Multitasking, Consequentialism and Practical Imagination

by

Gregory Brendan Cedar Strom

B.A., University of Chicago, 2002

Submitted to the Graduate Faculty of

the School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

School of Arts and Sciences

This dissertation was presented

by

Gregory Strom

It was defended on

June 13, 2011

and approved by

Edouard Machery, Associate Professor, History and Philosophy of Science

John McDowell, Distinguished University Professor, Philosophy

Kieran Setiya, Associate Professor, Philosophy

Michael Thompson, Professor, Philosophy

Dissertation Advisor: Robert Brandom, Distinguished Professor, Philosophy

Multitasking, Consequentialism and Practical Imagination

Gregory Strom, PhD

University of Pittsburgh, 2011

I maintain that consequentialism is committed to the view that there must be a fundamental bifurcation between sorts of value possessed by inner and outer elements of any exercise of agency, and I argue that this view is false because its adherents are unable to avoid global skepticism about the possibility of agentive efficacy. This result obliges us to articulate a concept of a thing that an agent can do in which her efforts are robustly united, rather than merely conjoined, with the achievement of her goals. I find such a concept in the notion of a task as it figures in the concept of multitasking. A task is a kind of unity that subsumes everything an agent is doing in order to achieve a certain end. I argue that empiricism about practical options—the view according to which a practical option is just given to our practical cognition rather than being somehow the work of our practical cognitive faculties—cannot account for the role that this sort of unity plays in our ordinary concept of multitasking. I also present an argument against a "command picture" of ethics according to which the ethical value of an exercise of agency can be determined by considering exclusively what its agent has wrought thereby, rather than necessarily by also considering the reasons from which she wrought it. Finally, I offer an alternative to this picture in the form of a series of adaptations and adjustments of Kantian moral doctrine: first, that respect for the moral law is a determinable having many determinations—honesty, benevolence, etc.—and that only actions performed from one of the various motives that are respectful of the moral law possess moral worth; and second,

that the moral law is also a determinable having many determinations and that an action can possess moral worth if the motive from which it is prompted is an attunement to the determination of this determinable that she bears. My most novel claim is that an agent's attunement to her determination of practical rationality is secured by the activity of a faculty of practical imagination.

TABLE OF CONTENTS

PREFACE

My mostly ruminative way of thinking would have gone nowhere in this dissertation without a steady diet of ideas and suggestions from texts and conversations so various and sometimes so far buried in my philosophical past that I have in many cases simply forgotten them or lost my ability to recognize their influence. I can say for certain that I was first dimly awakened to the possibility of philosophizing about the ethical and action-theoretic questions I've tried to address in this dissertation by my father James Strom, whose continuing willingness to discuss them with me has been of great value ever since. Their philosophical importance was made more explicit to me when as an undergraduate I took a course on action theory taught by Candace Vogler at the University of Chicago. But it was not until several years into my time at the University of Pittsburgh that I really came to see that what makes a philosophical problem of these and many other questions is the work that our treatment of them has to play in our conception of rationality. This finally dawned on me when institutional pressures compelled me to devote a summer's philosophical attention to Sellars—specifically, to his aphorism that "in characterizing an episode or a state as that of knowing, we are not giving an empirical description of that episode or state; we are placing it in the logical space of reasons, of justifying and being able to justify

what one says"[1]—and to the conception of philosophy implicit in Socrates' characterization of oratory as a form of flattery in Plato's *Gorgias*.

It has been my great fortune to work on this dissertation as a member of the philosophical community at the University of Pittsburgh. I should especially thank the five members of my dissertation committee: my adviser Robert Brandom, to whom I am boundlessly grateful for sharing with me what I think is a special genius for philosophical engagement, encouragement and advice; Edouard Machery, who kindly consented to be my outside reader and to press me on several important points; John McDowell, whose characteristically careful and insightful attention to my work has enriched it, or at least made it (I hope) far less poor than it would otherwise have been; Kieran Setiya, whose questions and responses to my work posed challenges with which I will continue to wrestle for a long time to come; and Michael Thompson, whose uncanny talent for infusing even the slightest remark with more philosophical insight than is contained in most philosophical articles, and whose lectures and writings on the topics I've tried to address in this dissertation have pervasively shaped my thinking.

I am grateful also to many of my fellow graduate students at the University of Pittsburgh for many helpful discussions of my work and my ideas: Daniel Addison, Jason Carnell, Jack Fortune, Joshua Hancox, Brandon Hogan, Kathryn Lindeman, Evgenia Mylonaki, Alexandra Newton, Hille Paakkunainen, Eric Palmer, Jonathan Surovell, and Christine Young.

