representation nectar there is blocked from interacting with their cognitive map, because the region on the map marked 'lake' can't receive any other markers" (299). If that is right, then the bees have a representational limit that renders them unable to accord with the relevant norm in one circumstance, even though their representational system is overall well-tuned to deliver success.

It is not my aim here to articulate the metaphorical ladder of agency in any fine-grained way, but it is worth mentioning that we seem to find psychological activity in some animals that approximates practical reasoning, without really qualifying as such. There is a level of psychological complexity in between that of rigid behavioral profiles in characteristic circumstances, and full-blown reasoning of the sort humans sometimes use well.

Anthony Dickinson has argued that many animals fail to instantiate imperial cognition, which he understands as taking "the form of propositional-like or explicitly symbolic representations that are deployed in the control of behavior by processes that conform to some normative standard, such as conditional and Bayesian reasoning" (Dickinson 2012: 2733). But Dickinson argues that these animals can approximate such cognition by deploying associative processes constrained by relatively sophisticated processing architectures. In certain circumstances, for example, rats display behavior that is interpretable as reasoning by disjunctive syllogism. If you give a rat an orange-lime drink sweetened with sugar, and then give it an unsweetened lime drink, the rat goes on to prefer an (unsweetened)

orange to an (unsweetened) lime drink: as though the rat reasoned from the thoughts that <the sweetness derives from the orange or the lime>, and <not the lime>, to the conclusion that <the sweetness derives from the orange>. It turns out, however, that the rule the rat is following is not best captured as reasoning by disjunctive syllogism. Rather, the rat depends upon associations between presently perceived items as well as associations between items retrieved by memory (items generated by the present context) (see Dwyer et al. 1998).[4] This allows a layered form of associative learning that, while not as rational across circumstances as disjunctive syllogism would be, is in fact pretty good. As Dickinson explains, "this form of learning greatly enhances the apparent inferential power of the associative system" (2735)—the associative processes are structured in such a way that though they do not mirror the relevant norm confronting the rat, they do "finesse a problem that would seem to require a rational solution" (2736). The rules the rat follows approximate norms of (practical) rationality rather than embodying them.

So we have an idea, now, of systems that at least approximate norms of practical rationality partly in virtue of their representational and psychological capacities. With higher animals, including humans, we come upon greater representational sophistication, and probably upon a new level of agentive sophistication.

6.4 Higher Animal Agency

The kind of psychological agency that many lower animals display is sensitive to applicable behavioral norms in many circumstances, even if the agents themselves do not represent, or intentionally follow, those norms. We expect better of some higher animals. Some higher animals are capable of deploying reasoning that embodies and mirrors the very norms of practical rationality that apply to the animal in her circumstance.

[4] The way the rat actually works through the problem appears to be this. The rat first associates the things it drinks (sweet orange-and-lime) with each other and the surrounding context. When the rat is then given unsweetened lime drink, sweet-and-orange is also brought to mind as a part of the associative link, which further strengthens the connection with orange and sweetness (in spite of the lime's being unsweetened). So when given the option to drink orange in the future, rats who had been given the lime drink would be more likely to drink it than rats who had not, simply in virtue of the strengthened connection between orange and sweetness.

Consider, for example, a system that forms rational plan-states in part by way of genuine practical inferences. These are genuinely inferential transitions from one set of mental states to another, where the sets of mental states involve practical premises and conclusions—premises about what might be done, and conclusions about what to do. On Quilty-Dunn and Mandelbaum's (2018) view of inferential transitions, inferential transitions between mental states are rule-based in that they "constitutively obey some logic" (537), and the explanation for their doing so is that the transitions are sensitive to—occur because of—semantic relationships between the discursive content of the mental states involved.

> If transitions between thoughts are sensitive to constituent structure, those transitions must obey some logic. This is true because a logical rule just is a kind of rule which is sensitive to constituent structures. For instance, suppose that a rule of mental logic is the following: If X is an AN, then X is an N. Suppose, further, that you token the thought BERTHA IS A BROWN COW. You will then, ceteris paribus, token the thought BERTHA IS A COW. The transition is logical because it occurs in virtue of the fact that the constituent structure of the input representation satisfies the antecedent of the rule, and the output is generated because its constituent structure satisfies the consequent of the rule. Transitions between discursive representations that are triggered because their constituent structures instantiate some rule of mental logic thus suffice to make those transitions rule-based and logic-obeying.
>
> (Quilty-Dunn and Mandelbaum 2018:538)

Practical inferential transitions that obey a basic logic are one example of a way that a system could follow norms of practical rationality. A system might have a goal to eat ripe fruit that, in conditions of foraging, is represented as a conditional: <if ripe, then eat>. Perception of ripe fruit might, in conjunction with this goal, lead to a practical inference: <eat (that) fruit>. A system might possess a goal to break a window, but represent the window as too thick to break by hand. This might lead to a transition from the goal to break a window to a sub-goal to scan of the room, by way of a reliable rule connecting the absence of means with a search process. The sub-behavior of scanning might thus be practically rational given the reliability of the rule followed. And this scan might lead to the perception of a hammer as salient to breaking windows, which might lead to a transition from the goal to break a window and the perception of a hammer that can break windows, to

the goal to break a window with that hammer. That kind of transition might follow a rule connecting means to ends, and it might do so because of the constituent structure of the relevant representations. That is, the system might follow the rule in this case because its representation of the hammer—as, e.g., hard, heavy, swingable, or whatever—has a structure that enables it to interact with the represented goal to break a window, and probably with background knowledge of physics, and thus to generate a rule-following transition to the goal to break a window with that hammer. Following such rules in these ways leads the system to accord with norms of practical rationality in the relevant circumstances, and following these rules across a broader set of circumstances is likely to lead to accordance with relevant norms more broadly.

Of course this discussion of following rules like modus ponens assumes that the system in question represents the world and its own abilities well enough to connect antecedent and consequent in appropriate ways. (Following the rule "if ripe, turn around three times and clap your hands" is not likely to lead to accordance with the norms of practical rationality in many circumstances.) It also assumes the system in question possesses a representational system (or systems) that operate in part via logical rules sensitive to the structure of its representations. As such, the deployment of such rules requires a fairly psychologically sophisticated system.

Importantly, though, the deployment of such rules does not require meta-cognition—the system need not represent its own (first-order) representations. It need not possess any explicit understanding of the norms it follows. As Peacocke has noted, much of our reasoning makes no reference to our own attitudes. The primary concern is relationships of reason that obtain in the world. Peacocke distinguishes between first-tier thought, which is about the world but involves no consideration of relations of reason, and second-tier thought, which "involves thought about relations of support, evidence or consequence between contents" (1996: 130) and which leads to rational mental state revision, without involving higher-order thought of any sort:

> It seems to me a four-year-old child could engage in the following simple piece of second-tier thought. First he thinks that a particular toy is in the cupboard, because that is where it is normally kept; then he remembers that his aunt is staying with them, and that she puts toys in the wrong place. He realizes that the cupboard's being the normal storage place does

> not mean the toy is there now, given the presence of Auntie, and so he no longer believes that the toy is in the cupboard. (Peacocke 1996: 130)

With second-tier thought, however, an additional, potentially very useful, layer of sophistication is acquired.[5] Consider Burge's distinction between critical reasoning and reasoning:

> A non-critical reasoner reasons blind, without appreciating reasons as reasons. Animals and small children reason in this way… Not all reasoning by critical reasoners is critical. Much of our reasoning is blind, poorly accessible, and unaware. We change attitudes in rational ways without having much sense of what we are doing. Often, we are poor at saying what our reasoning is. Still, the ability to take rational control of one's reasoning is crucial in many enterprises – in giving a proof, in thinking through a plan, in constructing a theory, in engaging in debate. For reasoning to be critical, it must sometimes involve actual awareness and review of reasons. (Burge 2013: 74)

At this stage—a stage of adult human sophistication that can involve reflection on our reasons as reasons, and that can involve considerations of relations of reason between our various psychological states, we find a level of agency that Michael Bratman has developed in much detail.[6] We find planning agents (see Bratman 1999, 2007). Bratman is not claiming that only planning agents possess plans. Rather, planning agents are agents with sophisticated psychological architectures, and agents the elucidation of

[5] Philip Pettit has designed an account of reasoning that relies on second-tier thought. According to Pettit, reasoning requires a capacity for meta-propositional attitudes—attitudes towards complex propositions "in which propositions may themselves figure as objects of which properties and relations are predicated" (2007: 498). Such attitudes allow agents to assess their first-order attitudes for features like truth, consistency, and entailment. Thus:

> To be able to reason, under the model I shall adopt here, is to be able to conduct an intentional activity that is designed—and perhaps explicitly intended—to raise the chance of satisfying [desiderata of rationality]. Specifically, it is to be able to ask oneself questions about certain propositional properties and relations; to be able thereby to let beliefs form on such matters; and to be disposed to adjust rationally to whatever beliefs one forms. (Pettit 2007: 499)

Of course human agents do sometimes follow norms in this way. But I think we do so in less explicit and intentional ways as well.

[6] See also R. Jay Wallace (1999): "A person is guided by their conception of their reasons when that conception is reflected in the content of the intention on which they act; in that case, one will be able to understand what the agent is doing only by grasping what speaks in favor of so acting, from the agent's own point of view" (239).

which requires great attention to the structure and normative underpinnings of planning activity:

> In support of both the cross-temporal and the social organization of our agency, and in ways that are compatible with our cognitive and epistemic limits, we settle on partial and largely future-directed plans. These plans pose problems of means and preliminary steps, filter solutions to those problems, and guide action. As we might say, we are almost always already involved in temporally extended planning agency in which our practical thinking is framed by a background of somewhat settled prior plans.
>
> (Bratman 2018: 202)

As Bratman's work elucidates, planning agents are veritably bathed in applicable practical norms. We spend much of our time working through implications of our commitments, testing them against other possible commitments, wondering whether some other course of action might be better in some way, wondering how the plan will impact others, or whether refinements to the plan might make profit along some unforeseen dimension.

The capacity of planning agents to recognize norms for what they are and to form plans influenced by these norms can give the impression that the rationality of planning agents is not sensitive to context.

But even planning agents—at least imperfect ones like humans—must be imbedded in appropriate contexts. The psychological apparatus of humans, after all, piggy backs on the apparatus of our ancestors. We, too, often rely on processes that push us at best toward approximation of the norms of practical rationality. We lean on heuristics, we are sensitive to nudges, we display systematic patterns of bias. This is a chief lesson of the social psychology and behavioral economics literatures of the past fifty years. The lesson, in part, is that certain types of circumstances shape our reasoning in systematic ways, and that while this reasoning can work well, there also exist circumstances in which the reasoning is quite poor indeed.

Jennifer Morton (2017) has recently offered research on reasoning in resource-poor and in resource-rich contexts as a vivid example. Agents in resource-poor contexts display different patterns of reasoning—different patterns of focus, different weighting of future versus near-term goods, different valuing of the resources at hand. Morton argues that the difference in circumstances changes—appropriately—the structure of an agent's practical reasoning:

> Deliberation is something that we do in the non-ideal and messy contexts in which we live. Theories of rationality that abstract so far from the context of deliberation as to render it invisible run the risk of ignoring or, worse, of distorting the experiences of those who exercise their agency in conditions of considerable disadvantage. This is not to endorse relativism about rationality. According to the account that I propose, there are facts of the matter about how it is better to deliberate in one context as opposed to another. (Morton 2017: 557)

Morton is right. The norms that apply to an agent are a function of the agent's constitution—their causal powers, their psychological faculties—embedded in a certain set of circumstances. Perhaps the agency of some angels or gods can be characterized with respect to the set of all circumstances. But imperfect agents like humans are, as all other animals are, embedded in sets of circumstances smaller than the set of all of them. This has important consequences for how we think of their agency.

6.5 Conclusion

In this chapter I have charted a metaphorical ladder of agency. I began by noting key elements of agency—the imputability of behavioral standards, the capacity to display coherence in meeting these standards, and the capacity to do so reliably. Steps up the ladder are marked by important structural and featural differences. Animals with certain representational capacities (e.g., perception) can set behavioral standards for themselves. Animals with increasingly complicated and sophisticated psychological architectures develop methods for meeting these standards with more flexibility, across greater differences in circumstance. And at a certain point we begin to find animals capable of reasoning, and even capable of monitoring and reflecting upon the quality of their own reasoning. At this level an understanding of agency and agentive excellence requires some understanding of activities of planning.

It is worth here considering an upshot of this discussion for the following item: the place of mental action in an understanding of agency.[7] Not every agent engages in mental actions undergirding practical reasoning. But

[7] In a forthcoming chapter (Shepherd forthcoming-a) I consider the implications of this general picture of agency for the "problem" of disappearing agents. It turns out not to be particularly troublesome.

humans frequently do. Is there anything special about the actional nature of such processes?

I do not think mental action is any more essential to agency than bodily action (see Levy 2018 for an interesting perspective). But I do think mental action is tied to human agency (and to agents with similar psychological structure) in an intimate way. In particular, I think the pervasiveness of mental action in our mental lives is a product of our particular computational, informational, and cognitive architectural limitations, as well as the solutions evolution seems to have bequeathed to us. I cannot argue the point in full here, but it seems to me that much of our mental action—and especially the actions that contribute to processes of practical deliberation—is driven by uncertainty and conflict (Shepherd 2015a). This uncertainty and conflict are related to our sense of the norms of practical rationality—often, in deliberation, we are engaged in a search to uncover what it is best to do, or how best to execute a pre-existing intention, or how best to navigate a conflict between various desires, or obligations, or commitments, or whatever. We deliberate because we are informationally fragmented in certain ways (Egan 2008)—it is a struggle to call to mind the relevant items, and to put them together in rationally satisfying ways.

This is a feature of our situation, but it need not transmit to agency qua agency. To see what I mean by this, consider a being constitutively incapable of uncertainty or conflict: an omniscient, omnipotent, and fully practically rational being. Call it Al. It is certainly conceivable that Al, in virtue of its supreme knowledge, never faces uncertainty. And Al, in virtue of its full practical rationality, never faces conflict (unless it be a conflict in the very norms of practical rationality). In whatever the situation, no matter how complex, Al need not deliberate to discern the best course of action. We might say that no matter the situation, no matter how complicated or fraught with moral gravity, Al simply sees—takes in at a glance—the thing to do. Al always acquires intentions reflective of Al's omniscience and full practical rationality. (In order to take in all the information required to always discern the thing to do "at a glance," Al will need some pretty amazing perceptual sensitivity and some pretty amazing cognitive sophistication. We can assume this is covered by Al's omnipotence.)

It seems to follow that neither the kind of uncertainty and conflict that is our normal situation, nor the actional processes of deliberation and decision via which we attempt to reduce uncertainty and accord with norms of practical rationality, are essential for agency (Arpaly and Schroeder 2012 and Morton 2017 make the same point). Further, it seems to follow that

uncertainty, deliberation and decision are important features of our—that is, human—agency precisely because human agency is far from perfect. We have perceptual, cognitive, and volitional limitations, and it is because of this that uncertainty, deliberation and decision play such a large role in our lives.

Even if these kinds of mental actions are inessential to the nature of agency, for agents like humans, the activity of practical reasoning that is essential to our agency is often conducted via mental actions—intentional mental activities like shifts of attention, inhibition of urges, imagination of possibilities for action or consequences of courses of behavior, comparison of action options, weighing of reasons, and so on. I want to suggest that for human agency, mental action is a rational response to the computational, informational, and architectural limitations we face. Mental action is a kind of rational glue—it is the way that we attempt to discover the norms of practical rationality, and to enforce rational coherence across the large but disjointed set of goals, preferences, and abilities that we tend to possess.

A final point turns us in the direction of the next two chapters. My concern there is modes of agentive excellence. Excellence is a kind of perfection of form. So it helps to understand the form of agency. As I have it, this form is that of a system whose behavior, internal and external, is integrated in a way that enables the meeting of behavioral standards. As we will see, skill is a refinement of this—a meeting and even an exceeding of behavioral standards along ever finer dimensions. And knowledgeable action is an important method for meeting standards in otherwise difficult contexts. Knowledge assists an agent in expanding the range of contexts in which she might succeed.

With apologies for waxing in this direction, the lesson of this chapter combined with the next two may be put thusly. Hamlet was only part right:

> What a piece of work is man, how noble in reason, how infinite in faculty, In form and moving how express and admirable, in action how like an Angel, in apprehension how like a god, the beauty of the world, the paragon of animals. (Shakespeare 1996: 112)

Of course Hamlet, with a penchant for drama, was overselling things. In truth human agents are deeply flawed. Not infinite in faculty. On some days, in action not much like an angel, in apprehension closer to a macaque. And yet, we do, in some contexts and in some domains, manage movement both express and admirable, and we do, in skilled and in knowledgeable action, close the gap between us and the gods.

7
Skill

7.1 Introduction

Pistol Pete Maravich was a skilled basketball player. You can look up his numbers. One thing the numbers don't show, however, is the unpredictability of his game. His passes, his shots, the way he moved the ball in space—somehow routinely, he violated expectation. Expectations are currency in basketball. If you know your opponent's expectations, you can plan to violate them. Violated expectations buy time, and create advantage.

Maravich's strange skill set was due in part to his father—the coach Press Maravich. From an early age Press had Pete going through unorthodox drills. The young Maravich may have just wanted to please his dad. But his training regimen instilled a unique set of skills in him. Pete's biographer Mark Kriegel explains:

> The gloves and blindfolds were just the beginning. There were so many other drills. Pete learned the fundamentals, of course: dribbling with either hand, chest pass, bounce pass, foul shots, jump shots, and hook shots. But as the basics could become monotonous, Press invented a more elaborate regimen… In all, there were about forty forms and exercises—'Homework Basketball,' as they would come to be known—to cultivate and harvest every bit of Pete's talent. Press and Pete gave them each names, like 'Pretzel,' 'Ricochet,' 'Crab Catch,' 'Flap Jack,' 'Punching Bag.' He would crouch, his arms moving in a figure-8 motion, between and around his legs, so rapidly that the ball looked as if it were suspended beneath his squatting self. (Kriegel 2007: 64)

Like his dad, Maravich was obsessed with excellence, albeit idiosyncratically. These drills were not geared to produce competence, but rather to push the boundaries of competence towards something better. In this regard Maravich is an exemplar of skill. That's because skill is a mode of agentive

The Shape of Agency: Control, Action, Skill, Knowledge. Joshua Shepherd, Oxford University Press (2021).
 DOI: 10.1093/oso/9780198866411.003.0007

excellence. It is a way that agents, qua agents, display excellence. Any account of skill has to accommodate this directionality.

Of course Maravich is only one exemplar of skill. There are many ways to be skilled, and many things at which to become skilled. I discuss (probably too many) examples throughout this chapter. The space is broad. An account of skill has to provide some unity to the diversity.

I used to have a game called *Infinite Stairs* on my phone. The game involves variations on a simple theme. Using an avatar, one climbs as many stairs as possible as quickly as possible. One does this by directing the avatar to climb stairs in one direction with one button, and in another direction with another button. In practice, this occurs as a series of taps on the screen with one's thumbs or other fingers. The game contains stairs going to infinity, but there are only a limited series of stair combinations that fit on the screen. So one can master, relatively quickly, the relevant combinations—left/right/right/left, or right/left/right/left, etc. One other relevant parameter involves the speed with which one can repetitively hit a single button. And a final parameter involves attention—given the distracting nature of some of the stairs or the background coloring or the motion of the avatar, it is necessary to focus attention away from distractors and only to the upcoming stair combinations. So, when playing *Infinite Stairs*, one improves one's capacity to recognize the relevant stair combinations, to perform the relevant button combinations with speed, and to focus attention in the right ways.

It should be uncontroversial that one can become skilled at *Infinite Stairs*. (After all, I did it.) Other, previously practiced abilities, such as familiarity with a tablet's buttons and with the focusing of attention, will be useful. But one has to structure these abilities in the right way. One has to generate novel visuomotor and visuospatial mappings. So there will be a period during which one is a novice, and then one will gradually improve until at some point, one has become skilled. This might not take very long. In my case, after about a month, the game began to get boring—there were small improvements I knew I could make, but they were very small. I was about as good as anyone should reasonably want to be. And I was about as good as I could be, absent, I don't know, performing quick-twitch exercises with my thumbs.

Other skills take much longer to develop, hold an agent's interest for far longer than a month, and cover more complex territory. Consider the skilled chess player. In cognitive science, this type of person has often served as a paradigm for research into skill—or at least "cognitive" skill—with the result that we know a bit more about skill at chess than in some other areas. Of course, the best chess players in the world today are not humans, and are

arguably not agents. They are special-purpose computers. One reason computers have been able to excel at chess is that skill at chess requires mastery of a very particular body of knowledge (see de Groot 1978; Chase and Simon 1973). It turns out, for example, that chess expertise does not correlate very well with expertise in cognitive subtasks more generally. In one series of studies, chess players were better on a measure of multi-step planning, but not at measures of fluid intelligence, nor at tests for visuospatial and verbal working memory (Unterrainer et al. 2006). Similarly, a psychometric test for chess expertise revealed that the best predictor of chess skill was not any general measure of cognitive skill, such as verbal knowledge or recall, but rather performance in a choose-a-move task. In other words, the best predictor of chess expertise is performance at chess (Van der Maas and Wagenmakers 2005). Instead of fluid intelligence, or problem-solving ability more generally, chess skill may involve something more like the recognitional capacities we deploy when we recognize faces or features of faces (see Boggan et al. 2012).