The friendship of my very good friend Paul Ellis, who is not a professional philosopher, also warrants an expression of gratitude here, since it has sustained and sparked my philosophical spirit through many bleak hours.

---

[1] Sellars, EPM, §36.

More than to anyone else, I am grateful to my wife Rachael Driver for her encouragement and her love, and for many philosophical conversations in which she has helped me find my way. Without her, I would not have had the courage, the occasion or the ability to write what I think are the very best parts of this dissertation. Thank you, Rachael.

1.0   INTRODUCTION

1.1   FIRST SECTION

Much of this dissertation was inspired by the idea that consequentialism could be shown to be incompatible with any adequate metaphysics of rational agency. One of the most frustrating obstacles on the path along which that inspiration drew me is the fact that "consequentialism" does not these days always mean what I think it should. What it should mean is what Anscombe meant by it when she coined the term, namely, the view that "'the right action' is the action which produces the best possible consequences (reckoning among consequences the intrinsic values ascribed to certain kinds of act by some 'Objectivists')"[2]—perhaps softened slightly, in order to honor the possibility that there may not be "best possible consequences" (since there may be infinitely many possible consequences or sets thereof), into the claim that an action is

---

[2] Anscombe, "Modern Moral Philosophy," p. 357. Anscombe does not explicitly mark this passage as a definition of consequentialism, and her official definition of the term is somewhat different, but it is clear in this passage nonetheless that she takes herself to be giving essential characterizations of the view she has in mind. This is also clear from a helpful footnote that she attaches to the parenthetical remark just quoted:

> Oxford Objectivists of course distinguish between "consequences" and "intrinsic values" and so produce a misleading appearance of not being "consequentialists." But they do not hold—and Ross explicitly denies—that the gravity of, e.g., procuring the condemnation of the innocent is such that it cannot be outweighed by, e.g., national interest. Hence their distinction is of no importance.

1

good insofar as its consequences are good.[3] That is what "consequentialism" should mean and that is what I thought it meant when I was inspired to write this dissertation.

So that—pending one further clarification that I will get to at the end of this section—is how I will use the term "consequentialism." And I realize that I owe the reader an account of why the current practice of characterizing consequentialism quite differently—in opposition to deontology as the claim that morality is agent-neutral—does not adequately capture its spirit. That debt will be paid in my first chapter. What I want to emphasize at this point is just that there at least was (and in my opinion still is) a certain consequentialist spirit, a spirit whose prevalence among all "the well-known English writers on moral philosophy from Sidgwick to the present day"[4] offended the young Anscombe and motivated her to coin a word for it and to characterize it in a way that I have refined into the "insofar as" claim with which I am working here. The fact that this label has been (as I will argue later) simply *reassigned* to a claim about agent-neutrality merely obscures without yet rendering obsolete the spirit of the view that Anscombe meant to indicate. For this reassignment had the effect of changing the subject, rather than finishing or refining the conversation Anscombe tried to start.

It should be clear that the content of consequentialism as I construe it turns on the significance of the distinction between an action and its consequences, so that one of the primary philosophical obligations of a consequentialist should be to clarify this distinction. And how she discharges this obligation constrains the strategies she may deploy to account for the supposed truth of the "insofar as" claim definitive of consequentialism. So we may distinguish these two

---

[3] For the sake of simplicity, I here disregard the importantly different valences of "right" and "good." I think this is dialectically fair because consequentialists are for the most part concerned, if at all, to define "right" only as a kind of afterthought, after establishing (that there is) an ethically illuminating method for ranking actions in proportion to the goodness of their consequences.

[4] Anscombe, "Modern Moral Philosophy," p. 351.

related questions: (1) what is the distinction between an action and its consequences? and (2) why is it true that an action is good insofar as its consequences are good?

Now, if one's answer to the first question is that an action just *is* the set of all its consequences—a fairly common stipulation—then consequentialism will of course turn out to be true, since it is a tautology that an action is good insofar as it is good. Perhaps this reduction of consequentialism to a tautology can be resisted by maintaining that although an action is identical to the set of its consequences, nonetheless the *sort of value* it has *as an action* is distinct from the sort of value it has *as a set of consequences*, and that it has the first sort of value insofar as it has the second. Since this view retains contentfulness for consequentialism by distinguishing two different aspects (as opposed to parts) of a single exercise of agency, I call this view "*aspectual* consequentialism." Aspectual consequentialism seems to be the only available way to maintain a non-tautologous consequentialism while identifying an action with its consequences. Notice, however, that by transposing the distinction between an action and its consequences into a distinction between two aspects of a single exercise of agency, an aspectual consequentialist does not address our request for a clarification of the distinction fundamental to her consequentialism. She has merely told us that this distinction is *not* a distinction between different parts of an exercise of agency, some of which precede others. We still need to understand what that distinction is. And the difficulty that aspectual consequentialism has not escaped by making this move is still pressing for non-aspectual forms of consequentialism as well: in order to be contentful, any form of consequentialism must somehow maintain—perhaps only at the level of aspect—that the distinction between an action and its consequences is suitably robust. The contentfulness of consequentialism requires there to be some sort of fundamental bifurcation in any exercise of agency between the action involved and its

consequences. I call the claim that there is such a bifurcation "bifurcationism," and my ambition in the first chapter of this dissertation is to show that bifurcationism is false.

I will return to that point, but first I would like to discuss the second question that I said any consequentialist must answer, namely, why is it true that an action is good insofar as its consequences are good? Here there are many possible answers. One might (for example) claim that an interventionist God has so ordered the universe that every good action is rewarded, at least in the long run, with a set of consequences somehow proportionally as good as that action, and every bad action is punished with a set of consequences somehow proportionally as bad as that action.[5] But of course this answer, which is true to the letter of what I've been calling consequentialism so far, is not true to its spirit. And that is because, setting aside points of theology, the notions of reward and punishment that figure in this answer require that the value an action possesses be logically *prior*, along some crucial dimension, to the value possessed by its consequences, whereas the spirit of consequentialism requires that the value possessed by an action be logically *posterior*, along some crucial dimension, to the value possessed by its consequences. This is so because consequentialism presents itself as a theory of what makes an action good (namely, its consequences) and whatever makes something have a certain property is prior to it along some crucial dimension.

For this reason, the formulation of consequentialism with which I have finally chosen to work in this dissertation is of it as the claim that the ethical value of an action is logically posterior to (because it is explained by, inherited from, or something like that) some sort of value possessed by its consequences. Traditional deontology—or as I prefer to call it, intentionalism— is the opposing doctrine according to which the ethical value of an action's consequences is

---

[5] This karmic view is perhaps a position that Voltaire meant to mock in *Candide*.

logically posterior to (because it is explained by, inherited from, or something like that) some sort of value possessed by the action itself. This opposition of views depends upon the robustness of the distinction between actions and consequences. And so I characterize the upshot of my first chapter as the claim that the distinction between actions and consequences cannot be robust enough to underwrite this debate. But if my argument is sound, then even a bifurcationist who does not assert a priority relation here, and who resolutely sticks to a mere "insofar as" claim relating the goodness of an action to the goodness of its consequences, will find her claim reduced either to a tautology (if she decides to assimilate what she calls "consequences" to the exercise of agency) or to a claim of pre-ordered harmony (if she decides to assimilate what she calls "consequences" to whatever is left entirely outside of the exercise of agency).

1.2   SECOND SECTION

My argument against bifurcationism takes the form of an argument for a disjunctivist response to a practical analogue of the Argument from Illusion that I call the Argument from Failure. Just as the Argument from Illusion tries to force us to conclude that we cannot *know* anything on the basis of appearances except perhaps that things *appear* to us to be thus and so, the Argument from Failure tries to force us to conclude that we cannot *do* anything by trying to do A except perhaps to *try* to do A. These arguments work as follows (an element of the Argument from Illusion precedes each dash; an element of the Argument from Failure follows each dash):

Just as there is always some slack between its appearing to us that things are thus and so, all on its own, and their actually being thus and so—so too there is always some slack between

5

our trying to do A, all on its own, and our successfully doing A. And just as we cannot know that things are thus and so on any basis such that there is slack between it and the fact that things are thus and so—so too we cannot do A by a certain means if there is slack between that means and our getting A done. So we cannot know on the basis of its appearing to us that things are thus and so, all on its own, that things are thus and so—and in like manner, it follows that we cannot do A by means of trying to do A, all on its own. The best we can accomplish on the basis of its appearing to us as if things are thus and so is to know that things *appear* to us to be thus and so—and the best we can accomplish on the basis of trying to do A is to *try* to do A.