This is not to say that chess players are unintelligent—clearly chess requires abilities to compute, to imagine, to reason, to recognize patterns, to assess one's own assessments, and so on. However, although this skill depends upon various cognitive abilities, the chess expert need not possess these cognitive abilities to a much greater extent than many others in the population.[1] Instead, expertise at chess can be largely put down to practice at chess. In this, chess expertise is similar to expertise at music, or at visual medical diagnosis (Ericsson and Lehmann 1996). The key is not necessarily general intellectual acuity, but intellectual acuity honed in a specific way, for a specific body of knowledge and set of action-types. This is why, as Ericsson and Lehmann (1996) note, "correlation between IQ and performance in a domain decreases over time, and after more than five years of professional experience the correlation is no longer reliable, even after appropriate statistical correction for restrictions of range" (281). Skill at chess often results from the structuring of cognitive ability sets within the normal human range into skill at chess via the acquisition of chess-specific knowledge.

One interesting feature of some activities is that they are variously composed by a wide and diverse set of abilities to perform more constrained activities. Basketball comes to (my) mind. Some of basketball's sub-skills, especially free throw shooting, have been extensively studied by sport

[1] With that said, a recent meta-analysis did find positive correlations between chess rating and measures of visuospatial, numerical, and verbal abilities (Burgoyne et al. 2016).

psychologists. It turns out that some principles important for explaining good performance at free throw shooting are shared with other aiming activities, e.g., putting in golf, archery, dart throwing, and penalty kicks in soccer (Harle and Vickers 2001; Vine et al. 2011). But of course basketball involves much more than free throw shooting. Indeed, it is possible to be truly great at basketball while truly horrible at shooting free throws (see, Shaquille O'Neal). Since basketball is a game involving different roles filled by people with different body-types and different sets of abilities, there are many ways to be skilled at basketball. Certainly almost all of these ways involve certain basic levels of ability to perform rudimentary physical activities—the ability to move with quickness and pace, for example. But more sophisticated abilities are involved as well. One must be able to catch and manipulate and accurately throw a ball of a certain size, to understand the rules and structures of the game and to discern good strategies for play from poor strategies, to recognize an opposing player's intentions and how these impact one's own strategies and intentions, and so on.

The development of high levels of skill, or even of high levels of control with respect to small components of complex activities like basketball, or violin, or oil painting, can take years even for very advanced agents. One must not only develop a range of abilities and master a range of actions. One must become familiar with complex relationships between circumstances, abilities, and success. Some activities require one to compete against highly intelligent opponents who anticipate one's moves and plan counters. One must develop abilities to deploy these behaviors flexibly and appropriately across a range of challenging circumstance-types.

This is all by way of extensive introduction. The extensiveness is due to the fact that, given the variety of ways an agent can become skilled, the proper diet of examples is important. What I'm doing in this chapter is developing an explanation of the nature of skill—of what it is to possess and exercise skill. The account is quite general, but it sheds some light on how humans come to possess and exercise skill. So it can be fruitfully integrated with relevant empirical work. And it places common elements of skill, like knowledge, in the right place.

7.2 Control

If we are building a skill into some agent, where should we start? I suggest control. That's because one clear element of all skills is control. One cannot

possess a skill S without possessing control over some sufficient range of behaviors involved in exercising S.

I developed an account of control in chapters 2 and 3. Perhaps a brief review of the account will benefit readers of those chapters, as well as folks reading this chapter alone.

I distinguish between control's exercise and control's possession. Control's exercise is displayed at a time, as one behaves or acts in a certain way. Control can be exercised to a greater or lesser degree. Consider Fish, who intends to execute the following plan: fake right, wait for the opponent to move right, then head left and around the opponent. Fish has this plan because he believes it is a good way to accomplish a goal of his, which is to shed the opponent. And so it is. Fish fakes right, the opponent begins to move right, and then, as Fish heads left, his feet get tangled and he trips.

Even though he failed to accomplish his goal, Fish has exercised some control over his behavior. As I put it, he exercised control with respect to some aspects of his plan. His fake right went fine. But Fish did not exercise perfect control—he tripped. How to understand the exercise of some degree of control? We focus on certain aspects of the plan, or upon the total plan, which Fish only partly followed. Fish exercises control to the degree that his behavior accords with his plan, or with parts or aspects of it. Or rather, Fish's control has to do with the behavior that non-deviantly conforms to his plan.

Control cannot be exercised unless it is possessed. That's where non-deviance enters in. To see why, suppose things unfold like this: Fish fakes right, his opponent knows the fake is coming and leans left, Fish fail to notice this and heads left anyway, but in so doing Fish's feet get tangled up with his opponent's, and after a bout of stumbling, somehow Fish emerges on the other side of his opponent, who now lies on the ground. Fish has accomplished his goal. But Fish got lucky. He exercised very little control, and we would probably say that he did not shed his opponent intentionally. Perhaps this was just one case. But perhaps the problem runs deeper. Perhaps Fish possesses very little control over his behavior. If so, we would expect Fish to commonly make mistakes like this. For the possession of control is about how one would exercise control across various sets of circumstances.

The possession of control is the possession of a package of causal properties that enable the agent to flexibly and repeatedly bring behavior to match the content of some pertinent plan-state, across some well-selected set of circumstances. The possession of some degree of control is a necessary condition on a plausible account of skill. Given that skill is a mode of agentive

excellence, one might think that the possession of high degrees of control is necessary for skill. The way I understand degrees of control, possession of high degrees of control is possession of dispositions to flexibly and repeatedly bring behavior to (nearly) perfectly match the content of relevant plans, across reasonably large and well-selected sets of circumstances. So, if Fish has high degrees of control at executing post moves, Fish tends to pull off exactly the moves he has planned.

One might now wonder whether high degrees of control are not only necessary, but sufficient for skill.

Say that Fish has a high degree of control at executing the moves he plans. Fish might still contain a fatal flaw. Perhaps his plans are no good. That is, perhaps the perfect execution of the plans he concocts are rarely conducive to the success of his basketball team. They may look nice, or be impressive as one-off athletic feats, or even lead to some success on occasion. But his coaches may in the end decide that they cannot play Fish, opting instead to play teammates whose control, while slightly lower, is nonetheless executed in service of plans that help the team.

Even so, one might think that Fish is skilled at something. He can perfectly execute his plans in a flexible, repeatable way. Certainly something has gone wrong with Fish—he doesn't seem to understand basketball. But something has gone right as well. Fish is very good at executing certain actions within the broader space of the sport.

We might say that what Fish has is skill at various action-types. This notion deserves brief elucidation.

7.3 Skill at Action

Skill essentially involves an agent's being excellent in some way. The skilled agent is skilled—possesses skill, exercises skill—at something. Normal talk is permissive about the boundaries of this something. Philosophers tend to talk of being skilled at an action-type. Indeed, most of the extant accounts of skill bound skill in exactly this way—an account of skill is an account of skill at A-ing, where A-ing is an action-type (e.g., Stanley and Williamson 2017).

If that's all we have to explain, then it's possible we're almost done.

Assume juggling is a single action-type, with whatever sub-actional components (tossing, catching) necessary to fill out the story. What is it to

be skilled at juggling? I suggest that to be skilled at juggling is to possess sufficiently high levels of control with respect to good plans to juggle. This would decompose into the possession of sufficiently high levels of control at sufficiently many of behavioral components of juggling. And this lends itself to a clean account of degrees of skill. One's level of skill at juggling increases along three dimensions. One dimension is completeness—the percentage of behavioral components with respect to which one has some very high degree of control. Another dimension is control—one's success-rate at plans to juggle, or at the components of plans that have juggling as a key goal, where the successes occur in virtue of the control that one possesses (that is, occur non-deviantly). A third dimension is plan quality—the overall success-conduciveness of the agent's plan for action, given her dispositional structure.

When thinking of action-types with relatively simple structure, it can seem as if control is all one needs to account for skill. This is in part because in the case of relatively simple action-types, it is easy to assume that good plans come for free. Perhaps in some cases they really do. Sometimes only minor familiarity with an action-type is needed to bestow the ability to form good plans for performing it.[2] But Fish's case suggests that this cannot work in full generality. It is possible to have high degrees of control with respect to plans, and actions, which are—in the broader context in which they are embedded—counter-productive.

Reflection on skill's variety suggests skills often run wider and deeper than talk of action-types alone can capture. A skilled debater is good at various kinds of reasoning, at listening, at a way of speaking, at synthesizing information, at presenting information. A skilled surgeon possesses a high degree of dexterity of hand and fingers, coupled with a refined understanding of the function of some part of the body, the ways this function may break down, the ways it may be repaired, as well as an ability to apply this understanding to a variety of case-types: to micro-differences in injury and damage and body-type.

The skilled agent is typically skilled at clusters of action-types. These clusters tend to hang together in a structured way. In order to understand skill we have to understand the structure of these clusters—of what I call action domains.

[2] In fact I think cases exist in which an agent can form good plans for A-ing despite having no concept of A-ing, and despite never having A-ed before. I discuss such a case in chapter 5.4.

7.4 Action Domains

The notion of an action domain is familiar from normal talk about skills. We can think of chess, or basketball, or surgery, or parenting, or teaching, or dancing as action domains at which an agent can be more or less skilled. Before saying what it is to be skilled at a domain, I want to illuminate some features of domains, and of how domains come together.

The most basic constituent of an action domain is an ideal of success. The ideal qualifies outcomes of action, including the actions themselves. Sometimes all you need to have an action domain is a particular ideal. Admittedly, this can make an action domain extremely broad. But consider a domain which has, as its only ideal, the complete domination of some range of entities (other agents, for example). As I envision it, this domain has no restrictions regarding how the domination is achieved. This domain is thus something like that of total war: *bellum omnium contra omnes*, war of all against all. In *Blood Meridian*, Cormac McCarthy's character The Judge seems to view all his actions as cantered towards success in this domain. "All other trades are contained in that of war," he says (1985: 246). And he views war as an all-encompassing domain. Consider the following two passages from one of The Judge's soliloquies:

> Men are born for games. Nothing else. Every child knows that play is nobler than work. He knows too that the worth or merit of a game is not inherent in the game itself but rather in the value of that which is put at hazard. Games of chance require a wager to have meaning at all. Games of sport involve the skill and strength of the opponents and the humiliation of defeat and the pride of victory are in themselves sufficient stake because they inhere in the worth of the principals and define them. But trial of chance or trial of worth all games aspire to the condition of war for here that which is wagered swallows up game, player, all. (246)

> It makes no difference what men think of war, said the judge. War endures. As well ask men what they think of stone. War was always here. Before man was, war waited for him. The ultimate trade awaiting its ultimate practitioner. That is the way it was and will be. That way and not some other way. (245)

I doubt The Judge is correct that all games aspire to the condition of war. The present point is that total war is plausibly a domain of action, constructed out of a single ideal of success.

In some cases the ideal is well known. But in many domains, more than one ideal exists. There may be controversy about how best to articulate the ideal, or about which ideals are most important.[3]

Whatever the ideals in play, often they will imply, or will explicitly contain reference to, constraints of various sorts. These constraints can be built out of different elements. Consider the following four elements, common to more familiar action domains: [a] goal(s), [b] an ordering over the goals in terms of centrality-to-success, [c] restrictions on circumstance-types, [d] restrictions on behavior- (and action-)types.

Typically, an action domain will include some restrictions on the circumstance-types that may arise, as well as the behavior-types and action-types an agent may perform within that domain. Many domains are highly restrictive in both ways. Think of stand-up comedy, synchronized swimming, various board games, cricket, as well as our earlier examples of *Infinite Stairs*, chess, and basketball. The presence of some restrictions is the normal case.

Within the space of more familiar action domains, there will often be more than one important goal, as well as an ordering on the importance or centrality of the various goals. This ordering depends upon some axiology for the domain—upon the ideals of success. Some goals are peripheral, and their satisfaction makes only minor contributions to success in the domain. Other goals are critical. In basketball, success is to score the most points. Given the rules of the game, this gives rise to two central sub-goals: scoring, and preventing one's opponent from scoring. Further sub-goals—goals to do with shooting technique, perhaps, or with defensive technique—make fairly direct contributions to the satisfaction of these goals. A goal involving crisp bounce-passes can contribute, but in a less direct way. All else equal, one should practice one's shot more than one's bounce-pass. If you agree, you share some of the implicit understanding I have regarding the way basketball's goals are ordered.

Many domains lack a single master-goal. They may have a central cluster of equally important goals. Or there may be some goals the satisfaction of

[3] Compare the suggestion made by some philosophers of sport that games, in particular, are governed by an ethos (Morgan 2004; Russell 2004; Simon 2000). According to Russell, for example, "games create opportunities for developing certain human excellences by presenting obstacles that must be mastered and overcome in order to achieve the goal set by the game" (Russell 2004: 146). Russell argues that game players should play the game "in such a manner that the excellences embodied in achieving the lusory goal of the game are not undermined but are maintained and fostered" (Russell 2004: 146). I discuss games as a target of skill in (Shepherd forthcoming-b).

which constitutes success, even if they are not the main goal within the domain. In complex domains, much of this will often be hotly contested. What constitutes success at philosophy, or at teaching, or at painting, or at improv comedy? I am not certain, but I have had enough conversations with some of these people to believe that beliefs differ. There may be a pluralism of top-level goals. There may be many routes to success. There may be multiple ways to exhibit skill.[4]

In some complex domains, one way to succeed is to develop sub-skills, or partial skill (as you like): e.g., to specialize at one important thing. In basketball, for example, this may be ball-handling, shooting, or shot-blocking. If one is good enough at one thing, the absence of (high levels of) skill at other elements of basketball can be forgiven. But of course a better way to succeed as a basketball player, if one can manage it, is to develop interesting combinations of sub-skills. These days, the ability to play defense against multiple positions plus the ability to make three-point shots will make one very rich. But look, it is not easy to develop both sub-skills in tandem.

[4] Because action domain construction is often a species of social construction, an additional point about domains deserves mention. The goals and restrictions that constitute a domain often have a kind of internal rationale. This rationale may be (and often is) contested. At times agents not only seek to perfect their skill within some domain. They may seek to tweak or manipulate or change the domain as well. What is it to engage in domain manipulation? I do not have an account, although I find the question interesting. It seems that an agent may manipulate a domain by manipulating constituents of that domain—behaving or arguing in a way that convinces others that new goals or new restrictions are consistent with the domain's rationale. Or an agent could change the way that people view the behavioral space by introducing unexplored behavior-types or action-types.

The acceptance of the changes agents wish to introduce is often a social matter. (Sometimes it happens without prompting by any one agent, of course.) An agent wants to show other agents what is possible regarding a domain, and so seeks to bring about something new. This has often happened amongst skilled artists, when a new technique is invented that changes the scope of what is possible within some genre of art (or that creates a new genre). This happens when musical artists invent a new technique or style of playing an instrument, or when they subtly change the composition of an instrument to make new sounds possible. Or consider how Asher Lev, a painting prodigy and the titular character in Chaim Potok's novel *My Name is Asher Lev*, describes an interaction with Michaelangelo's David:

> The following morning, I returned to the Accademia and stood for more than an hour drawing the David. I drew the head, with the eyes that reflected the decision to enter the arena of power; I drew the huge veined hands that would soon kill; I drew the shouldered sling being lifted in preparation for the delivery of death. The little man with the broken nose had created this sculpture in an act of awesome rebellion against his tradition and his teacher. Other Davids I had seen were small in size and represented David after the battle. This David was a giant and represented the decision to enter the battle. The little Italian had effected a spatial and temporal shift that had changed the course of art. (Potok 1972: 297)

It doesn't seem too farfetched to think that some manipulations of action domains are fruitful, and some less so.

Defending multiple positions calls upon a number of cognitive attributes, such as understanding the typical goals of different players, understanding where each player fits in a team's offensive scheme, and so on. And of course it calls upon physical attributes, requiring, almost as a necessary condition, speed and quickness and a physical frame between 6-foot-5 and 6-foot-9 with a wingspan longer than that. The ability to make three-point-shots involves an ability to reliably throw a ball through a hoop 18 inches in diameter from a distance of roughly 24 feet. What is involved in this reliability is the demand—at the highest level—that one make a good percentage (say, 38 percent) of one's shots across all relevant circumstances, where these include varying levels of pressure, with thousands of people yelling at you, with defenders running at you or sliding under you, while you are moving in one direction or other quite fast just before you shoot, probably breathing heavily and with a massively elevated heart rate, and so on. Here, small differences matter a great deal. Making 38 percent is good, but making 43 percent (something only ten out of 400 players in an insanely competitive professional association obsessed with making that shot can manage) will put one on a different level. JJ Redick keeps getting paid eight-figure salaries in his fourteenth season in the NBA, based largely on his shooting proficiency (this year he is at 46 percent, nearly a career high).

What are the processes that bring the constituents of an action domain together? It seems foolhardy to try to identify some special class of processes. Examining the many types of domains we might implicitly have in mind when discussing skill (human or otherwise), it seems that a pluralism of processes may be at work in any given case. Processes that lead to the development of particular games could come in for further investigation. But games carve out special classes of circumstances and goals, leading to relatively clean domains. Less clean—and in that way probably more interesting—are the domains that many practices and professions set up. The practices of woodworking, painting, sculpture, and playing music set up various domains. So do the professions—legal, medical, educational, etc. Other domains may be broader than these, in the sense that the goals proprietary to them are largely given by biological needs and drives, and the circumstances are largely given by the contexts in which those needs and drives get expressed or find satisfaction. Additionally, it seems that agents may construct novel domains out of existing resources. Many sports and games begin in this way, but domain construction is not limited to sports and games. Some agents may chart new territory in the space of skill by combining atypically combined domains. The skilled politician, for example,

may represent a fusion of the skilled conversationalist, skilled salesperson, skilled self-promoter, skilled debater, and so on.

This diversity of action domain construction is mirrored by a diversity of relationships action domains may bear to each other. Obviously some action domains will be entirely distinct. But many action domains will overlap to some degree. Action domains may share constituents: goals, orderings amongst goals, or various restrictions on circumstances or behaviors. Action domains may also, in spite of not exactly sharing constituents, call for closely analogous ability-types. Consider similarities between rugby, Australian rules football, and American football. Cases of partial overlap present interesting opportunities for the study of human skills. This is because of what skill researchers call transfer of learning (see, e.g., Schmidt and Young 1987). Some abilities, honed in the context of one action domain, may transfer cleanly. Others may only partially transfer, generating a need to fine-tune the ability in new directions—often, a difficult task given the ways human abilities get constructed by practice. Finally, some action domains may be fruitfully thought of as nesting within broader domains. This might be the case for many of the sub-skills I discussed under the heading of skill at basketball. And, although this phenomena is not as useful for thinking about skill, action domains can become entangled or embedded in others.

Recall Pistol Pete Maravich. He was a skilled basketball player. Arguably, he was a better entertainer. You can look up highlights. One thing I see in those highlights is the ambiguous usefulness of many of his moves. If they were designed to win, it's not entirely clear they hit the mark. But if they are designed to amaze, they are perfect.

I'm not the only one to have this thought. Press Maravich was friends and colleagues with UCLA's legendary coach John Wooden. At one point Pete Maravich's biographer has Wooden acknowledge, of a junior high-aged Pete Maravich: "I watched the Globetrotters with Goose Tatum and Marques Haynes. None of them could do more than Pete. Pete Maravich could do more with a basketball than anybody I have ever seen." The reference to the Globetrotters, who play basketball purely for entertainment, is a tell. Wooden was a stickler for playing the game the right way. And he wondered about the aim of Pete's practice. Wooden once asked Press:

> "How many hours does it take to learn all that? Wouldn't he be better off learning proper footwork for defense?"

> "You don't understand," said Press. "He's going to be the first million-dollar pro." (Kriegel 2007: 61)

It turns out, of course, that professional basketball is not all about winning. There are other metrics of success—other domains entangled within the socioeconomics of the thing. Pete's dad Press understood this.

Action domains are fluid. They are not anything like a natural kind, although some domains may prove more central to an agent's survival and flourishing than others. Action domains are often constructed and changed on the fly, based upon an agent's characteristic circumstances or needs or desires. When action domains stabilize, as happens in games and sports, and as happens partially in domains set up by various practices (e.g., various arts and crafts) and professions (e.g., law, philosophy, medicine), this is in part because of a need for coordination of expectations and activities with others.

Given the seeming multifarity of action-domains, and the diverse ways they can be constructed, why think the notion of an action domain interesting or fruitful for reflection on the nature of skill? One reason is that the most sophisticated natural agents we know of—humans—seem to regiment their practice, their habits, and their actions with respect to action domains. This is a natural strategy for any agent that falls short of omnipotence, omniscience, and omnirationality. By constructing an action domain, one is able to carve out a space of goals and circumstances that are manageable, and one is able to begin to map the relationships between the action domain's constituents in a way that facilitates achievement, and ultimately, excellence. A plausible suggestion at this point is that many paradigm human skills are the product of the way that humans carve the space of actions and goals. Humans are often skilled at (or in) some domain, and this is because skills often bear relationships of mutual support to action domains. Both skills and action domains inherit and contribute structure to each other.

A second reason to care about action domains is because there is no clean distinction between action domains and action-types that would warrant the claim that skill at action-types is fundamental. Some action domains could be reconstrued as action-types. Action-types are extraordinarily diverse, and some talk of action-types—baking, dancing, bank-robbing—may just as well be viewed as naming an action domain.

In spite of this last point, the best reason to think of skill in terms of skill at action domains is that this seems the category of appropriate generality

for skill. Skills may be restricted to particular actions. But most of the interesting and paradigmatic forms of skill involve integration of abilities, directed at interestingly clustered action-types and non-actional behavior-types. An account of skill at action alone is thus incomplete.

7.5 Skill at a Domain

We need an account of what it is to be skilled at a domain. Here is mine.

Skill at a domain essentially involves high success-rates at central goals. These high success-rates cannot be due to luck, or accidents. They must, then, be due in part to control. But not only control—at least not only control with respect to the plans an agent may token. Remember Fish's lesson. Good plans do not always come for free. Skill at a domain of much complexity must be due not only to control (that is, the capacity to execute plans), but also to the agent's capacity to (flexibly, reliably) form good—that is, success-conducive—plans.

The plans should be success-conducive for the agent. Some dreamers cook up plans that would be success-conducive for agents who could actually execute them. But dreamers whose plans outrun their own abilities by too much will not enjoy success.

Embedding veridical representations of the world, of nearby possibilities, of one's abilities, is one good strategy for meeting with success. This is, I think, especially true the more complex the domain is in which we are thinking about success. But veridicality is not the only strategy. Plans need not contain entirely veridical representations in order to be success-conducive.

Consider an agent constructed like so. They perceive drops from ledges to be shorter than they in fact are. This could conceivably benefit the development of skill at parkour, if it leads to less trepidation at jumping off of ledges.

Conversely, sometimes it may actually help to aim a bit high. This suggests that plans need not be perfectly executable by an agent to count as success-conducive. Perhaps success-conduciveness is consistent with some degree of self-deception, provided the deception is about the right thing. Human agents tend to do worse under pressure, for example. It may thus benefit an agent if she is able to somehow ignore, or misrepresent, or lie to herself about, the importance of the moment.

The general point is that when rating the goodness of a plan, success-conduciveness trumps veridicality. In many cases the world is such that plans need to embed accurate assumptions about the circumstances. But given divergences between plan veridicality, plan executability, and success, the important feature is the interaction of an agent's plan (or planning style) and the agent's abilities. If a plan (or a planning style) interacts with an agent's abilities to produce a net benefit for success, then the plan (or planning style) is good in that respect.

Here, then, is the picture.

Within a domain, and given some circumstance-type, a space of behavioral routes can be envisioned. These are mappings between possible agents, possible modes of behavior, and outcomes. These outcomes can be scrutinized in terms of the standards for success particular to the domain. The best outcomes will be ones either constitutive of success or else most conducive to success down the way. Depending on the circumstance, and the available outcomes, these will be outcomes that satisfy central goals straightaway, or that bring agents closer to the satisfaction of central goals, where in most domains the satisfaction of multiple goals would be preferable.

Since the diversity of possible agents is very large, the space of behavioral routes is very large. With the domain and constraints on circumstances specified, we can expect patterns to flow through this space. Some routes and some parts of some routes will appear more success-conducive for varieties of agent, while other routes will trend towards failure, or mediocrity. We can imagine toggling settings on the type of agent, and seeing the space change. One set of routes emerges as success-conducive if the agent is very fast on the ground. This set disappears and another emerges when we take away speed, but add flexibility. This will of course be heavily dependent upon the domain, and the circumstance at issue.

Now we focus on a particular agent. Now the massive space of behavioral routes is replaced by a space of available, at least partially executable plans. The notions of availability and executability are generous. These plans are concoctable by the agent—she could token these plans in the circumstance at hand. And the plans are not entirely unrealistic, in that she could execute at least parts of the plans she forms. Viewing this space, again a flow of patterns will begin to emerge. Some plans will seem more success-conducive than others. And we could toggle features of the agent—arousal levels, motivation levels, physical energy stores—and this would shift the success-conduciveness of the available plans.

The space of available, partially executable plans does not characterize the agent's skill. It may characterize something like her (raw) talent. Talent is associated with unrefined potentiality. It is coherent to say that one agent (Zion Williamson) is far more talented, while another (Julius Randle) is far more skilled. And unrefined potentiality sometimes actualizes. A young Edna St. Vincent Millay writes "Renascence," one of her best poems and one of the most memorable poems of the early twentieth century. She then goes to Barnard College. In a biography of Millay, Nancy Milford reports the following:

> William Tenney Brewster's composition classes were a legend at Barnard, where he was a professor of English and provost of the college. Tall and lean, he would sit with his feet coiled about the wastepaper basket, his fingers toying with a rubber band as he read his students' papers in a flat, dry voice. His comments about Millay's work, which were written in a cramped hand on the back page of her themes, were guarded and almost always on mark. He'd given 'Laddie,' about the death of the family dog, a B and said it verged on sentimentality. When she trotted out one of her old St. Nicholas poems, 'Friends,' he wrote 'Browningesque' and gave her a B. And in one of her less inspired themes, when she wrote, 'Why should it be imperative for me to write a theme? System is a fine thing… But even if I were a literary genius (which Heaven forbid) would I be able to—er—give, as it were, whenever System might choose to wiggle her finger at me? Decidedly not,' he marked 'coy' and added to his B, 'Pretty good for the sort; but capable of improvement.' But he continued to encourage her.
>
> (Milford 2001: 100–1)

The young Millay was clearly capable of great heights.[5] But she was an unreliable writer. Her talent was immature. Skill requires more than talent.

So we move from the space of available plans to the agent's performance profile at plan formation and plan execution. We want to see, across a range of circumstances within the domain, how often the agent forms success-conducive plans. And we want to see how success-conducive these plans are.

[5] From her "Renascence": But, sure, the sky is big, I said; | Miles and miles above my head; | So here upon my back I'll lie | And look my fill into the sky. | And so I looked, and, after all, | The sky was not so very tall. | The sky, I said, must somewhere stop, | And—sure enough!—I see the top! | The sky, I thought, is not so grand; | I 'most could touch it with my hand! | And reaching up my hand to try, | I screamed to feel it touch the sky (reprinted in Milford 2001).

The skilled agent reliably winds up with plans that are among the best of those available to her. The more skilled she is, the better the plans she forms. What makes a plan better than another is that, when the agent sets out along the paths it specifies, she is more likely to meet success. The quality of a plan is thus not independent of the agent's possession of control. This is no guarantee that the agent will successfully exercise her control in some particular circumstance. Even the best laid plans, you know.

So:

Skill at some domain D, for an agent J, consists in sufficiently high success-rates for J according to D's standards for success, where J's successes occur in virtue of J's facility at plan construction and J's control over behavior.

In order to have sufficiently high success-rates according to some domain's standards for success, an agent will need to have a facility at plan formation, and at plan execution, leading to satisfaction of at least some central goals in that domain. How much will be required to merit the judgment "J is skilled at D" will depend upon the domain. Skill at *Infinite Stairs* requires much less than skill at basketball, and skill at basketball may come about in a fairly wide range of ways. Skill at teaching is likely to be even more complex, and to permit more diverse manifestations.

Note well, then: skill at complex domains is likely to require elements like knowledge. I discuss knowledge below and in chapter 8. The point here is that skill itself does not require knowledge.

7.6 The Gradability of Skill

Skill is gradable; agents can possess more or less skill at some domain. I propose three principal dimensions along which skill at D may vary.

One dimension is the agent's actual success-rates at goals constitutive of a domain. We might say this is her height at goal satisfaction. All else equal, a higher success-rate at some goal will indicate greater skill. If all else is not exactly equal, there will be a preference for higher success-rates at more central goals.

A second dimension is the agent's success-rates considered across the goals constitutive of a domain. Fix, for example, some success-rate for an

agent. What percentage of the goals in some domain can an agent satisfy at this rate? This is the agent's breadth. All else equal, a greater average success-rate across all the goals will indicate greater skill. In complex domains, all else is rarely equal. So there will be a weighting for central goals.

A third dimension concerns circumstances. All else equal, the greater the range of circumstances along which an agent demonstrates good height and breadth, the greater the agent's skill. We could call this the agent's depth. Again, when all else is not equal, there may be a weighting for more common circumstance-types.

In some domains, there may be a weighting in favor of more difficult circumstance-types. This may enter in via a domain's conception of success—some games, for example, award more points for goal satisfaction in difficult circumstances. In many cases, I believe, the role of difficult circumstances can be understood in terms of success-rates more generally. Difficulty is impressive in part because it suggests that an agent's skill covers circumstances in which many fail, or because it suggests that an agent's skill is especially reliable. Difficult circumstances often offer evidence that is difficult to come by when times are easier.

Variance along any of these three dimensions can indicate greater or lesser skill. But these three dimensions interact. The ideal is an agent who covers all the goals (and, failing that, all the central goals), with extremely high success-rates, across very large sets of circumstances. But cases exist of agents with, e.g., relatively poor height, and relatively good breadth. Such cases may be difficult or impossible to adjudicate, especially regarding domains where the nature of success, or the centrality of goals, is a matter of legitimate dispute.

This way of understanding skill, and the gradability of skill, also illuminates partial skills. An agent's skill may be partial in virtue of excellence along any one dimension coupled with relatively worse performance along the other two. An agent that has an extremely high success-rate at one central goal in a domain is partially skilled. If the goal and the domain are right, this is a good way to make some money—think of long-snappers in American football. An agent that has good height (high success-rates), but limited depth (limited range of circumstances), is again partially skilled. Think of the clay court specialist in tennis. And an agent that has some middling degree of success across a wide range of circumstances, and a wide range of goals, is partially skilled. The utility infielder (Jose Oquendo, a.k.a. The Secret Weapon) in baseball comes to mind.

7.7 How Agents Possess Skill: Skill for Free

I have said very little about how agents come to possess skill. One complaint about the account presses at this point:

El Oso came to boxing later in life. But as soon as he entered the ring, he felt like he was home. El Oso is eight feet tall, weighs 575 pounds, and sports a muscular frame. He's not particularly good at many of the sub-components of boxing. He's not very fleet of foot. His hands are fairly slow. His punches aren't terribly accurate. But there are things El Oso can do. He can take a punch. In fact, he's taken the best punches the best challengers have to give. He shows no signs of damage. And El Oso can throw a punch. They don't always land, but when he's on target, it doesn't much matter if his opponent sees it coming. El Oso will break a rib, or blast right through an opponent's raised hands. The purists don't like it, but El Oso's the champ. He's 57-0, with 57 knockouts. There's no viable challenger in sight.

This story is a dramatization of a claim Stanley and Williamson (2017) make. The claim is that El Oso isn't really skilled at boxing. The reason is something like: El Oso's success is only, or primarily, due to raw physical ability. And raw physical ability is not skill. Here is what Stanley and Williamson say: "Someone's great strength may enable him to win a boxing match despite his lack of skill at boxing" (2017: 717). I agree, but I disagree with the point behind their claim—that factors like strength and speed are, they explicitly claim, "not themselves part of skill." Raw physical abilities make important contributions to skills, and the levels of skill agents possess. They are rarely the whole story regarding skill. But El Oso comes close. What to say?

Distinguish between skill for free and skill in virtue of tuning. Skill for free is still skill. It comes for free because it is attached largely—perhaps in some cases, entirely—to raw abilities. It still requires the possession of some control, and the capacity for constructing success-conducive plans. But in virtue of massive amounts of raw ability, the plans may be relatively easy to construct, and the control required may be undemanding. So it is with El Oso.

I get it. Normal talk of skill tends to reference skill-by-tuning. I am not trying to avoid violence to normal talk of skill at all costs. And it is a cost to neglect the importance of raw ability to skill. In many domains agents need

some combination of raw ability and finely tuned action production capacities. In many domains it is very difficult to come up with an El Oso. Downhill skiing, chess, basketball, gymnastics, philosophy, neurosurgery—in these domains raw ability alone won't cut it. But raw ability certainly helps. And we aren't shy about celebrating excesses of strength and speed and dexterity, or excesses of cognitive control or attentional capacity or foresight in planning. This is because these excesses, which provide the raw material for practice and learning, are themselves important parts of skill. We say of some athletes (Giannis Antetokounmpo) that they are "cheat codes." This is a compliment. To think otherwise is to cleave the agent in two in a way that leaves that agent's successes in exercising her abilities unexplained by the account of skill.

This is not to deny a drive towards fairness, and more watchable competition, in some domains. We manipulate some domains by introducing weight classes, or age restrictions, or whatever. This doesn't undermine the fundamental point.

Nor is this to deny the existence of ancillary reasons for celebrating skill by tuning. Sometimes we track the praiseworthiness of the training an agent undergoes, and the improvements she has made in virtue of her training.

Sometimes these reasons are aesthetic. There is a reason that, when trying to describe high levels of skill, writers reach for terms evoking artistic achievement. So, Paolo Uggetti writes of the guard James Harden, arguably the best guard to play since Jordan: "To call him simply 'methodical' is to do Harden a disservice. He's omni-intentional. Every offensive move seems calculated and artistic, a balletic performance fueled by emotion."[6] When sports writers offer analogies with artistic achievement in more recognizably aesthetic domains (e.g., "balletic"), they attempt to give voice to the fact that high levels of skill in many sports are aesthetic achievements in their own right. This has to do, it seems to me, with the fact that a drive towards excellence coupled with the complexity of many action domains leaves room for, and indeed, seems in some circumstances to call for, aesthetic creativity, the satisfaction of personal goals, and the expression of something like personal style, in exhibiting skill. Here is how Kyrie Irving—widely considered the best ball-handler alive—explains his approach: "A lot of thoughts that you have to put into action…It's just a constant masterpiece that you have to paint. Sometimes it's going to be all scribble and stuff

[6] From Uggetti (2018).

like that. It's okay to get out of the lines."[7] As usual, it's not entirely clear what Irving is saying. But in this case it should be expected. Agents driven to pursue excellence may find themselves navigating an increasingly fine-grained space of reasons for action, and in such circumstances the ability to imprint one's own style on a performance can be an expression of a high level of skill.

Sometimes, of course, in celebrating skill by tuning, we really are tracking a difference in skill. An agent's virtuosity frequently extends her height, breadth, or depth. I would think something like this is true of Lionel Messi, for years now the world's greatest football player. Messi displays virtuosic and complicated combinations of abilities. He is also prone to succeed in situations, and to see avenues for success, that are almost entirely unique to him. We celebrate the way he has tuned his skill for aesthetic reasons, but also because Messi is simply the best.

7.8 Skill and Knowledge

Skill and knowledge are bound up with each other in a variety of ways—more ways than I will chronicle here. My question is fairly specific. What is the place of knowledge-qualifying mind-to-world direction of fit states (paradigmatically: certain beliefs) in the explanation of skill's possession and exercise?[8] Different answers to this question generate two rival accounts of skill.

7.8.1 Pavese's View

According to Carlotta Pavese, skill is intimately related to (cannot be understood independently of understanding) the knowledge of certain propositions. Pavese does not often directly discuss skill—more often, she discusses knowledge how. But two views on skill can be found in her work. Allow me to briefly present both.

[7] Quoted in Chris Forsberg's (2017) story here: http://www.espn.com/nba/story/_/id/21696375/nba-kyrie-irving-again-rises-challenge-boston-celtics-victory.

[8] My question is thus not about the relation of world-to-mind direction of fit states that qualify as knowledge to skilled action. Some have suggested that intentions could be vehicles for knowledge (Campbell 2018; Dickie 2015), even if intentions do not involve beliefs.

In several places Pavese (2016a, 2016b, 2018) links skill and knowledge how as follows:

Claim 1: If one is skilled at Φ-ing, then one must know how to Φ.
Claim 2: Knowing how to perform a task sufficiently well entails that one is skilled at Φ-ing.
Entailment 1: So S knows how to Φ sufficiently well if and only if S is skilled at Φ-ing.

Claims 1 and 2 are presented as fairly intuitive, and do not commit anyone to a specific view of knowledge how. So for all I say here, they may be true. I do not have an account of knowledge how, nor do I want one. But if one wanted to think of knowledge how in terms of dispositions, then Entailment 1 would be unproblematic for my view of skill.

To the above, Pavese adds an intellectualist view of knowledge how, and thereby of skill. But it comes in two strengths. In her (2016a) and (2016b) she speaks of knowledge how "as a matter of" knowing certain true propositions: "According to intellectualism about know how, a subject S's knowing how to Φ, for some task Φ, is a matter of S's knowing a true answer to the question 'How could he himself Φ?' An answer to such a question is of the form 'w is a way he himself could Φ,' for some way w for S to Φ. Accordingly, S's knowing how to Φ is a matter of S's knowing, for some way w to Φ, that w is a way he himself could Φ" (2016b: 650).

Given Entailment 1, this suggests a strong connection between skill and propositional knowledge—skill would just be "a matter of knowing," where the knowledge in question is knowledge of propositional states regarding how to do something. This seems quite strong, but Pavese sometimes speaks of intellectualism about skill as a way of thinking of skill "directly in terms of standing propositional states," (644) and she characterizes the view in one place as the view that "skills are standing propositional states" (647). Call this strong intellectualism about skill.

A weaker view would be that propositional knowledge of the relevant sort is only a necessary condition on skill: S is skilled at Φ-ing only if S knows a true proposition regarding a way to Φ. This is suggested by Pavese (2018)'s endorsement of a belief/knowledge requirement on knowledge how that is only a necessary condition. Call this weak intellectualism about skill.

I am not sure which view Pavese herself currently holds. But I wish to briefly discuss both views.

Pavese frames her account in terms of skill at an action-type. One response to a key fragment of this argument—the fragment behind premise 4—came in chapter 4.4.4. There I rejected Pavese's argument for intellectualism about knowledge how, on which what is required is knowledge as only a necessary condition. I rejected the argument because it takes a belief requirement on intentional action for granted. But I showed that a negative belief requirement is far more plausible. The agent need not believe that w is a way he himself could Φ. The agent must simply fail to believe that the way she acts, w, is a way on which success is unlikely. The rest can be done by the agent's control in the relevant circumstances.

The chief problem with this version of intellectualism is that it needs belief about a way to Φ to be inseparable from the agent's ability to Φ, or control over Φ-ing. But these things are separable, rendering belief otiose in at least some cases.

Conceive of an agent who has never confronted another agent, nor ever considered how to signal anger to another agent, nor considered how to defend their burrow. Then they confront another agent for the first time, as that agent is raiding their burrow. Via specialized perceptual systems, they perceive a threatening agent. This generates a cascade of processes leading to a plan to defend the burrow. (They do not conceptualize the plan as a plan to defend the burrow, of course. But that is what they are doing.) Their brow lowers, their lips thin, their nostrils flair. It turns out this agent is very reliable at defending the burrow. They know how to defend (or, if you would rather, they are skilled at defending) the burrow without having any beliefs about how to do so.

Or, conceive of an agent who has been trained on a stimulus-response set that guarantees high success-rates in circumstances common to some (admittedly, likely very simple) action domain. This agent has no notion that their training was directed towards the domain, nor that it sets them up for success in the domain. Nonetheless, I find it plausible that they possess skill at this domain.

I may be wrong about the plausibility of such cases. Suppose we grant a belief requirement on knowledge how. If so, I'd now wager that the place knowledge has in skill depends upon the agent's control in an interesting way. For what role does the knowledge play? Pavese suggests that it may enter in at the planning stage: perhaps "the choice of appropriate means to ends is itself guided by a standing propositional knowledge state—say, a state of knowing what to do when" (Pavese 2016b: 645). But it is difficult to see how to think of the more determinate content of this state, and of its

characteristic functioning in any particular action domain, without assuming the agent has skill already in place. That is to say, if the knowledge is really going to guide the agent, that is because the agent will be able to deploy the belief in a controlled, non-deviant way. Such an ability does not come for free, simply in virtue of the knowledge's presence in the agent's mind.