I think that the way to block these arguments is to deny their first premises. Denying these premises amounts to a sort of disjunctivism, I think, because it amounts to affirming that when we are trying to do A (or when we think that things are thus and so on the basis of its appearing to us that things are thus and so), then *either* we are failing to do A (it merely appears that things are thus and so), *or else* we are indeed doing A (we do indeed know that things are thus and so) thereby. But this disjunctivist strategy is not how a bifurcationist will try to block the Arguments from Failure and Illusion. For a bifurcationist believes that there is a robust distinction between something she perhaps calls an "action" and something else she perhaps calls its "consequences"—an inner and an outer element of any exercise of agency. So she cannot deny, on pain of undermining this distinction, that there is some slack between this inner element and this outer element. And so her only option for avoiding skepticism is to deny the second premise of each argument—in the case of the Argument from Failure, that we cannot do A by a certain means if there is slack between that means and our getting A done.

The main problem with denying the second premise of the Argument from Failure, I maintain, is that doing so deprives us of the resources for maintaining that one of the inner

elements in the bifurcationist conception of an exercise of agency—whether we call it an "intending" or a "trying" or an "doing"—is an intending or a trying or a doing of one doable rather than another. I will argue for that in a moment, but first I want to emphasize why this is embarrassing for the bifurcationist if it is true. It is embarrassing because, if it is impossible to maintain that one of these inner elements is an intending or a trying or a doing of one doable rather than another, then it is impossible to maintain on any grounds that it ultimately *fails* as what it is an intending or a trying or a doing of. So the phenomenon of failure, on the reality of which the Argument from Failure depends, itself depends for its possibility on the contentfulness of whatever inner elements a bifurcationist chooses to believe in. If denying the second premise deprives the bifurcationist of the resources for maintaining that an intending or a trying or a doing is an intending or a trying or a doing of one doable rather than another, then the possibility of failure evaporates.

Why does denying the second premise of the Argument from Failure deprive the bifurcationist of those resources? In arguing for this point I take as a premise a modest version of Anscombe's bold claim that an action is intentional under a certain description only if its agent knows "without observation" that she is performing an action that description truly describes.[6] I happen to believe that this bold claim is true, but for my purposes all that I need is the more modest claim that an agent is *trying* to do A only if she knows she is trying to do A.[7] And a bifurcationist cannot affirm that. For the second premise of the Argument from Failure is that we cannot do A by a certain means if there is slack between that means and our getting A

---

[6] *Intention*, §8.
[7] A bit of circumstantial evidence that my claim really is more modest than Anscombe's is that Davidson clinches with an affirmation of it his presentation of the most famous purported counterexample against Anscombe's claim: "A man may be making ten carbon copies as he writes, and this may be intentional; yet he may not know that he is; *all he knows is that he is trying*." ("Agency," p. 50; my emphasis).

done. Denying this premise commits the bifurcationist to the claim that we *can* do A by a certain means if there is slack between that means and our getting A done. But if this is so, then a successful getting done of A always requires, in addition to an inner element of intending, trying or doing, a bit of luck that ensures an appropriate connection between this inner element and what it is an intending, trying or doing of. If the inner element is only ever connected by a bit of luck to the outer element that it is an intending, trying or doing of, then the agent can never know of the inner element that it will be appropriately connected to whatever outer elements she hopes to realize. And now the point is this: if that is impossible, then the agent can never know of the various ways of trying available to her that they are ways of trying to do one thing rather than another. So it is false, on the bifurcationist view, that an agent is trying to do A only if she knows she is trying to do A. That is not something she could know.

It may be thought that this argument goes too quickly, for it may be thought that the necessity of an element of luck to connect an intending, trying or doing to an appropriate outer element does not undermine the possibility that the agent may know of this inner element that it will be appropriately connected to whatever outer elements she hopes to realize. But this thought is plausible only insofar as we have a satisfactory response to the Argument from *Illusion*, since the sort of knowledge of an action's results that this thought hopes for can only be the defeasible sort that the Argument from Illusion purports to defeat. So how is the Argument from Illusion to be defeated? Either we deny its first or its second premise. If the bifurcationist is willing to deny its first premise, then her unwillingness to deny the first premise of the Argument from Failure would seem to lack motivation.[8] So she must deny its second premise, which asserts that

---

[8] This sentence goes proxy for a more detailed argument that I provide in my first chapter for the claim that if the Argument from Illusion can be overcome in the relevant contexts, then the first premise of the Argument from Failure must be false.