I noted something similar in chapter 3.3, footnote 12, when considering Gwen Bradford's (2015) account of competent causation. Bradford argues that an agent competently causes an outcome when that agent causes the outcome while having some "requisite amount" of justified true beliefs about how E is being caused. But, I noted, of course beliefs about the causation of outcomes can deviantly assist the causation of outcomes. One needs a solution to the problem of deviant causation—that is, one needs an account of control—in order to plug beliefs in correctly.[9] The same is true of knowledge. Knowledge can be misused. Knowledge has no magic in the causation of action that intention lacks. It is not enough to posit the presence of a knowledge state. That state must actually guide the agent's action. And it will not do so simply because the knowledge has the right kind of content. The agent must also have control over the use of that state to guide action.[10]

I turn to strong intellectualism about skill. It too is undermined by the above cases, but it merits discussion because, for one reason, in many of the most interesting and complex action domains, beliefs are required for skill. These beliefs will often be even more helpful if they amount to knowledge.

[9] Perhaps doing so would turn the relevant beliefs into knowledge. Dickie (2012) argues for a slightly different version of a skill-explains-knowledge view.

[10] The same is true of packages of knowledge states. I do not commit to any psychological account of how the control is achieved here, but a natural way to go would be to emphasize their structure. So, consider how John Bengson (2017) thinks of the structure of states of understanding that, he argues, contribute to skill. (For Bengson, understanding is a cognitive, epistemically evaluable state distinct from knowledge—but leave that aside.) Bengson argues that practical understanding undergirds manifestations of skill, and that in order to play this role, practical understanding needs several features. Practical understanding of some activity is a conception of the activity in question the content of which is (at least) [a] correct regarding the activity's features, [b] complete in adequately characterizing the activity's central features, [c] internally coalescent in identifying pertinent substantive connections between the activity's central features, [d] externally coalescent in being rationally consistent with alternative conceptions of the activity, and [e] content over which the agent displays mastery. Such a conception, Bengson asserts, is guiding for the agent: "an individual who has practical understanding will be in a state that is action-guiding, poised to underlie and explain the intentional execution of intelligent action" (43). I find Bengson's work here fruitful for further reflection. It seems particularly interesting to think about what sorts of practice, what sorts of mechanisms, and what sorts of capacities of thought might help realize (and to what degree) these properties of internal coalescence and mastery. (In this connection, see Mylopoulos and Pacherie 2017.)

So I think knowledge is rife in skill, even if it is not necessary for skill at an action, or at some action domains. But it is important to see that even in these domains, strong intellectualism cannot be true—skill is not simply a matter of standing propositional states.

Recall the claim that an agent is skilled at A-ing if and only if she knows how to A sufficiently well. If skill is just a matter of intellectualist knowledge how, then we should expect the agent's degree of skill to vary in lock-step with the agent's degree of intellectualist knowledge how.

To assess the viability of this view, we need an account of the gradability of knowledge how. Pavese has done interesting work on this very issue.

Pavese (2017) considers two ways to think of the degrees of what an agent knows how to do. One way, quantitative gradability, involves ascriptions of knowing in part how to A. Pavese offers a picture on which an agent knows in part how to A when that agent knows all of the propositions that are part of the answer to a question regarding how to A. And an agent knows in part how to A when that agent knows some of the propositions that are a part of the answer. One might think of the quantitative gradability of knowledge how as one dimension along which skill at a domain may vary. One agent may know in full how to A for many important action-types within a domain, and only in part how to A for others.

If this is how we think of knowing how to A sufficiently well, however, there are cases that force apart knowledge how and skill. The cases build upon considerations about the structure of action plans, and about the importance of the interaction between an agent's ability and her plans. They involve two agents, J and K. Both know in part how to A, in Pavese's sense of know how. J knows how a bit less—J knows less of the propositions regarding how to A than does K. But J is more controlled, and more successful, at A-ing than is K, in my senses of controlled and successful. How does this happen? This could be due to J's higher control at exercising movements that are important for success at A-ing. Or it could be due to the fact that the propositions J knows, while less than the number K knows, are far more important for successful A-ing across a wide range of circumstances. Or it could be due to the fact that, while J shares the same true beliefs as K, some of J's beliefs have been Gettiered.[11] I would submit that in such a case J is more skilled at A-ing than K, in spite of K knowing better how to A.

[11] It does seem like agents could possess beliefs that are enormously helpful to action execution, but that are not knowledge, due to Gettierization. There is a literature on this: see Poston (2009), Stanley (2011), Carter and Pritchard (2015), Pavese (2018), Carter et al. (2019).

Pavese's second way of thinking of degrees of knowledge how is qualitative gradability—i.e., "Louis Armstrong knew how to play the trumpet better than any of his contemporaries" (Pavese 2017: 369). Pavese observes that a plausible way of reading this claim involves "better than" modifying "knowing how to play the trumpet." As she puts it: "So, playing the trumpet. Who knew how to do it better than anybody? Louis Armstrong did, that's who" (370). Pavese offers a picture on which the way to think of knowing how better than someone else is in terms of the quality of the answers to relevant questions that one knows:

> s knows how to f better than/as well as s` knows how to f' is true (relative to a context c) if and only if there is a practical answer to How to f that s knows (every part of) (relative to c) and that (relative to c) is better than/as good as any practical answer (every part of which is) known by s` (relative to c). (Pavese 2017: 373)

An initial worry is that, if J is more successful at f-ing than K in spite of lacking knowledge regarding one part of an answer for how to f, while K knows every part of the answer, then J is plausibly more skilled at f-ing even though J cannot know how to f better than k. This might be fixed by dropping the requirement that s knows every part of the answer.

Pavese is here understanding the quality of knowledge how in terms of the quality of practical answers. What explains the quality of practical answers? In response to an ancillary objection, she offers this example:

> Suppose Carla and Ale both know several practical answers to the question How to make ravioli but one of the answers known by Carla is better than any of those known by Ale. One way that answer may be better is by being more detailed and precise; or it may be better by being about a better way of making ravioli (a better recipe); a further way her practical answer may be better is by practically presenting a recipe for making ravioli in a better way than any of Ale's answers . . . a practical sense may be better by being more efficient or simpler, just as certain computer programs can be more efficient than others; or it may be better by being more reliable, just like programs can be more or less likely than others to enable the successful execution of the task. By exploiting this further dimension of gradability for programs, my proposal can also make room for the intuition, voiced by Wiggins (2012: 121–2), according to which one may know how to

> perform a task better because, everything else being equal and under appropriate conditions, one tends to be more successful at the task.
>
> (Pavese 2017: 377)

We are given a few suggestions. The first—more detailed answers—is not necessarily associated with more success at A-ing. The second—being a better way of A-ing—seems uninformative. The third—practically representing a better way of A-ing, in virtue of enhanced efficiency or simplicity—does not quite get us to enhanced success in every case. Action in some domains benefits from more winding (less efficient) paths. Others may reward complexity of practical representations. Pavese's fourth suggestion, that a practical representation is better in virtue of being more reliable, sounds like a potential definition of better in terms of success. But although a knowledge-qualifying practical representation may be more reliable in many circumstances, here I reiterate the point that a plan, or a practical representation, may be reliable without being perfectly veridical, or qualifying as knowledge.

Further, the part of an agent's ability that is independent of her practical representation, or her plan, may contribute significantly to her levels of success or reliability, plausibly changing her level of skill without influencing her level of knowledge how.

The strong intellectualist about skill might respond as follows: If the agent's ability is cognitive, then it must just be further knowledge how. And if it is non-cognitive, then it is not a part of her skill. Pavese considers the following objection to her account of qualitative gradability:

> Could not two subjects possess the same amount and quality of propositional knowledge and yet differ in the degree to which they know how to perform a task? . . . If so, one may be better at a task than another because one's ability to perform the task is superior, independently of what propositional knowledge one possesses. "Ability" here means mental or cognitive ability, not simply strength or fitness. (Pavese 2017: 375)

So strength and fitness play no role in quantitative knowledge how. Pavese continues:

> On the general picture outlined thus far, the [mental] ability component cannot vary independently of the knowledge component, for it is knowledge of the relevant practical answer that endows one with the relevant

> ability and corresponding counterfactual success. Thus, on this proposal, it simply cannot be the case that two subjects have the same relevant kind of propositional knowledge about a task—and in particular, knowledge of the same practical answers—and yet differ in their ability to intentionally perform the task (although, of course, they may differ in their nonmental strength or fitness). (Pavese 2017: 376)

In a footnote she adds motor acuity, that is, the changes due to motor-skill learning that enable an agent to execute an action "with more precision and accuracy" (Krakauer et al. 2019: 651), also makes no difference.

Now, these seem false as claims about mental ability. But I do not wish to argue the point. For even if non-mental abilities, and motor acuity, make no difference to one's level of knowledge how, I submit that these plainly make a difference to skill. Indeed, it may very well be because of these features that a particular practical representation contributes to the agent's reliability, or levels of success.

So skill is not just a matter of intellectualist knowledge how. For, in part, the degrees of knowledge how are not the same thing as the degrees of skill. This is consistent with the thought that knowing how sufficiently well—in the ways Pavese's excellent work illuminates—is critical for understanding the structure of many human skills.

7.8.2 Stanley and Williamson's View

Stanley and Williamson (2017) differ from Pavese. But they too would make knowledge prior to skill: "Skill at Φ-ing is a state whose nature is constituted by the knowledge relation" (721). How so?

Stanley and Williamson argue that a skill is "a kind of disposition to know"—that is, "to be skilled at the action type of Φ-ing is to be disposed to form knowledge appropriate for guiding tokens of Φ-ing" (715). It is important that skill is identified with a disposition. This allows Stanley and Williamson to avoid circularity—a skill is not a competence (or any other skill-seeming ability) to acquire knowledge.

Stanley and Williamson do not explain what guidance ultimately comes to. But the notion is important for their account of skilled action, which piggybacks on the account of skill. They draw a distinction between the direct manifestation of a skill, which is knowledge appropriate for guidance, and the indirect manifestation of a skill, which is action guided by acquired

knowledge states: "any skilled action is guided by knowledge that manifests [in the direct sense] possession of skill at that activity" (718). One can discern, then, two different accounts. Skill is a disposition for certain cognitive changes to occur, leading to the acquisition of knowledge. Skilled action is action guided by knowledge the acquisition of which is a manifestation of skill.

One odd feature of this account, of which Stanley and Williamson are aware, is this. One might have thought that the essential manifestation of skill occurs in skilled action. But since they hold that skill and skilled action are separate things, Stanley and Williamson's account of skilled action makes action an inessential manifestation of skill. Stanley and Williamson demote skilled action's role in an understanding of skill in order to emphasize "what is distinctively mental about skill" (721).

One kind of response to this feature of their account is given by Weatherson: "There's something suspicious about a theory of physical skill that divorces it so strongly from the physical" (Weatherson 2017: 382). I think there is something true in that, though skilled action need not be and is not always bodily. We could restate the point like so. There is something suspicious (i.e., false) about a theory of skill at action that divorces it so strongly from the execution of action.

A further problem for Stanley and Williamson concerns the gradability of skill. My points here are similar to those made in response to Pavese, so I will be brief.

Stanley and Williamson mention three ways their account might incorporate gradability. First, one might become disposed to acquire the guidance-apt knowledge more quickly. Second, one might become disposed to acquire more of the relevant facts in a given situation. Third, one might become disposed to acquire qualitatively better information. (They say nothing about what "quality of information" comes to outside of citing Pavese's work on qualitative gradability. Since we have discussed that, I set it aside.)

I think we can agree with Stanley and Williamson that a disposition to acquire knowledge can be graded. I also agree with them that comparisons of skill are sometimes difficult due to the multiple dimensions involved in the relevant assessments. As they say, comparisons of skill often requires "marrying distinct scales" (Stanley and Williamson 2017: 723). However, the account I offer enables a much fuller sense of why this is so. If skills are skills at domains of action, we can see why difficulties in comparisons of skill often emerge. There is often vagueness in the ways action domains get fixed, leading to verbal disputes. So, arguably of course, Jordan is the

greatest player ever, but LeBron may end up with the greatest career; Nadal is better on clay, but Federer is better on grass; Hemingway's use of terse sentence structure makes for thrilling reading, but his female characters are often flat; Einstein was brilliantly insightful, but would be rubbish at running a high-powered modern physics lab; Francis Bacon's portraits are dark, troubling, and great, but David Hockney's almost whimsical portraits may on the whole stay with you for longer; Philosopher A (no names!) is a lovely synthesizer; Philosopher B has the most devastating counter-examples; Philosopher C's ideas are alluring but good grief does C use some imprecise metaphors. Moreover, the account I offer is able to capture the performative element in skill, explaining how two different performances can exemplify different degrees of skill independently of any question about knowledge acquisition during those performances.

Stanley and Williamson's remarks on gradability are hampered by the fact that they separate the knowledge acquired with its role in guiding action. They say that one's disposition to acquire guidance-apt knowledge may improve if one becomes disposed to acquire this knowledge more quickly. But why think quicker equals better? There is no reason, qua disposition to acquire, to prefer speed. Insofar as the knowledge is guidance-apt, one might think it depends on the context of use. Sometimes quicker is better, sometimes slower—it depends on the kind of action at issue, and accordingly on the kind of knowledge and how and when one needs it. The same point can be made regarding their claim that one's skill might improve if one becomes disposed to acquire knowledge of more facts. But more facts do not always mean more control over behavior. Sometimes more facts swamp or distract attention. The point here is that assessments of skill should be linked to some plausible standard. One natural one is the guiding function Stanley and Williamson regard as crucial. But to make that the essence of skill's gradability pulls against their account, on which the guidance of action by knowledge is only an indirect manifestation of skill. As a result, their account seems to imply cases in which an increase in skill undermines the execution of skilled action. The result is that knowledge's prime value in this context—that it tends to guide action better than some representational state that falls short of knowledge, or that is not a knowledge-qualifying state (perhaps, e.g., an intention)—is not well explained.

This is not to deny the importance of cognition, and indeed of knowledge, for most (if not all) of the most interesting skills. One often does need the disposition to acquire knowledge that can—in conjunction with other states such as intentions and perhaps less safe but truth-apt states like

predictions—guide one's behavior. But it seems more plausible to say that acquiring guidance-apt knowledge is something at which one can be more or less skilled. If so, these dispositions are not themselves skill, though they can be structured in ways that constitute an example of skill, and that contribute to skilled action more broadly.

7.9 Conclusion

My overarching concern in this chapter has been to understand skill. In the broader context of this book, I wish that understanding to fall into place as a mode of agentive excellence. Excellence is a kind of perfection of form. So it helps to understand the form of agency. As I have explicated it, this form is that of a system whose behavior, internal and external, is integrated in a way that enables the application of, and the system's meeting of, behavioral standards. Skill can be seen as the possession of structure by a system that enables excellence according to the behavioral standards that action domains set.

I closed the chapter by discussing rival accounts of skill. These accounts overestimate the role of knowledge. But look: knowledge is clearly critical for many human skills. That is not in dispute. In fact, once we see clearly how skill and knowledge are distinct, there is room for a view on which action that intimately involves knowledge—what I call knowledgeable action—forms a distinct mode of agentive excellence. This is the subject of the next chapter.

8
Knowledgeable Action

8.1 Introduction

"Trust me, I know what I'm doing" was the catch phrase of Sledge Hammer, the star of the satirical 1980s cop show *Sledge Hammer!* The joke was that he didn't. *Urban Dictionary* makes the same joke. Its top entry for "trust me, I know what I'm doing" is this: "A phrase often used by people who don't really know what they're doing." This is kind of funny, because it's kind of true.

Now, human agents often know what they are up to. When they do, things are going pretty well for them. In this chapter I want to make sense of how knowledge contributes to a particular mode of agentive excellence—what I will call knowledgeable action.

Knowledgeable action is intentional action that embeds knowledge in specific ways. I say more below, but to preview: sometimes agents act intentionally. Sometimes when they do, they also act knowledgeably. Knowledgeable action is a mode of agentive excellence in extending the agent's success in ways I elucidate.

Two distinct targets of knowledge are of interest—knowledge of action and knowledge of how to act. Knowledge of action signifies the agent's knowledge that she is A-ing, as and when she is A-ing. Knowledge of how to act signifies knowledge states that are about ways to A.

My guiding concern in introducing knowledgeable action is askew from mainstream rationales for caring about the role of knowledge in action. One rationale is driven by an aim to understand knowledge how to A. I do not share this concern, although I will discuss connections and divergences between my view of knowledgeable action and two views on knowledge how in section 8.9.

A different rationale is driven by the thought that knowledge of A-ing contains unusual epistemic features. Knowledge of action is epistemically special. I share an interest in understanding these features. And things will turn out fortuitously. For as I build an account of knowledgeable action,

The Shape of Agency: Control, Action, Skill, Knowledge. Joshua Shepherd, Oxford University Press (2021).
 DOI: 10.1093/oso/9780198866411.003.0008

I will also develop a view on the epistemic specialness of the knowledge embedded in knowledgeable action.

In this chapter I give more attention to knowledge of action than to knowledge of how to A. Until section 8.7 I speak primarily of knowledge of action. At that point it becomes possible to bring knowledge of how to A more directly into play. This is a little bit clunky. Trust me, I know what I'm doing.

8.2 A Word about the Knowledge State

Some philosophers think of knowledge as involving belief. There are different ways knowledge of action could do so. Perhaps in many cases relevant beliefs about what I am doing are very closely associated with my intention (Pavese 2018). Perhaps my intention necessarily involves a belief about what I am doing, or how I am doing it (a la cognitivists about intention, like Marúsič and Schwenkler 2018).

Some philosophers think of knowledge as a *sui generis* cognitive state (Williamson 2002).

Some have argued that knowledge of action could involve (non-belief-involving) intention alone (Campbell 2018).

I am agnostic, in this chapter, about all of this. But I will discuss various philosophers who hold, and speak out of, different views on the kind of state knowledge of action is. It is most natural, I think, to take the view I articulate as a kind of belief-involving, belief-is-distinct-from-intention view. But it could just as well be understood in terms of a *sui generis* knowledge state, or in terms of an intention if intention includes the right kind of belief.

8.3 Knowledge of Action's Epistemic Features

Start with self-knowledge. An agent's knowledge about certain aspects of her mind is in certain ways special or distinct from other sorts of knowledge. Typically, the literature on self-knowledge restricts discussion to self-knowledge of mental states. So, in a recent introduction to the topic, Brie Gertler writes that "In philosophy, 'self-knowledge' standardly refers to knowledge of one's own sensations, thoughts, beliefs, and other mental states" (Gertler 2015: 1). And in a recent book-length treatment of the

recent literature on self-knowledge (Coliva 2016), one finds no discussion of knowledge of action. But arguments are not given for omitting discussion of knowledge of action; the omission is arguably artefactual.[1]

An agent's knowledge of her own action is often said to be special or distinct in much the way self-knowledge of other elements of mind is said to be. Indeed, in my view the consideration of knowledge of action as a part of self-knowledge of aspects of mind more generally proves useful in setting the epistemological task facing the philosopher. In both cases, the task is to clarify what is distinctive about the knowledge in question, to demonstrate that the knowledge really is distinctive in the elucidated way(s), and to elucidate the grounds for claiming that what is distinctive amounts to knowledge.

Move to what might be distinctive about knowledge of action. Elizabeth Anscombe famously asserts that an agent's knowledge of her own action is in some sense non-observational. Anscombe introduces this claim in order to explain the criterion she proposes for distinguishing intentional from non-intentional action.[2] Intentional actions "are the actions to which a certain sense of the question 'Why?' is given application; the sense is of course that in which the answer, if positive, gives a reason for action" (Anscombe 2000: 9). But what is it to give a reason for action? For Anscombe, one cannot give a reason by simply citing a cause of the action. Nor can one give a reason by citing "evidence for supposing the thing will take place" (9). Instead, she suggests we understand the giving of reasons in terms of the acceptance or refusal of the application of the question "Why?" to some bit of behavior.

Two ways of refusing application of this question to one's behavior involve knowledge. First, we refuse application of the question if we answer "I did not know I was doing that" (11). Second, we refuse application if we answer "I knew I was doing that, but only by observation" (14). As Anscombe has it, then, non-observational knowledge of one's action is necessary (though insufficient) for legitimate acceptance of the request for reasons. This is meant to highlight a revelatory truth about the nature of intentional action: unlike our non-intentional behavior, our intentional

[1] In this connection, Lucy O'Brien writes: "Given the recent debates about self-knowledge and first-person authority it is surprising that there has not been more discussion about our knowledge of our actions. It is surprising because our knowledge of our own actions seems, prima facie, to share many of the features of our knowledge of beliefs and perception, that have given rise to these debates" (2009: 156).

[2] In explicating Anscombe's view, I need not endorse her distinction between non-intentional and intentional action, nor her way of drawing it.

actions are (normally) those things we intimately know, in the sense that we easily accept and easily answer the request for reasons.