we cannot know that things are thus and so on any basis such that there is slack between it and the fact that things are thus and so. If there *is* always slack between tryings and doings, then only by denying this premise from the Argument from Illusion can we claim to know on the basis of our trying in a certain way that we are thereby trying to do A rather than B. But all of the points I have raised against denying the second premise of the Argument from Failure apply with equal force to this second premise of the Argument from Illusion. For in order for there to be the possibility of *illusion* from which this Argument proceeds, it must be possible for us to have an illusory appearance with some (misleading) content, say *p*. But if the second premise is false and the connection between the fact that *p* and the appearance on the basis of which we know that *p* always involves an element of luck, then it is mysterious how an appearance could have this content *p* at all. For we must have ways of "really" knowledgably comparing this sort of appearance with how things really are in order "really" knowledgeably to assess the reliability of this sort of appearance in representing various kinds of facts. And that requires us, at least sometimes, to have "real" knowledge-constituting access to the facts represented by appearances, access that is not subject to the whims of the element of luck. So I have not covertly stipulated a solution to the Argument from Illusion in my treatment of the Argument from Failure—it would be more accurate to say that I am offering a solution that works equally well for both Arguments.[9]

---

[9] I should emphasize that this approach to these Arguments is deeply indebted to what I take to be McDowell's treatment of the Argument from Illusion in "Knowledge and the Internal."

1.3   THIRD SECTION

That is the argument of my first chapter. If that argument is correct, then bifurcationism is false, and any view committed to it—in particular, consequentialism and intentionalism—must be false as well. But I am not content merely to refute these views by way of refuting bifurcationism, because I see them as expressions of some fundamentally mistaken philosophical orientations toward the practical realm that themselves survive the refutation of bifurcationism. In the second and fourth chapters of this dissertation, I have provided definitions, labels and refutations for two of these orientations of thought.

The first of these is an orientation that I call "pre-critical empiricism" about the concept of a practical option. What I mean by this is a conception of practical options according to which practical options are things that the world independent of us somehow does all the work of *putting* into our minds, as opposed to a (post-"Copernican revolution") conception of practical options according to which they are best understood as somehow the *work* of our own cognitive faculties. Every occasion on which we are prompted to manifest practical excellence, on this conception, is one in which we are *given* various options, and what it is to manifest practical excellence on one of these occasions is always for us merely to *select* the right one.

A pre-critically empiricist conception of practical options can often seem to be an implicit commitment of the very practice of casuistry, by which I mean moral philosophizing that proceeds by attending to thought-experiment cases like the trolley problem, since thought-experiments depend for their usefulness on the fact that in them various facts are given: in the absence of given facts, there *is* no thought-experiment. Casuistry as such, however, is innocent of this implicit commitment. What is true is that the practice of casuistry necessarily involves abstracting away (sometimes for legitimate purposes) from the facts attention to which is

10

necessary to avoid falling into pre-critical empiricism about practical options. And since casuistry is so widespread in current philosophical literature, it should not be a surprise that pre-critical empiricism about practical options often seems inevitable.

Very roughly speaking, pre-critical empiricism about any concept can be refuted by pointing to a kind of unity for the necessity of which it cannot account that is characteristic of the items falling under that concept. The refutation of bifurcationism that I have already sketched, which uncovers a kind of unity characteristic of exercises of agency, can for that reason be thought of as a refutation of pre-critical empiricism about the concept of an *ethically valuable consequence* (or if you like, as a refutation of pre-critical *rationalism* about the concept of an *ethically valuable intention*). Ethically valuable consequences are not given to us as ethically valuable consequences independently of everything else, so that the ethical value that other elements of our practical lives possess can be derived from theirs; rather, ethically valuable consequences enjoy a sort of unity with ethically valuable intentions that is effected in a bit of the work of our practical cognition that I call an exercise of agency.

The key to refuting pre-critical empiricism about practical options similarly turns upon clarifying the way in which various things are unified in any practical option, a thing that we can do. I claim that we can find the relevant sort of unity of a thing that we do in the concept of a task implicit in our everyday talk of multitasking. It should be plain that this concept involves a special kind of unity from the fact that when we say that someone is multitasking, we say that she is doing more than one thing at the same time. For if this can be so, then it must be that *some* of what she is doing at that time is unified together into one task while the rest of what she is doing at that time is unified together into another task, so that the concept of a task certainly involves a kind of unity that we may seek to clarify. The various constraints that I articulate on

how this unity is to be understood hinge upon what I call the *promotion* relation, which is the relation in which two of the things an agent can do stand when she does one of them *in order to* do the other. For this promotion relation enables us to distinguish what one is doing in order to do A from what one is doing in order to do B. So, if one is doing A not in order to do anything else and one is also doing B not in order to do anything else, then we may (I claim) use these "doables" A and B to mark the distinct tasks that an agent is doing at a time by saying that what is unified under the task marked by the doable A is whatever the agent is doing *in order to* do A, while what is unified under the task marked by the doable B is whatever the agent is doing *in order to* do B. (So it may be that there is significant overlap in the contents of two different tasks, since there may be quite a lot that one is doing both in order to do A and also in order to do B.)