But why does knowledge of behavior drawn from observation render one unable to legitimately accept the request for reasons? In her reconstruction of Anscombe's reasoning on this point, Hanna Pickard (2004) invokes the "intuition that one's knowledge of one's own actions bears some affinity to one's knowledge of one's own mind" (207). This is an interesting thought. As Richard Moran (2004) observes, central to Anscombe's thinking on knowledge of action is the thought that "the agent's own conception of what he is doing is not just another description, side by side with all the others . . . There is a privileged relation, though not incorrigible, between what the agent is doing and what he takes himself to be doing" (Moran 2004: 44). The connection between knowledge's non-observational nature, and an agent's authority in knowing it, follows closely behind this observation. Pickard notes that one's knowledge of an action "is not something about which one is especially authoritative if it is based on observing, as anybody suitably positioned can, what gets done" (Pickard 2004: 207). The idea that we can understand the purpose of Anscombian knowledge without observation as part of an explanation of an agent's authority in knowledge is promising.

What is it, though, for an agent to be authoritative in knowing what she is doing? Alex Byrne (2005) helpfully distinguishes between peculiar epistemic access to an item and privileged epistemic access to an item. Privileged access is understood in terms of an increased likelihood that one's beliefs regarding the item in question will amount to knowledge. And the increase in likelihood is understood as relative to some contrast class. As Byrne's focus is knowledge of one's own mental states, he notes that one could have privileged access to one's own mental states in comparison with one's own access to other items (e.g., beliefs about one's environment), and in comparison with another's access to one's mental states. Peculiar access is understood as "a special way or method of knowing" the item in question (Byrne 2005: 81). One reason for drawing this distinction is that these two types of access are logically distinct. As Byrne notes, some maintain that our access to our own mental states is non-peculiar—fundamentally the same as the access others enjoy—although we may enjoy privilege due to our increased familiarity with ourselves. By contrast, it is possible to maintain that we have peculiar access to our mental states (perhaps via some faculty of inner sense), but that this access is underprivileged, perhaps because "one's 'inner eye' is very unreliable by comparison with one's outer eyes" (81). Another reason for drawing the distinction is that in the good case

there seems to be an explanatory relationship between these two modes of access. As Byrne notes, in good cases "they are connected: the kind of peculiar access that we enjoy presumably explains why we have privileged access" (Byrne 2011: 202, see also O'Brien 2009: 158).

We might apply the same distinction to a difference between one's epistemic access to what one is doing and another's epistemic access to the same thing. On this application, one has peculiar access if one's way of coming to know what one is doing is not available to others. And one has privileged access to what one is doing if one's beliefs about what one is doing are more likely to amount to knowledge than the beliefs of others regarding what one is doing. Ideally, an explanation of one's authority in knowing what one is doing will capture the explanatory relationship between these two modes of access. As I will understand it, then, an agent's authority in knowing what she is doing is constituted by the agent's having both peculiar and privileged access to what she is doing, where the peculiar access helps explain the privileged access.

Thus understood, an agent's authority in knowing what she is doing is in part explained by the knowledge's non-observational nature, because if one's epistemic grounds for this knowledge were fundamentally observational, then it would fail the above explication of authority. If an agent had her knowledge in a purely observational way, she would not have it in a peculiar way. But peculiarity is a necessary condition on my understanding of authority.

Now we ask how an agent comes by this authoritative knowledge. Philosophers have tried various routes. Controversy abounds. In what follows I discuss some of these attempts, and I explain why I find them wanting. Then I offer my own proposal, and explain how it illuminates knowledgeable action as a mode of agentive excellence.

8.4 Intention and its Limitations

Say an agent intends to A, and is in the midst of executing the intention. He is A-ing. In discussing knowledge of A-ing, Kevin Falvey (2000) and Sarah Paul (2009a) appeal to the agent's intention to A to explain his knowledge of A-ing. Their views, while distinct, both appeal as well to the openness of certain act-descriptions (the so-called atelic progressives). Sarah Paul offers an example:

> The truth of the statement 'Jen is crossing the street' does not entail that Jen crossed the street; she might turn around and go back, or she might

> get run over by a truck, without its ceasing to be true that she was crossing the street. (Paul 2009a: 16)

Paul argues that for actions under such descriptions, "if the agent is or has been doing them at all, it is true that he has done them," and thus "the agent can count as so acting even if the goal of the activity is never successfully achieved." For such actions, then, "given that the agent need not ever achieve some terminal endpoint in order to count as doing [such actions]…he also need not have observational evidence that he is succeeding to know he is doing it" (16). Paul claims further that something similar can be true of telic progressives: in their case, "the agent can count as so acting even if the goal of the activity is never successfully achieved. As Anscombe puts the point, a man can be doing something which he nevertheless does not do (as when Jen gets killed halfway through crossing the street)" (16).

How does the agent come by this (non-observational, and therefore at least peculiar) knowledge? For Paul, the agent can do so by inference from his intention,[3] when appropriately situated. Paul elaborates upon the agent's situation as follows. The agent must have favorable background beliefs regarding: a) their ability, b) their circumstances, c) their history as a reliable agent, and d) their "understanding of the way some action descriptions apply partly in virtue of [their] intention in acting" (Paul 2009a: 18).

This certainly seems like one way an agent can come by knowledge of what she is doing. Just as certainly, there seem to be many cases that fail one of these conditions, in which the agent knows what she is doing. Imagine that you have suffered from a recent bout of Miller Fisher Syndrome—a rare reaction to a virus that causes nerve de-myelination, resulting in muscle weakness and disorders of balance and motor coordination, among other things. Say that you have been unable to descend stairs without help for a few weeks, but you awake one morning and decide to try. The prospect of doing so successfully is much less certain for you than it was a few weeks ago. In fact, you may be uncertain whether you can succeed, and you may only have an intention to try to A. (We could run the case with an intention to try to A, or to A.)

[3] While Paul's account knowingly discards some of the epistemic specialness, Falvey's account is rather different. He suggests that knowledge of A-ing consists in knowledge of an intention to A, provided one is entitled to assume the execution of the intention is going well. My worry about this account is similar to my worry about other intention-based views. The scope is too narrow.

What happens when you descend the stairs? Speaking from experience, I can say that one descends with an intentionally heightened awareness of the progress of one's actions. Where once one descended unthinkingly, now one grips the rail, monitoring footfall, muscle tension, balance, and speed. One is very aware of what one is doing in part because one is aware of the potential for difficulties. One is poised to observe mistakes, to make corrections. (This may happen to healthy agents in mundane cases as well, such as descending stairs in traffic, or descending stairs the height of which is unfamiliar.) It may happily be that no difficulties are encountered. One descends the steps exactly as planned. As one does, one knows what one is doing. One not only knows that one is trying to descend the stairs. One knows that one is descending the stairs. An inference based upon intention plus favorable background beliefs does not seem to capture this.[4]

None of this is to deny that intention, given the right background conditions, can be an important epistemic resource for agents, and can generate knowledge of action. My contention is that intention cannot be the only

[4] Kieran Setiya also argues that agents come by their knowledge of A-ing via their intentions. But Setiya does not lean on inference. And he emphasizes slightly different background conditions. For Setiya (2008), agents must know how to A, and must know they are able to A. These features justify a decision—the formation of an intention to A, and a corresponding belief. How might this get us all the way to knowledge of action?

In response to the trenchant criticisms of Sarah Paul (2009b), Setiya (2009) considers a case in which an agent is Φ-ing by performing basic intentional actions ABC. Setiya claims the agent knows she is Φ-ing by knowing she is ABC-ing. So we have to ask how an agent can (for Setiya: non-inferentially, non-observationally) know she is ABC-ing. Setiya says this:

> How do I know that I am doing ABC? We can say at least this. If I know how to take those basic means, this knowledge consists in the disposition to execute the corresponding intentions. Since I have this disposition, it is no accident that, when I intend and thus believe that I am doing ABC, I am doing so in fact. Exercising basic knowledge how ensures non-accidentally true belief. To say this is not to endorse an epistemology on which its being no accident that a belief is true suffices for it to count as knowledge. But it does pre-empt a residual source of skepticism, that the truth of beliefs formed without sufficient prior evidence could only be a matter of luck. When they are constituted by intentions and one knows how to perform the relevant actions, that is not the case. (Setiya 2009: 136)

When I know how to ABC, the belief that I am ABC-ing can be non-accidentally true because the knowledge how disposes me to execute the intentions that constitute the relevant beliefs.

Against accounts like Setiya's, Schwenkler argues that "it is hard to see how [such accounts] can work in cases where failure is quite likely, e.g. ones involving especially hard-to-perform actions, challenging conditions, extremely weak-willed agents who tend not to follow through on their intentions, and so on" (Schwenkler 2012: 733). Since agents who act successfully in difficult circumstances do often have knowledge of their actions, it looks like Setiya's account leaves out important cases in which we know what we are doing.

epistemic resource available to acting agents. Nor can it undergird the account of knowledgeable action I introduce below.

In any case intention-based accounts are in a way unnatural. They have difficulty capturing our non-observational knowledge of what is actually happening—of how the action is actually unfolding. The agent who knows she is A-ing by intention alone remains very much in the dark.[5]

8.5 Perception and its Limitations

If intention-based accounts leave agents in the dark, it is natural to wonder if perception provides a needed supplement. Most philosophers think not. Perception, they think, can be at best an enabling condition, and not a part of any justification, for knowledge of action (Falvey 2000, 32). John Schwenkler (2011, 2012, 2015) disagrees, claiming the following:

> There seem to be cases where the role of perception in agent's knowledge is plainly justificatory, as evidenced by the way perception would be cited in answer to a question of the form "How do you know that you are X-ing (say, replenishing the water supply of the house)?"—"Because I can see that I am," the agent might reply, thereby identifying perception as the justificatory ground of his or her knowledge, and not just something that enables him or her to know this thing non-empirically. (2015: 26, fn. 63)

To understand more fully Schwenkler's position, I should bring out an element of his broader picture. Schwenkler grounds the formal causal

[5] The same criticism applies to Thor Grünbaum's (2013) interesting discussion of a role for perceptual demonstratives in the knowledge of action. This is because perception's role, for Grünbaum, is to help specify the content of an intention:

> Usually when acting manipulatively on objects in our environment, intentions which make general or descriptive reference to objects of action are often not specific enough to initiate and guide our actions. Intentions to throw the paper in the waste basket or to write one's signature on some form are not specific enough. The agent needs to know which particular object is the basket or form in question before she can initiate her action. We can think of perceptual attention as having the role of selecting the right object in the world and thereby specifying the prior intention. The role of attention is to transform the prior intention (too unspecific to initiate and guide action) into a demonstrative intention: the intention to φ with this o. (Grünbaum 2013: 301)

role of knowledge emphasized in some interpretations of Anscombe (e.g., Moran 2004) in its efficient causal role. Consider the following passage:

> To explain what makes agential knowledge practical it is not enough just to say that this knowledge is 'what gives the descriptions under which what is going on is the execution of an intention' (Anscombe 2000: 87)—or rather, we seem not to have explained how our movements ever could be so described, until we explain how the knowledge of what one is intentionally doing makes a difference to the character of these movements themselves, and is therefore a knowledge of oneself as an agent, and not a merely passive subject of motion. (2015: 9)

In setting out his account of practical knowledge, Schwenkler is also diagnosing a flaw in Anscombe's. While sufficiently powerful agents need not check in on the progress of their actions, human agents can do very little without perception. Anscombe is aware of this, but argues that perception is often a mere aid for action execution, not an indispensable basis for knowledge of action. But, as the failures of intention-based accounts make clear, it does not look like perception is a mere aid. Knowledge of what a human agent is actually doing appears to require information that could only be gleaned from perception.

It is at this place that the grounding of knowledge's formal causal role in its efficient causal role does important work for Schwenkler. He wants to explain how knowledge of action could be genuinely practical and non-observational while being grounded in perception. He claims that in many cases "perception may serve as an 'aid' to action just by informing the agent of whether she is or is not acting as she means to: its guiding role is inseparable from its grounding role as a source of agential knowledge" (2015: 26). Ultimately, then, for Schwenkler knowledge grounded in perception "is not a mere 'monitoring' of an independent event, but an integral part of the process by which a person brings about the very action that it is a representation of" (2015: 27).

I like much in this view, but I have a few worries. First, such language suggests that Schwenkler thinks that perception represents what is happening as an action. It is debatable whether it does.

Second, even if perception sometimes represents bodily movements or other events as action, it is unclear that it does so in a way that could underwrite an agent's authority in knowledge of action, or an agent's claim to have non-observational knowledge of what she is doing. For if perception can

represent bodily movements as action, presumably it can represent your as well as my movements as action. If perception is the sole ground of non-observational knowledge of action, how is my knowledge of what I am doing distinct from my knowledge of what you are doing? The worry is that this account jettisons rather than explains knowledge's independence from observation, or an agent's authority in this knowledge.

Third, not all information drawn from perception could play a formal causal role in determining what one is doing, even if all information drawn from perception could, in principle, play an efficient causal role in influencing what one does. For consider the following four roles for perception in action. First, some information drawn from perception signals that one is succeeding in action. One kicks a football, feels the impact at just the right spot on one's boot, sees the ball take off at the right angle. One is doing as one intends. Second, some information drawn from perception signals that one is making a mistake in action. One kicks a football but feels one's leg out of place in the swing, hears something odd in the way the boot strikes the ball. Third, some information drawn from perception signals unpredicted effects of one's action. One kicks a football and notices a bystander watching its flight. Fourth, some information drawn from perception signals features of the world not immediately relevant to one's action. One kicks a football and notices storm clouds on the horizon.

While the first two types of information might play a formal role with respect to an experience of what one is doing—either succeeding or not (perfectly) succeeding—the second two types of information may play no such role. Certainly as one acts one continues to experience the world, and not every part of the world is relevant to what one is doing or to one's knowledge of what one is doing. The latter two kinds of information may in principle, however, contribute to how one acts, even though neither is represented or experienced as something the agent intentionally does.

Fourth, no one should deny that perception can add shape to action. What the agent does, and how she does it, cannot be fully explained in many cases without appeal to perception, and to perception-action and perception-cognition links. But what does knowledge have to do with this? Schwenkler wants to explain "how the knowledge of what one is intentionally doing makes a difference to the character of these movements themselves" (2015: 9). But the knowledge here seems superfluous. Perception does all of the work.

I return to this last worry below, where I argue, in effect, that perception can be important for knowledge of action—as well as for knowledge of how

to act, and for knowledgeable action—when and because the agent puts the relevant information to use in a certain way.

8.6 Conscious Awareness and its Limitations

8.6.1 Pickard and Bodily Experience

Hanna Pickard (2004) is motivated in part to explain what she calls an "overlooked feature of our knowledge" of action. She explains this feature as follows:

> Actions are events: they occur during periods of time. Knowledge of such events must be sensitive to their course through time. Part of what is distinctive about one's knowledge of one's own actions is the nature of this sensitivity. Not only does one normally believe that one is acting if one is, but if one believes that one is acting, then normally one is. One's knowledge keeps near perfect time with the actions themselves.
>
> (Pickard 2004: 206)

I agree. When acting we do not experience ourselves as acting under a loose description that involves only our intention and related beliefs. We experience ourselves acting—doing things in the world. Pickard is sensitive to this important phenomenon:

> Certainly we do seem to experience ourselves acting. If one is building something, it is not just that one knows what one intends, and can see what in fact results: intuitively, one can be aware of doing the building.
>
> (Pickard 2004: 212)

How best to capture this phenomenon? Pickard's proposal adverts to the content of bodily experience. For Pickard, "one can be aware of acting in virtue of being aware of one's body—that with which one acts—from the inside" (Pickard 2004: 212).

The claim is not that bodily experience contributes to one's awareness of acting, but rather that it grounds one's awareness of acting. Is this true? Certainly Anscombe seemed to reject such a view. For Anscombe, the sensations one has in one's body—sensations of pressure at various body points, or of stretching in various muscles, for example—are not sufficient

to explain the non-observational knowledge one has of the position of one's body parts. Anscombe seems to ignore the possibility that bodily experience could have higher-level content than such sensations, however. Pickard seeks to capitalize on this possibility. She asserts that not only does bodily experience have high-level content, it has content that describes the action one is performing at a time:

> We hear not just sounds but words; we feel not just an impenetrable surface, but a table. Given that body awareness is a form of perception, we should be willing to consider the possibility that its content is comparably rich: when one acts with one's body so as to pick up a cup, the content of the awareness one has of one's body is as of picking up a cup.
>
> (Pickard 2004: 220)

In order for bodily experience to ground knowledge of action, Pickard not only needs bodily experience to have rich (or so-called "high-level") contents (see Bayne 2009). She needs the contents to ground beliefs about action rather than about mere bodily movement. It would be more straightforward to understand how this happens, and more in line with Pickard's intentions here, if bodily experience presents the body as acting. Pickard makes the following claim: "one can be aware from the inside not just of one's arm rising, but of raising it" (Pickard 2004: 206). Later in the same paper she claims that "the awareness one has of one's body as acting not only facilitates the practical execution of action. It can also serve to sustain one's present-tensed knowledge that one is acting over the course of the action" (227).

Does bodily experience present the body as acting rather than as moving? This claim is difficult to assess. One reason for this is that bodily experience is typically accompanied by a range of other experiences, both perceptual and non-perceptual. One rarely if ever has an experience that only involves the elements that make up bodily experience. So it is difficult to find or to develop the kind of contrast that we need.

One might think that anarchic hand syndrome (AHS) could shed some light on Pickard's claim. Agents with AHS perform actions with an anarchic hand while denying that they are the driving force of the actions (Marchetti and Della Sala 1998). These agents still experience their hand as moving, but they do not experience its movements as their actions. One might think, then, that such cases are a problem for Pickard. But it remains possible that these agents experience their anarchic hand as acting, and that they distance themselves from the action for other reasons.

Even if this is right, however, the fact that anarchic hand patients commonly describe themselves as surprised by the actions of their anarchic hand indicates that the bodily experience accompanying the hand's movements is not sufficient for non-observational knowledge of action. In the agentive phenomenology literature, it is often claimed that the conscious performance of an action carries with it a sense of mineness. The way Horgan, Tienson, and Graham (2003) explicate this aspect, "You experience your arm, hand, and fingers as being moved by you yourself, rather than experiencing their motion either as fortuitously moving just as you want them to move or else as being transeuntly caused by your own mental states" (329). One problem raised by AHS is whether bodily experience is alone sufficient to capture this important aspect. Although forms of AHS can have sensory consequences (e.g., AHS that stems from damage to the parietal lobe (Jenkinson et al. 2015)), AHS is not specifically a deficit of bodily experience. This indicates that bodily experience is not sufficient for mineness phenomenology. If we wish to capture non-observational, practical, authoritative knowledge of action, bodily experience must at the very least be supplemented or integrated into other grounds.

8.6.2 Peacocke and Awareness of A-ing

Christopher Peacocke (1999, 2003) advances an account of knowledge of action on which an agent's conscious awareness of A-ing provides a reason for judging that she is A-ing.

Peacocke argues that conscious awareness has the right kind of content and the right kind of connection to an agent's judgments to provide the required link to knowledge. For when an agent takes elements of her conscious experience at face value, and her conscious experience is related to the world in the right way, the judgment that results is not only truth-conducive. In addition, it can "be philosophically explained in terms of the nature of the intentional contents and states involved in the transition" (2003: 52). Reliability or truth-conduciveness alone is not sufficient. Reliability must come about in virtue of the "nature of the intentional contents and states involved" (52).

I do not wish to disagree here with Peacocke's epistemic proposal. My worry concerns his account of the awareness of A-ing. I have criticized this account in an earlier paper (Shepherd 2017b). Here is the gist.

First, the experience of bodily action seems, phenomenologically, to involve perception to some extent. When we experience ourselves moving

our bodies, a feature of the experience is that the body is moving. We need not deny that illusions of movement can be caused by tryings in some cases. But it is implausible that a trying alone can account for the actional experience of the body moving.[6]

Second, Peacocke appeals to work by Anthony Marcel and colleagues to support his view (see Marcel 2003). I argue that the work does not, in truth, offer the required support. My earlier explanation is, I hope, sufficient:

> Marcel's basic result is this. Using vibro-tactile stimulation to the elbow, it is possible to put agents in a position in which they falsely experience their arm in one location and so intend to move their arm in one direction in order to accomplish a goal. But due to the illusion induced by stimulation, these agents must actually move their arm in the other direction to accomplish the goal. It turns out that agents are accurate in moving their arms in the correct direction. But 60–70 percent of agents in Marcel's study show no awareness of having moved in a direction opposite to their consciously reported intention. From this Peacocke concludes 'The content of their trying (or some event causally related to it) seems to cause the content of their impression of action, even though the actual motor instruction issued requires, and produces, movement in the opposite direction' (Peacocke 2003: 122). But it is not clear how strong this evidence is. For, as Marcel notes, it is possible that these agents had no awareness of the specific character of their arm movements: their strong belief that they were moving their arm in one direction might have overridden whatever arm-related experience

[6] I would also apply phenomenological complaints to Lucy O'Brien's (2009) account of knowledge of A-ing. Like Peacocke, O'Brien draws a link between awareness and knowledge. Her account of awareness is different. For O'Brien the awareness of A-ing is not caused by a trying to A. Rather, to consciously act is to act in "a certain way or mode"—a conscious way or mode. And this is to act "on the basis of an evaluation of the possibilities open to us" (O'Brien 2009: 184). When an agent acts on the basis of an evaluation of action possibilities, O'Brien claims she acts with a sense of control. O'Brien further claims that to act with a sense of control is to act with knowledge-enabling "agent's awareness":

> I act consciously when I am agent aware of my action, and…I am agent aware of my action when it is something I actively control. Our actions are those things we know, not by observing them, or by reflecting about them, or accepting some presentation of them, but rather by actively engaging in them. (O'Brien 2009: 183)

I do think agents sometimes act with knowledge-enabling agent's awareness. This is what was going on, I think, in my earlier stair-descent case. My chief problem with O'Brien's account is that it seems the wrong description of the contents of the experience of acting. The evaluation of action-possibilities takes place prior to the enacting of whatever possibility one settles upon. So if we want to explain our active engagement in the action, we need to say something about the content of the experience as we engage.