So the unity characteristic of a task is a kind of unity in which various doables may be bound together under another doable by the mediation of the promotion relation. Now since our practical options are the various tasks that we might perform in any situation, the question we must pose is how a pre-critical empiricist about practical options can regard the not-yet-selected options with which a practical situation furnishes an agent as having this kind of unity. And pre-critical empiricism about practical options is inconsistent with what I claim is the right answer to this question. For since pre-critical empiricism about practical options regards the unity of a practical option as merely *given* to us by the world, it is committed to denying that this unity is already the work of our cognitive faculties. But that is exactly what it is. For this unity, in which various doables are bound together by the mediation of the promotion relation, derives from how the agent may do one thing *in order to* do another. That is, it derives from the instrumental-justificatory relations in which various doables can stand for her.

So in order for a practical option to have content (to be one practical option rather than another), we must know how to situate it in instrumental-justificatory relations. But just knowing how to do this with that doable is *already* recognizing something of its justificatory valence. And this fact goes to refute pre-critical empiricism about practical options because recognizing some justificatory valence is *already* doing something—forging the practical option according to our own conception of the promotional relations among doables and thereby taking responsibility for a justificatory relation between one thing we can do and another thing we can do. Even though (perhaps) there are occasions on which we can act in a certain way "for no particular reason," it would not be *acting* in that way (as opposed to some other), as opposed to merely having something of that sort (as opposed to some other) happen in our lives, unless we were practically deploying the concept of acting in that way; and having the concept of acting in that way ready for deployment *is* knowing how to connect the concept of acting in that way instrumentally to concepts of acting in other ways. So our recognition of these orienting possible promotional relations must be written in to the practical options themselves, and since the recognition of these promotional relations is *ours* and not the *world's*, it must be our work and not that of the world that writes it into them.

1.4   FOURTH SECTION

Before refuting the second of the philosophical orientations that tempt us into consequentialism or intentionalism, I have two points to make about the notion of a non-instrumental reason that I think fall out of the distinctions deployed in my account of multitasking. The first point is about the possibility of multi-justification of ends—justification, that is, by more than one non-

13

instrumental reason. I defend this possibility against any view that attempts to deny it by exhibiting what I call the "vector sum structure." A view exhibits the vector sum structure if it maintains that there is a special property Φ such that the only non-instrumental practical reason an agent can have for doing a doable A is that *A is* (or would be) *Φ*. Candidates for this special property Φ could perhaps be the properties of *being the agent's duty*, *being to the agent's advantage*, *being such as to promote world utility*, or perhaps something else. It would *seem* that adopting a view of this kind would be an effective way to avoid recognizing the possibility of multi-justification of ends. But I argue that, in at least one significant sense, any view with the vector sum structure must recognize nonetheless that the multi-justification of ends is possible.

This argument proceeds on the premise that practically justified multitasking is possible. A Vector Sum theorist can only countenance the possibility of practically justified multitasking by admitting that, although a doable must have the property Φ in order for an agent to be justified in doing it as an end, nonetheless more than one doable can have the property Φ at the same time. If that is so, then the Vector Sum theorist must grant that a single situation can furnish the agent with distinct constellations of considerations, one of which justifies one of these doables A by entailing that A is Φ while the other constellation justifies the other doable B by entailing that B is Φ. And if the Vector Sum theorist grants that it is possible for there to be multiple constellations of considerations sufficient to justify an agent in doing a doable as an end, then she will have undercut her aspiration to claim that there cannot be more than one non-instrumental reason for an agent to do a single doable A. For once she grants that a single situation can furnish us with multiple distinct constellations of considerations each of which is sufficient to confer a justificatory status on the doing of doables as ends, the claim that there

could not be more than one such constellation that attached its justificatory force to a single doable A would then be entirely unmotivated.

This is not to say that the dispute between advocates and opponents of the vector sum structure is *entirely* verbal. It is not, for its advocates maintain, while its opponents deny, that no consideration ever confers non-instrumental justificatory status on a doable A except through the conduit furnished by the consideration, if it obtains, that *A is Φ*, where the property Φ is sufficiently contentful to *explain why* doing A is justified instead of merely saying that it is. Our point is just that this dispute has limited significance for the question of whether multi-justification of ends is possible, since as long as they think that practically justified multitasking is possible, advocates and opponents of the vector sum structure will have to agree that *indirect* multi-justification of ends is possible, by *distinct* constellations of considerations each of which confers non-instrumental practical justificatory force *sufficient* to constitute a non-instrumental reason to do an end, and that is good enough for us.