> was present, leading to their false reports of arm movement. Second, as Marcel also notes, 30–40 percent of participants did have conscious access to the conflict between intention and arm movement. Marcel reports that 'All of the latter subjects commented in one or another way that there was something peculiar, or that while they had intended to move in one direction they had the impression that they had moved in the other' (2003: 166). That the relationship between trying to A and having an experience of A-ing Peacocke asserts is not universal—even if it is present in most of the 60–70 percent of subjects in Marcel's study, a claim I have questioned—casts doubt on this as a general account of the experience of acting.
>
> (Shepherd 2017b: 428)

A further problem is that temporarily paralyzed agents can experience tryings without experiencing any associated bodily movements (Shepherd 2016). So it seems that this account of the experience of acting will not work.

8.6.3 My Suggestion and its Limitations

The account I offer of the experience of acting's contents appeals to work on informational integration in the etiology of our conscious experience. I argue that just as some of our perceptual experiences result from the integration of contents in separate modalities (see O'Callaghan 2015), such as vision and audition, our experiences of acting result from the integration of contents due to intention (and perhaps motor command) as well as various perceptual modalities.

I claim:

> The experience of acting's distinct unity stems at least in part from a sub-personal process of construction that integrates spatial and temporal contents drawn from agentive and perceptual processing and attributes the result to the same unfolding event: the action.
>
> (Shepherd 2017b: 443–4)

The upshot of this account—the "distinct unity" claim—is that the experience of acting is legitimately of acting. When consciously acting the agent is experiencing doing a thing—it is probably more accurate to say that she is experientially doing a thing—in virtue of the contents drawn from

intention and motor command, and the agent is experiencing doing a thing, in virtue of the information gleaned from perception.

One might run an epistemic angle on this proposal as Peacocke does. Or one could try out a kind of dogmatism (a la Pryor 2000), as I suggest in (Shepherd 2017b). This can, I believe, get you to knowledge of action in some cases—the cases that involve this kind of experience of acting. So we now have at least two sources that may very well generate epistemically special knowledge of action—intention and the experience of acting.

But this is not the full story. For not every case of action involves the experience of acting. Sometimes this is because the kind of action at issue does not involve the right kind of intention-perception links. Sometimes it is because consciousness is not involved in the action in the right kind of way. So the knowledge one gets via this route is, again, restricted.

And there is a further issue, somewhat separate from the epistemic account of the knowledge of action. I am ultimately interested in knowledgeable action. And the role of knowledge in action is broader than these sources could indicate.

Notice: the knowledge acquired via intention or the experience of acting seems otiose for the purposes of action execution. The relevant work is done by states and processes—intention, perception, experiences—that do not amount to knowledge. What would be good, I argue presently, is if an agent could use experiences of acting, along with other resources at her disposal, to come by a knowledge that contributes to her exercise of control in acting.

8.7 Practical Reasoning in Action

Here's the big secret. An agent has knowledge of action that she can use—knowledge of A that contributes to the A-ing, as well as knowledge of how to A—when she is poised to engage in practical reasoning in action.[7] This covers knowledge of A-ing, as well as knowledge of how to A.

Practical reasoning in action is the familiar notion of practical reasoning, applied to and during action. I do not endorse a specific account of practical

[7] Rödl (2011) develops an Anscombian view of intentional action that depends upon claims about practical reasoning and knowledge of action. There are some affinities of form between his claims and mine, but important differences exist. Mine is not an account of intentional action, but of intentional action that uniquely involves knowledge. Further, my account is liberal about the structure of practical reasoning—no particular claims about the nature or form of the premises or conclusion of practical reasoning are assumed.

reasoning, or of reasoning, here. In general, practical reasoning is a response to uncertainty about some element of the action: about whether to behave in a certain way, whether to continue one's plan or modify or abort, or about whether one can succeed exactly as planned (Camp 2009, Shepherd 2015a). It is then a set of psychological processes directed towards figuring out whether one is succeeding, or whether one can succeed, or what the right answers to one's questions about whether and how to act might be.

So, as one slowly descends the stairs, one's heel may clip a stair leading to a lurching feeling. This may lead to an automatic reaction, to be sure. But in some cases, it will lead to a rapid form of reasoning. One will envision the possibility of gripping the rail harder, and then envision the possibility of halting downward progress by leaning back or sitting down, and then one may envision the possibility of skipping the next step to better land on the step just beyond. Very rapidly, one is calibrating levels of confidence in these possibilities, and the risk of failing at any of them. One may judge that the best move is the least risky, and decide to sit down. As Marúsič and Schwenkler (2018) put it, "practical reasoning is the reasoning by which an agent determines what she is going to do, insofar as what she will do is up to her. That is, insofar as matters are up to the agent, she is in a position to determine how they will unfold, and she does this through her practical reasoning" (310).

It is important to note that I do not require that practical reasoning be actually engaged—that the agent be actively figuring out how to act. I require that the agent be poised to engage in practical reasoning. The poise I have in mind entails activated beliefs about what the agent is doing. This is because in some cases, agents anticipate trouble. Or perhaps agents simply wish to act in a way that involves heightened awareness (where this awareness entails poise to engage in practical reasoning). In such cases agents act with poise. This is a state of readiness that involves awareness of what is going on related to the action, often leading to pick-up of knowledge via various available sources where the pick-up is a part of a disposition to use what is picked up. Poise can be triggered by registration of difficulty, leading to practical reasoning. But it can be in place beforehand as well, before practical reasoning is actually engaged.

The set of psychological processes that qualify as practical reasoning are, in my view, diverse. Some views of reasoning that emphasize attitudes towards complex propositions "in which propositions may themselves figure as objects of which properties and relations are predicated" (Pettit 2007: 498). Such attitudes may play important roles in some circumstances. Tyler Burge speaks of agent's deploying higher-order attitudes in order to "take

rational control of one's reasoning" (Burge 2013: 74). These are useful tricks for limited agents, like human agents. Given limitations in knowledge and functioning we sometimes need to explicitly review our first-order reasoning. Of course there is rarely time to do so. So in many cases our practical reasoning does not take this form.

I noted in chapter 2 that the philosophy of cognitive science identifies a number of ways an agent's practical reasoning may involve (without, perhaps, being fully exhausted by) non-propositionally structured representational states: map-like representations (Camp 2007), or analogue magnitude representations (Beck 2014), or mental imagery (Gauker 2011), or combinations of these (Shepherd 2018a).[8] In other work I have argued that while reasoning via such states may limit agents in some ways, there may be important benefits as well. In skill development agents navigate a tension between efficiency and flexibility—automatization increases efficiency (in the good case), while decreasing flexibility and sensitivity to reasons (Shepherd forthcoming-c). Something similar is true of reasoning via non-propositionally structured states. The types of transitions afforded by map-like or imagistic states may be limited. But the manner of transitions afforded may be just what the agent needs. Some formats are, in some circumstances, or embedded in some systems, more usable than propositionally structured states (Camp 2007; Shepherd 2018a).

Consider the use of what Robert Briscoe (2008) calls "make-perceive." Agents often voluntarily or intentionally generate mental imagery in order to augment the current experiential field or to adumbrate nearby possibilities. This imagery may be motoric or involve various perceptual modalities, depending on the needs of the agent. Developing a facility with this capacity is extraordinarily helpful for the smooth control of action. Briscoe (2018) considers the example of a rock-climber. When climbing a wall, a climber needs to execute several sub-actions that balance accuracy of reach, distribution of weight, strength of grip, position of the feet, accessibility and climbability of upcoming patches of rock, and so on. Skilled rock climbers rely on enhanced perceptual recognition of elements on the rock wall, and enhanced abilities to deploy motor and perceptual imagery as they work through the best ways to execute the action of climbing the wall (see Pezzulo et al. 2010). Plausibly, doing so involves assessments of relations of reason

[8] I suppose these states may relate the agent to propositions, without themselves being propositionally structured. Whether one cares will likely depend upon how invested one is in certain proposition-involving construals of knowledge how.

between ways of executing the action, goals, one's own abilities, and available physical resources.

Notice that the involvement of knowledge in these instances of reasoning may depend upon context- or domain-specific forms of control or skill. Katalin Farkas (2017) has recently emphasized just how context-dependent much of the knowledge driving action may be. In particular, Farkas argues that many of the answers to embedded questions that constitute "knowledge-wh" may be available only "through, or with the aid of, a complex sensory experience" (863). (A quick aside: it is fine to think of my "knowledge of how to A" as a kind of knowledge-wh—knowledge of what to do, when to do it, and so on, are ways of knowing how to A.) To illustrate:

> When the sailor knows she can lean out this far, the content involves a proprioceptive presentation. If I pointed at her and said 'one can lean out that much,' I would not manifest the relevant knowledge-wh, because the content of my knowledge would involve a visual, third person, rather than a proprioceptive presentation. It seems very likely that the ability to perform the maneuver, at least in some cases, requires a proprioceptive presentation. If so, then my knowledge may be sufficient for some purpose, but would not be sufficient for the purpose of performing the maneuver. The reason is that the extent one can lean out would not be presented under the right mode of presentation. (Farkas 2017: 864)

In applying Farkas's insights to knowledge of (bodily) action, I would emphasize an addition. First, it is plausible that much of the knowledge relevant to action control relies on content given not just in sensory, but in sensorimotor experiences—experiences of acting. So we highlight an additional way that conscious experience is relevant to the knowledge of action, even though it cannot be the full story. The conscious experiences one has while acting not only furnish one with materials relevant to one's activity of practical reasoning. In addition, epistemically relevant properties of the action itself are given in a way that only the relevant experience can provide.

Now, practical reasoning in action embeds and depends upon judgments about what one is doing, and how one is doing it. One engages in practical reasoning usually because uncertainty arises about whether what one is doing will lead to success, or whether one's plan for how to do so will lead to success. The engagement of practical reasoning in action gains structure from the action one is executing. In such reasoning, one is seeking to answer questions about whether to proceed as planned (as intended), or about how best to change the plan in light of extenuating circumstances. Sometimes

one is engaging in a form of practical reasoning that is more like monitoring, or waiting: the agent is asking whether she will pull off the next step as planned, and she is poised to intervene in case of a mistake. In all of this various judgments about what one is doing, and how, are more or less taken for granted. But they are there. They set the background against which one reasons. One is reasoning about the best way to A, or about whether to continue A-ing. Such reasoning embeds judgments that one is A-ing, and often embeds judgments about how the A-ing is going at a time, at a relatively fine grain of detail.

These judgments about what one is doing, about how one's action is progressing, may be informed by a range of sources—one's intention, elements of perception, one's experience of acting. But the already-discussed accounts of the epistemic credentials of knowledge via these sources do not transfer. When one acquires knowledge of what one is doing due to a process of practical reasoning, the epistemic account is importantly distinct.

8.7.1 The Epistemic Account

When an agent is engaged in practical reasoning, she is making a series of judgments. These judgments form beliefs. We care in this instance about the beliefs about what the agent is doing, and about how the agent is doing it or will be doing it. These beliefs amount to knowledge when, and because, the agent has a sufficient degree of control over her practical reasoning.

What is it to have control over practical reasoning? The basic features of the account of control developed in chapter two apply. But there I was concerned with control as indexed to an agent's plans, or to her goals. By now we have a sense of the broader spaces in which agents move. So we understand, for example, that agents can be assessed according to goals that they do not currently possess—goals proprietary to action domains. A more expansive anchor for assessing control over practical reasoning is possible.

An agent has a sufficient degree of control over her practical reasoning in action when her reasoning, in the relevant circumstances, reliably finds ways to successfully A, or, if the circumstances are genuinely not favorable to A-ing, finds the best available options (given background goals and preferences and additional goals accessible to the agent). Often action circumstances afford multiple actions, and sometimes practical reasoning suggests a way to satisfy background goals or desires, or to promote background values, by updating or changing existing plans. Agents may reason their way to very different courses of action, but this may plausibly be considered a

success if the result is preferable to the possible world in which the agent did not engage in reasoning, given the agent's pre-existing plans, preferences, and commitments, or given the action domains in which she is currently engaged.

I said the agent's beliefs about what she is doing and how amount to knowledge when, and because, the beliefs result from a process of sufficiently controlled practical reasoning. Whence comes the because?

When an agent has sufficient degree of control with respect to her practical reasoning in action, in the relevant circumstances and the action domain they fall under, and when she is exercising this control, then she has entitlement to the judgments she makes while conducting that process.

Entitlement is a kind of warrant, an epistemic status a cognitive state can possess that, given the right additional features (about which I'm here going to be neutral), can contribute to that state's qualifying as knowledge.[9] Entitlement differs from justification—another type of warrant—in the following way. While justification requires a subject to be able to articulate the reasons she has for some belief or judgment, entitlement does not require a subject to be able to articulate or conceptualize her reasons for a belief or judgment. All that is required for entitlement is that a process that confers the status of entitlement provides a good route to the truth of the entitled state. So, for example, Burge writes that our entitlement should be "connected to the truth of our judgments about our thoughts" in a non-accidental way (1996: 103). And Peacocke writes that an account of entitlement must show how it is "reasonable to employ the ways of reaching beliefs that the theory describes as entitling, and those ways must be appropriately connected with the truth of the beliefs reached in those ways" (1996: 117–18).

The way I have set this up, we have a good case and a bad case. In the good case, the agent has sufficient control over her behavior, and she exercises sufficient control, by way of having and exercising sufficient control with respect to practical reasoning. In the good case the agent has entitlement, and because of that, her judgments amount to knowledge. The agent only has entitlement in the good case.[10]

[9] So, Tyler Burge says of warrant that "Being epistemically warranted in having a belief is having the belief in a way that is good for having true beliefs, given limitations on one's information and cognitive capacities…if one is warranted, one's belief is held through a natural competence that is epistemically good—conducive to the belief's being true and, usually, to the belief's constituting knowledge" (2013: 489).

[10] Readers might note similarities between my account of entitlement and Burge's account of entitlement for self-knowledge of attitudes like belief, desire, and intention. The accounts are similar in some ways, but distinct in others. Burge's applies primarily to knowledge of the

In the bad case, the agent lacks sufficient control over behavior. (This entails that she will not exercise sufficient control.) This may be because he lacks sufficient control regarding what he intends to do. Or the problem could lie somewhere in his cognitive apparatus. Either way, practical reasoning in those circumstances would not be a good route to finding ways to A. Nor would the agent's judgments about what they are doing amount to knowledge.[11]

For example, if you were to place me on the television show *Dancing with the Stars*, I would be very uncertain about how to perform. In my case, practical reasoning would not be a rational response to uncertainty. For I am so bad at formal dancing, and I know so little about it, that my practical reasoning is unlikely to turn up any solution that could guide me away from uncertainty and towards success.[12] At this low level of control, I am unlikely to judge correctly regarding what I am (intentionally) doing. For I am doing little except for flailing about, and that unintentionally.

8.8 Knowledgeable Action

Knowledgeable action is a subset of intentional action. It is intentional action guided by knowledge of A-ing, and knowledge of how to A, where the knowledge states qualify as knowledge in virtue of the agent's entitlement to her judgments, the entitlement comes about in virtue of the agent's

attitudes; mine to knowledge of action as well as to knowledge of ways of acting. Burge rejects the thought that the entitlement obtains in virtue of "the reliability of some causal-perceptual relation between cognition and its object" (1996: 98). He emphasizes instead the role of an agent's judgments in critical reasoning, and a constitutive relation between "the judgments about one's thoughts and the judgments being true" (98). I actually emphasize the possibility of such a constitutive relation (see section 8.8.1). But it obtains in virtue of control, which is understood, broadly, as involving a reliable causal relationship between the agent's practical reasoning and her behavior. So I doubt Burge would accept my account, even though I think it could be used—via the account of control—to illuminate self-knowledge of the attitudes as well. For Burge's argument relies on the idea that critical reasoning, which involves knowledge, is a kind of rational control over reasoning. And one could understand rational control as control over reasoning that, in virtue of the entitlement the control provides, involves knowledge of relevant reasons.

[11] At least, not because of the entitlement I am here articulating. Perhaps the agent could infer from intention or perception what she is up to—likely, failing, or only trying to A—but in such a case, the knowledge of trying to A would not come from the agents control regarding practical reasoning.

[12] I would need a good teacher, and the absence of an audience—maybe under a bridge somewhere—before I made any progress.

control over practical reasoning plus behavior, and the control over practical reasoning plus behavior explains the fact that the knowledge guides the agent's action.

8.8.1 The Radical Case

This account allows for a radical case in which the agent becomes uncertain about her ability to succeed, begins to engage in practical reasoning, makes a judgment that is disconnected from her prior evidence about what she was doing or trying to do, and so now holds a belief that she is A-ing in way X (when before she was A-ing in way Y, or B-ing in way Z). The case is radical because it allows that if the agent has control over her A-ing in way X in virtue of having control over the practical reasoning that can direct her A-ing in way X, she now knows that she is A-ing, and she knows that X is a way to do so, and the way she is doing so. And this in spite of whatever she was doing before the judgment.

So it goes. I think in such a case the agent redirects herself—when she judges she is A-ing in way X, that's what she is then doing, in virtue of the judgment. Such a judgment need not be a part of her intention. Instead, what happens is that the judgment causes a relevant intention to A in way X. Of course the connection is not foolproof. But control over practical reasoning is in part constituted by these kinds of connections. In the good case—the case in which the agent exercises sufficient control—such a judgment causes a corresponding intention. If such a link fails, the agent is not exercising sufficient control over practical reasoning, and the entitlement disappears.

8.9 Knowledgeable Action and Related Proposals

I wish to draw some connections, and note some departures, from related work that has helped me to see the relevant issues more clearly.[13]

Imogen Dickie (2015, building upon her 2012) offers an account of practical knowledge. She begins with an example evocative of the problem

[13] I came too late upon Kim Frost's (2019) view (based upon a reading of Anscombe) that practical knowledge can be understood as a kind of capacity. This view seems to bake control into exercises of the capacity for practical knowledge in a way that I find interesting. I hope to address this kind of view in later work.

facing intention-based accounts of knowledge of action. Robin Hood forms an intention to hit a target some distance off. Does Robin know that he is going to hit the target? We could extend this to the timespan of action—as he is shooting, does Robin know that he is hitting the target? Dickie says yes, and her explanation reverses the traditional order of explanation from truthmaker to knowledge-qualifying state. (Dickie speaks of the relevant state as both a belief and the cognitive commitment component of an intention.) She comments:

> In this case, the status of the belief as more than just luckily true is not secured by any role the truthmaker plays in the belief's formation. It is secured by the role the belief plays in bringing about the truthmaker. Given Robin's skill, when he commits himself to hitting the target, it is not just a matter of luck that this commitment ends up fulfilled… Rather, Robin's skill enables his intention to be a non-lucky generator of its own fulfilment… So whether an intention counts as knowledge does not depend on how it is formed. It depends on whether it plays an appropriate role in generating its fulfilment. (Dickie 2015: 89)

Dickie is aware that this account depends upon some account of the "appropriate role." She doesn't wish to commit to one. However, given similarities between her account of skill, and my account of control,[14] I would offer my account of non-deviant causation as a useful strut. (I don't know whether Dickie would accept this.)

I like Dickie's idea of appealing to the agent's control (Dickie says skill) to explain why some relevant state qualifies as knowledge. I have used this idea, but with a different target. Dickie's concern here is the attitude of cognitive commitment to some plan. That's either an intention, or a component of some intention. Such a state may qualify as knowledge (I'm neutral), but even if so, it's role as knowledge seems secondary to its role in action execution. For Dickie, controlled action manifests knowledge. But the knowledge does not do work in explaining the agent's success—it is rather explained by her capacity to succeed. (This point leads Habgood-Coote to complain that

[14] Here is how Dickie states things in her 2012: "S is a skilled Φ-er iff, for appropriate range Σ of situations r, in most cases, if S were to intend to Φ in r, S's intention would lead S to act in some [way] $w \in f(r)$, where f is a function taking each r to the set of reliable ways for S to Φ in r" (Dickie 2012: 739) What Dickie is calling skill I would call control. There are slight differences, and potential disagreements in how we would fill out "appropriate range of situations," and in how we would understand the role of intention.

"Dickie's view makes knowledge-how into an epiphenomenal by-product of skill" (2019: 14).)

I am interested in a broader range of states. When engaged in practical reasoning, the agent's action not only expresses knowledge (if it does), it embeds knowledge of action and knowledge of how to act, which then plays a role in the control of that very action. That's one thing I wish to capture. Furthermore, when engaged in practical reasoning, the agent must not simply look for the best ways to execute her intention. They must also remain open to radical changes of plan—changing or dropping intentions, for example. Such may be the demands of practical rationality, or the force of the agent's reasons, in a situation. I also wish to capture this additional layer of possibility. This is why we need a broader notion of controlled practical reasoning, as opposed simply to a notion of control at executing a plan-state.

I apply Dickie's idea that the explanation of the knowledge involved in action runs through the agent's control. In my case the knowledge involved is broader, and the control involved is different.

I turn to a recent discussion of knowledge how by Joshua Habgood-Coote. Habgood-Coote offers an account of knowledge how called the interrogative capacity view:

> For any context c, subject S, and activity V, an utterance of 'S knows how to V' (in its practical-knowledge ascribing sense) is true in c iff c has associated with it a set of practically relevant situations {F1, F2, . . . }, and, for all (or at least most) Fi that are members of {F1, F2, . . . }, S has the capacity to activate knowledge of a fine-grained answer to the question, how to V in Fi?, in the process of V-ing. (Habgood-Coote 2019: 92)

The capacity to activate knowledge is, for Habgood-Coote, in a way disjunctive. An agent might activate knowledge by learning new information, or she might exercise pre-existing knowledge. And there is an additional layer of complexity. For the agent might activate knowledge of how to V in the process of V-ing, Habgood-Coote states, simply by V-ing in some specific way. We are told that "Merely being able to think an answer to a question is not enough for knowing how" (91). The knowledge and the action are intertwined. "The ability to answer the question how to V? on the fly is at the same time an ability to answer the question how to V? by doing V, and an adverbial ability to V by answering the question how to V?" (91).

Here is one way to see this. Habgood-Coote has offered a sufficient condition for the ability to V. There may be other ways to possess this ability,

but this one involves the ability to V in a specific, knowledge-involving way. Habgood-Coote offers this as an account of knowledge how, but how it fares in the conceptual space of the knowledge how debates is not my concern. For this also sounds (to me) like a covert account of knowledgeable action.

I said knowledgeable action is intentional action guided by knowledge of A-ing, and knowledge of how to A. The crucial part of my account is the explanation of how knowledge gets involved in the action, and of why it is specifically knowledge that is involved. I think this is just what Habgood-Coote's account needs, were we to think of it as an account of knowledgeable action. Early in his paper the question of knowledge arises. Habgood-Coote notices that one could find the right answer to a question of how to V deviantly (by guessing, for example). He proposes to put the issue to the side, instead singling out "the kind of answering that involves getting to the correct answer in the epistemically right way" (89). What way is this? Habgood-Coote writes "I will gloss this as the ability to know answers, since knowing an answer entails having got to the correct answer in the epistemically right (non-Gettiered, non-lucky) way" (89).

I'm aware that a man with a hammer sees too many nails. But again we see the importance of an account of non-deviant causation. In explicating knowledgeable action, such an account is crucial in being able to explain the controlled practical reasoning that provides entitlement to knowledge of A-ing and of how to A. Without it, we are forced to assume knowledge is involved, without an explanation. But an explanation is important, since agents are wily things and may very well find shortcuts and tricks enabling skill in the absence of knowledge. If it is knowledge specifically that we are interested in, it helps to understand how this knowledge comes about, and thereby helps to constitute this particular mode of agentive excellence.

8.10 Conclusion: Knowledgeable Action and Agentive Excellence

I have articulated a view on which the engagement of controlled practical reasoning in action embeds knowledge-constituting judgments regarding what the agent is doing, and how. The knowledge of action plays an important role in a process of rational control over that very action. I call action that involves knowledge of action in this way knowledgeable action. Knowledgeable action is a mode of agentive excellence. For in acting knowledgeably, the agent extends the reach of her control by way of her

controlled use of practical reasoning. She finds and implements good ways to execute her intentions (or desires or whatever), or to satisfy nearby background goals. The ways she finds, when she engages in controlled practical reasoning, are ways she might miss if she depended upon rigid routines, or upon whatever practice in the relevant domain might instill in her. This use of practical reasoning permits the exercise of control even in difficult circumstances—circumstances that render the agent uncertain, leading to practical reasoning.

There is a connection here with Anscombe's well-known claim that knowledge of action is in some sense practical as opposed to speculative. Here is Anscombe's central example of the practicality of an agent's knowledge:

> Imagine someone directing a project, like the erection of a building which he cannot see and does not get reports on, purely by giving orders. His imagination (evidently a superhuman one) takes the place of perception that would ordinarily be employed by the director of such a project. He is not like a man merely considering speculatively how a thing might be done; such a man can leave many points unsettled, but this man must settle everything in a right order. His knowledge of what is done is practical knowledge. (Anscombe 2000: 82)

For Anscombe, it seems that an agent's knowledge is practical at least in part because the agent relies on this knowledge to give his orders and to settle everything in the right order. So the agent knowledgeably determines the details of his project's execution.[15]

What I have done in this chapter is offer a fuller account of the way that knowledge is involved in the agent's execution of action, and of why it is knowledge that is involved, as opposed to something else. Knowledgeable action is a subset of intentional action: Anscombe over-generalized from the knowledge-involving cases. It is in knowledgeable action that the agent

[15] I am not here attempting to make full sense of the things Anscombe says regarding knowledge's practicality. Many have found Anscombe's view on these matters puzzling. Some representative remarks: "Despite its centrality to Anscombe's project, her concept of practical knowledge has not been widely understood" (Schwenkler 2015: 1). "Anscombe's account of 'practical knowledge' is hard to interpret" (Candlish and Damnjanovic 2013: 700) [Anscombe's view of practical knowledge is] "not just causally perverse but epistemically mysterious" (Velleman 2007: 103). "Anscombe is right to claim that there is something special about knowledge of what one is doing intentionally; but her explanation is not particularly helpful" (Setiya 2008: 392).

is well-placed to answer why-questions about what is going on. For in these cases the agents are posing those same questions to themselves, and answering them.

What distinguishes the actions that involve knowledge from those that need not? I suggest that the distinction can be drawn in terms of a requirement of practical reasoning for the execution of a certain (large) class of actions (and this requirement in turn seems to track a requirement upon flexibility in action execution). This requirement may stem from the complexity inherent in an action, or from difficulties inherent in the environment within which an agent acts. Either source may generate difficulties that agents need to work through as they carry out an action plan. When agents need to work through difficulties, they begin to reason practically, and at this point knowledge is generated and engaged.

This is an empirical claim. I and others have defended it (Shepherd 2015b; Christensen et al. 2016). Agents often need to engage practical reasoning at some point as they execute plans. This is a function of the complexity of human plans, the human behavioral control apparatus, and the nature of the circumstances in which humans operate. The result is that knowledgeable action, while a subset of intentional action, is nonetheless pervasive.

This is a different explanation for the pervasiveness of knowledge of action than one offered by Velleman (2007). Velleman maintains that agents have strong and persistent ("intrinsic") desires to know what they are doing and why they are doing it, and that (roughly) these desires motivate them to deliberate and act in ways that generate this knowledge. My explanation is simpler. Whether agents have such desires or not—and whether these desires play a crucial role in deliberation and action or not—I claim that knowledge of action is pervasive because execution of action usually requires the practical reasoning that generates it.

9
Conclusion

How these things end. In Cormac McCarthy's novel *Blood Meridian*, some young recruits in a militia of questionable legality spend an evening drinking and carousing. One of them ends up dead. As they observe the body in the early morning, an old Mennonite whom they had, while carousing, berated, joins the observation. McCarthy has him comment:

> There is no such joy in the tavern as upon the road thereto, said the Mennonite. He had been holding his hat in his hands and now he set it upon his head again and turned and went out the gate. (McCarthy 2001: 43)

We're off the road to the tavern now. The book is basically done. One hardly knows whether to let sound the mournful, almost disappointed tones of Pearl Jam's "Long Road"—*but still something's missing… cannot say*—or to play instead the triumphalist conclusion of Haydn's "The Creation":

Vollendet ist das große Werk!

We should take a step back, and consider the various claims and accounts offered in this book in broader context. I began by invoking a contrast one often finds in the philosophy of action, between activity attributable to an agent and a plane of mere passivity. I said I had an aim to explain this metaphor's allure. Agents behave differently than non-agents. Something is unique and special about agents.

I tried to capture this difference—to develop a perspective on the shape of agency—by way of interlinked accounts of key (one might say pivotal) agential phenomena. Three of these are basic building blocks of agency. Two express modes of agentive excellence.

The basic building blocks are control, non-deviance, and intentional action.

Control is essentially a matter of an agent disposed, in certain contexts or circumstances, to behave in a certain way. The way the controlled agent

The Shape of Agency: Control, Action, Skill, Knowledge. Joshua Shepherd, Oxford University Press (2021).
 DOI: 10.1093/oso/9780198866411.003.0009

behaves—the way her dispositions manifest in the relevant circumstances—is such that they repeatably and flexibly produce situations (movements, thoughts, or events more broadly) that match their plan for behavior.

The account I offered delved, in different ways, into details that are important for understanding what I mean by talking of repeatable, flexible behavioral approximation to the content of a plan. The notion of a plan state is important. So, too, is the notion of a circumstance-type. For an agent's control cannot be understood in isolation from the circumstances in which she possesses or exercises the control.

Nor can an agent's control be understood in isolation from non-deviant causation. For non-deviance is, in essence, the expression of the control that an agent possesses. In order to understand what happens when non-deviant causation happens, I developed the notion of a comprehensive set of circumstances. This is a set of circumstances that is derived by building a causal model that includes an agent, a plan, and the agent's location in a particular situation. What is special about the model is that it gets the causal parameters of the particular situation right. Non-deviant causation then turns out to be the normal production of behavior that is, for the agent, normal given the plan and across the comprehensive set of circumstances.

After discussing non-deviant causation, I turned to a specification and to a discussion of different varietals of control. For it should be the case that a satisfying account of control has the potential to shed light on different usages of control. I focused on two, and argued that the account of control I offer can be extended to illuminate direct control as well as voluntary control. I also, probably foolhardily, offered an explication of what is "up to" an agent that my friends who think about free will may read, and quickly ignore.

With accounts of control and non-deviance in hand, I offered a new account of intentional action. That chapter may seem, in retrospect, to be particularly rough sledding. But the basic idea is not. Intentional action is, in essence, the exercise of a sufficient degree of control in bringing behavior to approximate a good plan. The relative simplicity of the account should appeal to causalists, and the ability of the account to handle common objections to causalist accounts should dismay anti-causalists. In broader perspective, I think the account locks onto something of importance. The divide between activity and passivity is metaphorical. But the divide between systems incapable and those capable of intentional action is not. This account

locates the elements—namely, control and capacities to form good plans in at least some circumstance-types—that make the divide precise.

After specifying these basic building blocks I articulated a view of agency on which agents contain capacities that warrant the application of behavioral standards—usually, practically rational norms, or practical reasons—and that enable the meeting of at least some of these standards. I did not offer an analysis of agency. In chapter 6, where I discussed the nature of agency, I was deliberately non-committal in a way that I was not in other chapters. What I wanted to do was to lay out a space of different forms of agents that could serve as a hinge on which the book might turn, towards a discussion of modes of agentive excellence.

The modes of excellence I discussed enable the expansion of the space of behavioral standards in some ways, and the meeting of a wide range of standards, in others. Skilled agents structure themselves so that the standards that apply to their performance become very fine-grained. They display excellence so refined that at times it seems skilled behavior is its own reward—excellence for excellence's sake, along whatever dimension of evaluation.

Knowledgeable action is a different mode of excellence. By way of practical reasoning, sophisticated agents use knowledge of what they are doing to find ever better ways of doing it—or, failing that, to find ways to stop doing that and to do something better, or next best, as the case dictates.

I began with a rough contrast one often finds in the philosophy of action, between the activity of an agent and a plane of mere passivity. The metaphor is alluring, I think, because agents contain, to greater or lesser degrees, capacities that enable the transformation of the world by way of standards that the agent sets—by way of plans, and goals embedded in plans. When an agent can do this across even minute differences of circumstance, then it seems the result is due to something stable, something capable of transmuting the flux of stimulus and response into something more than just the ongoing flux of stimulus and response—something that combines reliability and flexibility in a way suggestive of something behind the flux. Something like a mind, or at least a plan. With the exercise of control, the passive becomes active, and plans give rise to intentional action. With the development of skill, capacities for planning and for exercising control and for executing intentional action begin to cover broad differences of circumstance. With the acquisition of knowledge, agents are able to impose their will on parts of the world that their practice may not have adequately prepared

them for. In knowledgeable action, agents exert change in the world in part by figuring out the world they change.

Well, look at me. I'm rambling. The story is becoming excessively metaphorical at this later stage. But, like the Stranger says at the conclusion of *The Big Lebowski* (Coen and Coen 2009), it was a pretty good story, don't you think? It made me laugh to beat the band.

Parts, anyway.

References

Adams, F. (1986). Intention and intentional action: The simple view. *Mind & Language*, 1(4), 281–301.

Adams, R. M. (1985). Involuntary sins. *The Philosophical Review*, 94(1), 3–31.

Aguilar, J. H. (2012). Basic causal deviance, action repertoires, and reliability. *Philosophical Issues*, 22, 1–19.

Aguilar, J. H., and Buckareff, A. (eds) (2010). *Causing human actions: New perspectives on the causal theory of action*. MIT Press.

Amaya, S. (2018). Two kinds of intentions: a new defense of the Simple View. *Philosophical Studies*, 175(7), 1767–86.

Anscombe, E. (2000). *Intention*. Harvard University Press.

Armstrong, D. M. (1968). *A materialist theory of the mind*. Routledge.

Arpaly, N., and Schroeder, T. (2012). Deliberation and acting for reasons. *Philosophical Review*, 121(2), 209–39.

Austin, J. L. (1956). A plea for excuses: The presidential address. *Proceedings of the Aristotelian Society*, 57, 1–30.

Bayne, T. (2009). Perception and the reach of phenomenal content. *The Philosophical Quarterly*, 59(236), 385–404.

Beck, J. (2014). Analogue magnitude representations: A philosophical introduction. *The British Journal for the Philosophy of Science*, 66(4), 829–55.

Bengson, J. (2017). The unity of understanding. In *Making sense of the world: New essays on the philosophy of understanding*, Grimm, S. R. (ed.), 14–53. Oxford University Press.

Bird, A. (1998). Dispositions and antidotes. *The Philosophical Quarterly*, 48, 227–34.

Bishop, J. (1989). *Natural agency: An essay on the causal theory of action*. Cambridge University Press.

Boggan, A. L., Bartlett, J. C., and Krawczyk, D. C., (2012). Chess masters show a hallmark of face processing with chess. *Journal of Experimental Psychology: General*, 141(1), 37.

Bradford, G. (2015). *Achievement*. Oxford University Press.

Brand, M. (1984). *Intending and acting: Toward a naturalized action theory*. MIT Press.

Brand, M. (1986). Intentional actions and plans. *Midwest Studies in Philosophy*, 10(1), 213–30.

Bratman, M. (1984). Two faces of intention. *The Philosophical Review*, 93(3), 375–405.

Bratman, M. E. (1999). *Faces of intention: Selected essays on intention and agency*. Cambridge University Press.

Bratman, M. E. (2007). *Structures of agency: Essays*. Oxford University Press.

Bratman, M. E. (2018). *Planning, time, and self-governance: Essays in practical rationality*. Oxford University Press.

Briscoe, R. (2008). Vision, action, and make-perceive. *Mind & Language* 23(4), 457–97.

Briscoe, R. (2018). On the uses of make-perceive. In *Perceptual memory and perceptual imagination*, F. Dorsch and F. Macpherson (eds), 161–85. Oxford University Press.

Buehler, D. (2019). Flexible occurrent control. *Philosophical Studies*, 176(8), 2119–37.

Burge, T. (1996). Our entitlement to self-knowledge: I. *Proceedings of the Aristotelian Society*, 96, 91–116.

Burge, T. (2009). Primitive agency and natural norms. *Philosophy and Phenomenological Research*, 79(2), 251–78.

Burge, T. (2013). *Cognition through understanding: Self-knowledge, interlocution, reasoning, reflection: philosophical essays (Vol. 3)*. Oxford University Press.

Burgoyne, A. P., Sala, G., Gobet, F., Macnamara, B. N., Campitelli, G. and Hambrick, D. Z. (2016). The relationship between cognitive ability and chess skill: A comprehensive meta-analysis. *Intelligence*, 59, 72–83.

Byrne, A. (2005). Introspection. *Philosophical Topics*, 33(1), 79–104.

Byrne, A. (2011). Transparency, belief, intention. *Proceedings of the Aristotelian Society*, 85(1), 201–21.

Camp, E. (2007). Thinking with maps. *Philosophical Perspectives*, 21(1), 145–82.

Camp, E. (2009). Putting thoughts to work: Concepts, systematicity, and stimulus-independence. *Philosophy and Phenomenological Research*, 78(2), 275–311.

Campbell, L. (2018). An epistemology for practical knowledge. *Canadian Journal of Philosophy*, 48(2), 159–77.

Candlish, S. and Damnjanovic, N. (2013). Reasons, actions, and the will: The fall and rise of causalism. In *The Oxford handbook of the history of analytic philosophy*, M. Beaney (ed.). Oxford University Press.

Carnap, R. (1950). *Logical foundations of probability.* University of Chicago Press.

Carter, J. A., and Pritchard, D. (2015). Knowledge-how and epistemic luck. *Noûs*, 49(3), 440–53.

Carter, J. A., Pritchard, D., and Shepherd, J. (2019). Knowledge-how, understanding-why, and epistemic luck: An experimental study. *Review of Philosophy and Psychology*, 10, 701–34.

Chase, W. G., and Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55–81.

Choi, S. (2006). The simple vs. reformed conditional analysis of dispositions. *Synthese*, 148, 369–79.

Choi, S. (2008). Dispositional properties and counterfactual conditionals. *Mind*, 117, 795–841.

Choi, S., and Fara, M. (2018). Dispositions. *The Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/fall2018/entries/dispositions/.

Christensen, W., Sutton, J., and McIlwain, D. J. (2016). Cognition in skilled action: Meshed control and the varieties of skill experience. *Mind & Language*, 31(1), 37–66.

Chuard, P., and Southwood, N. (2009). Epistemic norms without voluntary control. *Noûs*, 43(4), 599–632.

Clark, A. (2007). Soft selves and ecological control. In *Distributed cognition and the will: Individual volition and social context*, Ross, D., Spurrett, D., Kincaid, H., and Stephens, G. L. (eds), 101–22. MIT Press.

Clarke, R. (2009). Dispositions, abilities to act, and free will: The new dispositionalism. *Mind*, 118(470), 323–51.

Coen, E., and Coen, J. (2009). *The Big Lebowski*. Faber & Faber.

Coffman, E. J. (2018). Do We Decide Intentionally? *Australasian Journal of Philosophy*, 96(4), 1–6.

Coliva, A. (2016). *The varieties of self-knowledge*. Palgrave-Macmillan.

Costa, M. J. (1986). Acting intentionally and minimal abilities. *Analysis*, 46(3), 144–7.

D'Oro, G. and Sandis, C. (eds) (2013). *Reasons and causes: Causalism and anti-causalism in the philosophy of action*. Palgrave-Macmillan.

Davidson, D. (1980). *Essays on action and events*. Oxford University Press.

Davidson, D. (2004). *Problems of rationality*. Oxford University Press.

Davis, W. A. (1984). A causal theory of intending. *American Philosophical Quarterly*, 21(1), 43–54.

Day, K. A., Roemmich, R. T., Taylor, J. A., and Bastian, A. J. (2016). Visuomotor learning generalizes around the intended movement. *eNeuro*, DOI: 10.1523/eneuro.0005-16.2016.

de Groot, A.D. (1978). *Thought and choice in chess*, 2nd edition. Mouton Publishers.

Dennett, D. C. (1969). *Content and consciousness*. Routledge.

Dennett, D. C. (1984). *Elbow room: The varieties of free will worth wanting*. MIT Press.

Dickie, I. (2012). Skill before knowledge. *Philosophy and Phenomenological Research*, 85(3), 737–45.

Dickie, I. (2015). *Fixing Reference*. Oxford University Press.

Dickinson, A. (2012). Associative learning and animal cognition. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 367(1603), 2733–42.

Di Nucci, E. (2009). Simply, false. *Analysis*, 69(1), 69–78.