The second point I have to make about the concept of a non-instrumental practical reason is that we may derive the unity of a single non-instrumental practical reason (if such a thing exists) from the unity of a set of doables into a single task that it justifies. The argument for this claim proceeds by noting that if it were not true, then we would face an infinite regress in which the practical justificatory force of every non-instrumental reason to do a doable A was always divisible without remainder into bits of justificatory force that attached to the doables an agent might perform in order to do this doable. The unity characteristic of a non-instrumental practical reason reason is therefore coordinate with a kind of unity characteristic of a task.

1.5   FIFTH SECTION

My label for the second of the philosophical orientations I have identified that tempt us into consequentialism or intentionalism when conjoined with bifurcationism is the "command picture of morality." This is a conception of morality on which an agent essentially relates to morality as subordinate to a system of commands, constraints, requirements, or something of the kind, for someone in the grip of this picture thinks that morality requires the agent always to do certain things while it is in principle indifferent to whether she does them on the basis of an understanding of why she should. An agent could in principle meet these criteria for ethical excellence even if all of her actions flowed from the motivations of obedience; and the task of ethical philosophy that proceeds from this conception is merely, as Mill says, "to tell us what are our duties, or by what test we may know them"[10]—to *command* us rather than to *connect* us to those duties by revealing them as proceeding from our authentic recognition of reasons to act. Perhaps no one actually thinks that a rational agent could always have the motivations of obedience, and that for that reason we must cultivate other sorts of motivation; but the command picture of ethics is committed to the claim that an agent would count as ethically excellent if, *per impossibile*, she always did act from those motivations.

But I say that it is not enough for an agent merely to do what would in some sense count as the right thing to do. She must also do those things for the right reasons if her action is to have ethical value. For a theory of ethical excellence for *rational* agency must do justice to the rationality of that agency—the fact that such agents characteristically act as they do on the basis of their own appreciation of why it would be good to act in that way. Cases in which the

---

[10] *Utilitarianism*, p. 18.

motivations of obedience happen to be consistent with the ethical excellence of the agent—as when military personnel obey orders in a just war, if there is such a thing—are necessarily anomalies, for in these cases there is a *division* of rational labor. The single phenomenon of responding practically to a reason is here split up between a commander, who is responsible for appreciating the practical justificatory force of a reason, and a subordinate, who is responsible for doing what counts as a response to it. To think that ethically excellent action could always be like that is to alienate rational conduct from its rational ground.

I therefore take it as a guiding principle for ethical philosophy that it must make sense of the fact that an agent must do the right things *for the right reasons* in order for her action to possess ethical value. The *telos* of ethical philosophy is not to arrive in *any way* at a specification of what ethics requires of us, but to arrive at such a specification in a way that moves us to pursue that end for the right reason, the reason for which a good person would pursue it. I try to make this claim vivid by formulating it as the claim that no action can possess ethical value unless its agent's relation to her reasons for acting is inconsistent with the possibility that the "ought-ought gap" might open up for her with respect to them. This is the gap expressed when someone says something of the form, "I know that I ought to do A—but why should I do it?" An agent who says such a thing may know very well how to demonstrate, perhaps from secure *a priori* principles, that she has a duty to do A, but she nonetheless finds this demonstration inadequate to move her in the right way, and that is unacceptable. She can explain very well that a good person would do A in her circumstances, but what she needs is to relate herself to the concept of a good person in such a way that, for her, recognition that a good person would do A in her circumstances for such-and-such a reason *is the same thing as* practical recognition of the rational force of this reason that manifests itself in her doing A. An agent in

17

whose life the ought-ought gap has opened up may in some sense have the concept of a good person—the concept of someone who acts in this way—but she is insufficiently moved by her possession of this concept to act as a good person does. The bulk of my final chapter is dedicated to understanding how this gap in her motivational structure can be overcome, and thereby to deepen our understanding of the conduct of ethically excellent agents who do not suffer from this difficulty.