Di Nucci, E. (2013). Action, deviance, and guidance. *Abstracta*, 7(2), 41–59.

Dwyer, D. M., Mackintish, N. J., and Boakes, R. A. (1998). Simultaneous activation of the representations of absent cues results in the formation of an excitatory association between them. *Journal of Experimental Psychology: Animal Behaviour Processes*, 24(2), 163–71.

Egan, A. (2008). Seeing and believing: Perception, belief formation, and the divided mind. *Philosophical Studies*, 140(1), 47–63.

Enç, B. (2003). *How we act: Causes, reasons, and intentions*. Oxford University Press.

Ericsson, K. A., and Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology*, 47(1), 273–305.

Falvey, K. (2000). Knowledge in intention. *Philosophical Studies*, 99(1), 21–44.

Fara, M. (2008). Masked abilities and compatibilism. *Mind*, 117(468), 843–65.

Farkas, K. (2017). Practical know-wh. *Noûs*, 51(4), 855–70.

Fodor, J. A. (1984). Semantics, Wisconsin style. *Synthese*, 59(3), 231–50.

Foran, S. T. (1997). *Animal movement*. ProQuest Dissertation Publishing.

Forsberg, C. (2017). Kyrie Irving again rises to challenge in Celtics' comeback victory. *Espn.com*,http://www.espn.com/nba/story/_/id/21696375/nba-kyrie-irving-again-rises-challenge-boston-celtics-victory.

Frankfurt, H. (1988). *The importance of what we care about*. Cambridge University Press.

Fridland, E. (2014). They've lost control: Reflections on skill. *Synthese*, 191(12), 2729–50.

Fritz, K. G. (2018). Moral responsibility, voluntary control, and intentional action. *Philosophia*, 46(4), 831–55.

Frost, K. (2019). A metaphysics for practical knowledge. *Canadian Journal of Philosophy*, 49(3), 314–40.

Gauker, C. (2011). *Words and images: An essay on the origin of ideas*. Oxford University Press.

Gert, J. (2004). Intentional action and nearly certain success. *Ratio*, 17(2), 150–8.

Gertler, B. (2015). Self-knowledge. *The Stanford encyclopedia of philosophy*, E. N. Zalta (ed.), https://plato.stanford.edu/archives/sum2015/entries/self-knowledge/

Gilovich, T., Griffin, D., and Kahneman, D. (eds) (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge University Press.

Ginet, C. (1990). *On action*. Cambridge University Press.

Goldman, A. I. (1970). *Theory of human action*. Princeton University Press.

Greco, D. (2016). Safety, explanation, iteration. *Philosophical Issues*, 26(1), 187–208.

Green, E. J., and Quilty-Dunn, J. (2017). What is an object file? *The British Journal for the Philosophy of Science*, DOI: 10.1093/bjps/axx055.

Grice, P. (1974). Method in philosophical psychology (from the banal to the bizarre). *Proceedings and addresses of the American Philosophical Association*, 48, 23–53.

Grünbaum, T. (2013). Seeing what one is doing. *Philosophy and Phenomenological Research*, 86(2), 295–318.

Gould J. L., and Gould C. G. (1988). *The honey bee*. W. H. Freeman.

Habgood-Coote, J. (2019). Knowledge-how, abilities, and questions. *Australasian Journal of Philosophy*, 97(1), 86–104.

Harle, S. K., and Vickers, J. N., (2001). Training quiet eye improves accuracy in the basketball free throw. *The Sport Psychologist*, 15(3), 289–305.

Harman, G. (1976). Practical reasoning. *The Review of Metaphysics*, 29(3), 431–63.

Helton, G. (forthcoming). If you can't change what you believe, you don't believe it. *Noûs*.

Horgan, T., Tienson, J., and Graham, G. (2003). The phenomenology of first-person agency. In *Physicalism and mental causation: The metaphysics of mind and action*, S. Walter, and H.-D. Heckmann (eds), 323–41. Imprint Academic.

Hornsby, J. (2013). Basic activity. *Aristotelian Society Supplementary Volume*, 87(1), 1–18.

Hurley, S. (2003). Animal action in the space of reasons. *Mind & Language*, 18(3), 231–57.

Hyman, J. (2014). Desires, dispositions and deviant causal chains. *Philosophy*, 89(1), 83–112.

Hyman, J. (2015). *Action, knowledge, and will*. Oxford University Press.

Hyman, S. E. (2007). The neurobiology of addiction: implications for voluntary control of behaviour. *The American Journal of Bioethics*, 7(1), 8–11.

Jeannerod, M. (2006). *Motor cognition: What actions tell to the self.* Oxford University Press.

Jenkinson, P. M., Edelstyn, N. M., Preston, C., and Ellis, S. J. (2015). Anarchic hand with abnormal agency following right inferior parietal lobe damage: a case report. *Neurocase*, 21(4), 471–8.

Justus, J. (2012). Carnap on concept determination: Methodology for philosophy of science. *European Journal for Philosophy of Science*, 2(2), 161–79.

Keijzer, F. (2013). The Sphex story: How the cognitive sciences kept repeating an old and questionable anecdote. *Philosophical Psychology*, 26(4), 502–19.

Kittle, S. (2017). Robustness and up-to-us-ness. *Disputatio*, 9(44), 35–57.

Krakauer, J., Hadjiosif, A. M., Xu, J., Wong, A. L., and Haith, A. (2019). Motor learning. *Comprehensive Physiology*, 9(2), 613–63.

Kriegel, M. (2007). *Pistol: The life of Pete Maravich.* Free Press.

Kung, C., and Saimi, Y. (1982). The physiological basis of taxes in Paramecium. *Annual Review of Physiology*, 44(1), 519–34.

Langton, R. (2006). Kant's phenomena: Extrinsic or relational properties? A reply to Allais. *Philosophy and Phenomenological Research*, 73(1), 170–85.

Lavin, D. (2013). Must there be basic action? *Noûs*, 47(2), 273–301.

Levy, Y. (2013). Intentional action first. *Australasian Journal of Philosophy*, 91(4), 705–18.

Levy, Y. (2018). What is "mental action"? *Philosophical Psychology*, 32(6), 971–93.

List, C., and Pettit, P. (2011). *Group agency: The possibility, design, and status of corporate agents*. Oxford University Press.

Ludwig, K. (1992). Impossible doings. *Philosophical Studies*, 65(3), 257–81.

McCann, H. J. (1987). Rationality and the range of intention. *Midwest Studies in Philosophy*, 10(1), 191–211.

McCann, H. J. (2011). The Simple View again: a brief rejoinder. *Analysis*, 71(2), 293–5.

McCarthy, C. (2001). *Blood meridian, or, The evening redness in the west*. Modern Library.

McHugh, C. (2014). Exercising doxastic freedom. *Philosophy and Phenomenological Research*, 88(1), 1–37.

Mackie, J. L., (1977). Dispositions, grounds, and causes, *Synthese*, 34, 361–70.

Maier, J. (2018a). Abilities. *The Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/archives/spr2018/entries/abilities/.

Maier, J. (2018b). Ability, modality, and genericity. *Philosophical Studies*, 175(2), 411–28.

Manley, D., and Wasserman, R. (2008). On linking dispositions and conditionals. *Mind*, 117(465), 59–84.

Marcel, A. (2003). The sense of agency: Awareness and ownership of action. In *Agency and self-awareness: Issues in philosophy and psychology,* J. Roessler and N. Eilan (eds), 48–93. Oxford University Press.

Marchetti, C., and Della Sala, S. (1998). Disentangling the alien and anarchic hand. *Cognitive Neuropsychiatry*, 3(3), 191–207.

Martin, C. B. (1994). Dispositions and conditionals. *The Philosophical Quarterly*, 44, 1–8.

Marušić, B., and Schwenkler, J. (2018). Intending is believing: A defense of strong cognitivism. *Analytic Philosophy*, 59(3), 309–40.

Mayr, E. (2011). *Understanding human agency*. Oxford University Press.

Mele, A. R. (1987). Intentional action and wayward causal chains: The problem of tertiary waywardness. *Philosophical Studies*, 51(1), 55–60.

Mele, A. R. (1989). Intention, belief, and intentional action. *American Philosophical Quarterly*, 26(1), 19–30.

Mele, A. R. (1992). *Springs of action: Understanding intentional behaviour*. Oxford University Press.

Mele, A. R. (2001). Acting intentionally: Probing folk notions. In *Intentions and intentionality: Foundations of social cognition,* Malle, B., Moses, L. J., and Baldwin, D. (eds), 27–43. MIT Press.

Mele, A. R. (2003a). Agents' abilities. *Noûs*, 37(3), 447–70.

Mele, A. R. (2003b). *Motivation and agency*. Oxford University Press.

Mele, A. R. (2017a). Direct control. *Philosophical Studies*, 174(2), 275–90.

Mele, A. R. (2017b). *Aspects of agency: decisions, abilities, explanations, and free will*. Oxford University Press.

Mele, A. R. (2019). Causalism: On action explanation and causal deviance, in *Explanation in action theory and historiography: Causal and teleological approaches*, Schumann, G.(ed.), 45–58. Routledge.

Mele, A. R., and Moser, P. K. (1994). Intentional action. *Noûs*, 28(1), 39–68.

Melis, A. P., Hare, B., and Tomasello, M. (2006). Chimpanzees recruit the best collaborators. *Science*, 311(5765), 1297–300.

Milford, N. (2001). *Savage beauty: The life of Edna St. Vincent Millay*. Random House.

Miracchi, L. (2015). Competence to know. *Philosophical Studies*, 172(1), 29–56.

Miracchi, L. (2017). Perception first. *The Journal of Philosophy*, 114(12), 629–77.

Moran, R. (2004). Anscombe on "practical knowledge," in *Royal Institute of Philosophy Supplement*, J. Hyman and H. Steward (eds), 43–68. Cambridge University Press.

Morgan, W. J. (2004). Moral antirealism, internalism, and sport. *Journal of the Philosophy of Sport*, 31(2), 161–83.

Morton, J. M. (2017). Reasoning under scarcity. *Australasian Journal of Philosophy*, 95(3), 543–59.

Mylopoulos, M., and Pacherie, E. (2017). Intentions and motor representations: The interface challenge. *Review of Philosophy and Psychology*, 8(2), 317–36.

Nadelhoffer, T. (2006). On trying to save the simple view. *Mind & Language*, 21(5), 565–86.

O'Brien, Lillian. (2012). Deviance and causalism. *Pacific Philosophical Quarterly*, 93(2), 175–96.

O'Brien, Lillian. (2015). *Philosophy of action*. Palgrave-Macmillan.

O'Brien, Lucy. (2009). *Self-knowing agents*. Oxford University Press.

O'Brien, Lucy. (2017). Actions as prime. *Royal Institute of Philosophy Supplements*, 80, 265–85.

O'Callaghan, C. (2015). The multisensory character of perception. *The Journal of Philosophy*, 112(10), 551–69.
O'Shaughnessy, B. (1973). Trying (as the mental "pineal gland"). *The Journal of Philosophy*, 70(13), 365–86.
Paul, S. K. (2009a). How we know what we're doing. *Philosopher's Imprint*, 9(11): 1–24.
Paul, S. K. (2009b). Intention, belief, and wishful thinking: Setiya on "practical knowledge." *Ethics*, 119(3), 546–57.
Pavese, C. (2015a). Practical senses. *Philosophers' Imprint*, 15(29), 1–25.
Pavese, C. (2015b). Knowing a rule. *Philosophical Issues*, 25(1), 165–88.
Pavese, C. (2016a). Skill in epistemology I: Skill and knowledge. *Philosophy Compass*, 11(11), 642–9.
Pavese, C. (2016b). Skill in epistemology II: Skill and know how. *Philosophy Compass*, 11(11), 650–60.
Pavese, C. (2017). Know-how and gradability. *Philosophical Review*, 126(3), 345–83.
Pavese, C. (2018). Know-how, action, and luck. *Synthese*, DOI: 10.1007/s11229-018-1823-7, 1–23.
Peacocke, C. (1979). Deviant causal chains. *Midwest Studies in Philosophy*, 4(1), 123–55.
Peacocke, C. (1996). Our entitlement to self-knowledge: II. *Proceedings of the Aristotelian Society*, 96, 117–58.
Peacocke, C. (1999). *Being known*. Oxford University Press.
Peacocke, C. (2003). *The realm of reason*. Oxford University Press.
Pettit, P. (2007). Rationality, reasoning, and group agency. *Dialectica*, 61(4), 495–519.
Pezzulo, G., Barca, L., Bocconi, A. L., and Borghi, A. M. (2010). When affordances climb into your mind: advantages of motor simulation in a memory task performed by novice and expert rock climbers. *Brain and Cognition*, 73(1), 68–73.
Pickard, H. (2004). Knowledge of action without observation. *Proceedings of the Aristotelian Society*, 104(1), 205–30.
Poston, T. (2009). Know how to be Gettiered? *Philosophy and Phenomenological Research*, 79(3), 743–7.
Potok, C. (1972). *My name is Asher Lev*. Knopf.
Pryor, J. (2000). The skeptic and the dogmatist. *Noûs*, 34(4), 517–49.
Quilty-Dunn, J., and Mandelbaum, E. (2018). Inferential transitions. *Australasian Journal of Philosophy*, 96(3), 532–47.
Robinson, M. (2014). The limits of limited-blockage Frankfurt-style cases. *Philosophical Studies*, 169, 429–46.
Rödl, S. (2011). Two forms of practical knowledge. In *Essays on Anscombe's Intention*, Ford, A., Hornsby J., and Stoutland F. (eds), 211–41. Harvard University Press.
Ross, D. (1978). He loads the gun, not the dice. *Analysis*, 38(3), 114–15.
Ross, L. N. (2018). Causal concepts in biology: How pathways differ from mechanisms and why it matters. *British Journal for Philosophy of Science*, DOI: 10.1093/bjps/axy078.
Russell, J. S. (2004). Moral realism in sport. *Journal of the Philosophy of Sport*, 31(1), 142–60.

Salinger, J. D. (1961). *Franny and Zooey*. Little, Brown.
Sandis, C. (2006). The explanation of action in history. *Essays in Philosophy*, 7(2), 12.
Schlosser, M. E. (2007). Basic deviance reconsidered. *Analysis*, 67(3), 186–94.
Schmidt, R. A. and Young, D. E., (1987). Transfer of movement control in motor skill learning. In *Transfer of learning: Contemporary research and applications*, Cormier, S. M., and Hagman, J. D. (eds). Academic Press.
Schwenkler, J. (2011). Perception and practical knowledge. *Philosophical Explorations*, 14(2), 137–52.
Schwenkler, J. (2012). Non-observational knowledge of action. *Philosophy Compass*, 7(10), 731–40.
Schwenkler, J. (2015). Understanding "practical knowledge." *Philosophers' Imprint*, 15(15), 1–32.
Sehon, S. (2016). *Free will and action explanation: A non-causal, compatibilist account*. Oxford University Press.
Setiya, K. (2008). Practical knowledge. *Ethics*, 118(3), 388–409.
Setiya, K. (2009). Practical knowledge revisited. *Ethics*, 120(1), 128–37.
Shabo, S. (2014). It wasn't up to Jones: Unavoidable actions and intensional contexts in Frankfurt examples. *Philosophical Studies*, 169(3), 379–99.
Shakespeare, W. (1996). *Hamlet*. T. J. B. Spencer, (ed.). Penguin Classics.
Shepherd, J. (2014a). The contours of control. *Philosophical Studies*, 170(3), 395–411.
Shepherd, J. (2014b). Causalism and intentional omission. *American Philosophical Quarterly*, 51(1), 15–26.
Shepherd, J. (2015a). Conscious control over action. *Mind & Language*, 30(3), 320–44.
Shepherd, J. (2015b). Deciding as intentional action: Control over decisions. *Australasian Journal of Philosophy*, 93(2), 335–51.
Shepherd, J. (2016). Conscious action/zombie action. *Noûs*, 50(2), 419–44.
Shepherd, J. (2017a). Halfhearted action and control. *Ergo*, 4(9), 259–76.
Shepherd, J. (2017b). The experience of acting and the structure of consciousness. *The Journal of Philosophy*, 64(6), 422–48.
Shepherd, J. (2018a). Intelligent action and the use of mixed representational formats. *Synthese*, DOI: 10.1007/s11229-018-1892-7.
Shepherd (2018b). Intending, believing, and supposing at will. *Ratio*, 31(3), 321–30.
Shepherd, J. (2019). Skilled action and the double life of intention. *Philosophy and Phenomenological Research*, 98(2), 286–305.
Shepherd, J. (forthcoming-a). Disappearing agents, mental action, rational glue. In *Mental action and the conscious mind*, Brent, M.(ed.). Routledge.
Shepherd, J. (forthcoming-b). The targets of skill and their importance. In *Routledge handbook of skill and expertise*, Fridland, E.and Pavese, C.(eds). Routledge.
Shepherd, J. (n.d.). Skill and sensitivity to reasons.
Shepherd, J., and Justus, J. (2015). X-phi and Carnapian explication. *Erkenntnis*, 80(2), 381–402.
Simon, R. L. (2000). Internalism and internal values in sport. *Journal of the Philosophy of Sport*, 27, 1–16.

Sinhababu, N. (2017). *Humean nature: How desire explains action, thought, and feeling*. Oxford University Press.

Smith, A. M. (2005). Responsibility for attitudes: Activity and passivity in mental life. *Ethics*, 115(2), 236–71.

Sosa, E. (2015). *Judgment and agency*. Oxford University Press.

Stanley, J. (2011). *Know how*. Oxford University Press.

Stanley, J., and Williamson, T. (2017). Skill. *Noûs*, 51(4), 713–26.

Steward, H. (2012). *A metaphysics for freedom*. Oxford University Press.

Stich, S. P. (1978). Beliefs and subdoxastic states. *Philosophy of Science*, 45(4), 499–518.

Tännsjö, T. (2009). On deviant causal chains—no need for a general criterion. *Analysis*, 69(3), 469–73.

Tautz, J., Zhang, S., Spaethe, J., Brockmann, A., Si, A., and Srinivasan, M. (2004). Honeybee odometry: performance in varying natural terrain. *PLoS Biology*, 2(7), e211.

Thalberg, I. (1962). Intending the impossible. *Australasian Journal of Philosophy*, 40(1), 49–56.

Thalberg, I. (1984). Do our intentions cause our intentional actions? *American Philosophical Quarterly*, 21(3), 249–60.

Thompson, M. (2008). *Life and action*. Harvard University Press.

Uggetti, P. (2018). The Rockets feel like the NBA's best team. *The Ringer*, https://www.theringer.com/nba/2018/3/1/17067474/houston-rockets-los-angeles-clippers-best-team.

Unterrainer, J. M., Kaller, C. P., Halsband, U., and Rahm, B. (2006). Planning abilities and chess: A comparison of chess and non-chess players on the Tower of London task. *British Journal of Psychology*, 97(3), 299–311.

Van der Maas, H. L., and Wagenmakers, E. J. (2005). A psychometric analysis of chess expertise. *The American Journal of Psychology*, 118(1), 29–60.

Velleman, D. (2007). *Practical reflection*. CSLI Publications.

Vetter, B. (2014). Dispositions without conditionals. *Mind*, 123(489), 129–56.

Vihvelin, K. (2004). Free will demystified: A dispositional account. *Philosophical Topics*, 32(1/2), 427–50.

Vine, S. J., Moore, L., and Wilson, M. R. (2011). Quiet eye training facilitates competitive putting performance in elite golfers. *Frontiers in Psychology*, 2(8), 8.

Wallace, R. J. (1999). Three conceptions of rational agency. *Ethical Theory and Moral Practice*, 2(3), 217–42.

Wasserman, R. (2011). Intentional action and the unintentional fallacy. *Pacific Philosophical Quarterly*, 92(4), 524–34.

Weatherson, B. (2017). Intellectual skill and the Rylean regress. *The Philosophical Quarterly*, 67(267), 370–86.

Wiggins, D. (2012). Practical knowledge: Knowing how to and knowing that. *Mind*, 121(481), 97–130.

Williamson, T. (2002). *Knowledge and its limits*. Oxford University Press.

Williamson, T. (2017). Acting on knowledge. In *Knowledge first: Approaches in epistemology and mind*, Carter, J. A., Gordon, E. C., and Jarvis, B. (eds), 163–83. Oxford University Press.

Wilson, G. M. (1989). *The intentionality of human action*. Stanford University Press.
Winning, R. J. (2019). *The mechanistic and normative structure of agency*, Doctoral dissertation, University of California-San Diego.
Woolridge, D. E. (1963). *The machinery of the brain*. McGraw-Hill.
Wu, W. (2011). Confronting Many-Many problems: Attention and agentive control. *Noûs*, 45(1), 50–76.
Wu, W. (2016). Experts and deviants: The story of agentive control. *Philosophy and Phenomenological Research*, 93(1), 101–26.

Index

For the benefit of digital users, indexed terms that span two pages (e.g., 52–53) may, on occasion, appear on only one of those pages.