1.6   SIXTH SECTION

It is helpful at this point to observe that someone in whose life the ought-ought gap has opened up has a problem analogous to the problem that faced Achilles in Lewis Carroll's famous dialogue between that hero and the Tortoise.[11] Achilles wants to perform an inference from two premises, the first of the form *p* and the second of the form *if p then q*, to a conclusion of the form *q*. But the Tortoise dupes him into thinking that before he can conclude *q* he has always to write down, as another premise, his understanding of the justificatory relation between the premises he has already affirmed and the conclusion *q*. Carroll makes it clear that no rational transitions from premises to what they justify could ever take place under this constraint. Achilles' understanding of the justificatory relation between his premises and his conclusion must indeed be somehow operative in the rational transition that he proposes—but not as any collection of premises, for as premises this understanding will be no help. What he needs is therefore a way of affirming that *p* and *if p then q* together entail *q* that is enough to move him to

---

[11] Carroll, Lewis. "What the Tortoise Said to Achilles". Mind, n.s., 4 (1895), pp. 278–80.

conclude *q*. And in general anyone who tries to draw some conclusion (practical or theoretical) from a collection of reasons must affirm the justificatory relation between those reasons and that conclusion in a way that is *sufficient* to move her to judge or act in accord with that justificatory relation. There must always be reasons affirmed in the usual way and a way of taking them up *affirmed in another way*.

I use the word "attunement" for this other way of affirming, and I say that just as Achilles needed to *attune* himself to the cogency of *modus ponens* in order to bridge the gap that opened up for him between the conclusion *q* and the premises *p* and *if p then q*, an agent in whose life the ought-ought gap has opened up must *attune* herself to the concept of a good person in order to bridge it. For the concept of a good person, like that of *modus ponens*, is the concept of a rational nexus of a certain sort through which the justificatory force of reasons generates what they are reasons to do or to think. A good person is a person who is attuned to the concept of a good person—the reasons on which such a person acts will constitute one side of this rational nexus, and her conduct will constitute its other side. And to be attuned to the concept of a good person in this way it is precisely not enough for an agent merely to say or think of herself that she is a good person. That can help her no more than it helped Achilles to add as a premise that *p* and *if p then q* together entail the conclusion *q*.

I think that the foregoing remarks on the centrality of questions of motivation to ethical philosophy articulate, in perhaps unfamiliar terms, Kantian thinking about these matters. Kant's starting point in the *Groundwork* is the idea that the good will is the only thing that can be conceived as good without qualification, and he emphasizes that the good will possesses this value entirely within itself; this is what I affirm when I say that a good person *is* a person who is attuned to the concept of a good person—she need not *wait* for her goodness to manifest itself in

good action in order for her goodness to be complete. I take Kant's claim that an action must be performed from the respect for the moral law in which the good will consists if it is to have any moral worth to be his way of affirming my guiding principle that no exercise of agency can possess ethical value unless its agent's relation to her reasons for acting as she does is inconsistent with the possibility of her being subject to the ought-ought gap with respect to them. And his characterization of this respect as an *a priori* feeling that an agent generates within herself from the representation of a worth that thwarts her own self-love[12] is his way of affirming that a good person is attuned to the concept of a good person in a way that moves her to act as good people do, for a good person disregards whatever inclinations she may have to act otherwise than as the relevant reasons are reasons to act.

I depart from Kantian doctrine, however, in two important respects. First, I maintain that Kant was overhasty to maintain that there is only one appropriate way of being related to the reasons to which a good person is responsive, and which consists in being motivated by mere respect for the moral law. I do agree that an action possesses ethical value only insofar as it manifests respect for the moral law (or in other words, a relation to the reasons there are to act that is inconsistent with the opening up of the ought-ought gap), but I prefer to put this point by saying that any way of being moved to act as a reason requires must be respect*ful* of that reason, rather than that it be *mere respect itself*. This leaves open the possibility that *respect* is a determinable having many determinations (*courage*, *benevolence*, *justice*, and so on), and can never be a property of an agent except insofar as one of its determinations is a property of her. Just as an object in no way fails to be a *colored* thing in virtue of being *colored because*

---

[12] *Groundwork*, 401n1

chartreuse, I say that a way of being moved to act in no way fails to be *respectful of the reasons that prompt it* in virtue of being *benevolent*.

I also dispute Kant's doctrine that there is only one concept—the concept of excellence in *mere* practical rationality, or of a good person *in general*—in the attunement to which being a good person consists. It is because he endorses this doctrine that the only concept to which he tries to attune us in his ethical works is the concept of practical rationality. Here again I prefer to think of this concept as a determinable having many determinations of which the concept of a good human being may be one. Kant says that the concept of excellence in mere practical rationality must be the concept in the attunement to which being a good person consists because he is impressed by the fact that we may know not only of human beings, but of any practically rational being at all, that it is wrong for them to tell a lie. But this conclusion does not follow. What follows is that the access one has to that concept the attunement to which constitutes obviating the ought-ought gap has to be access, *perhaps among other things*, to the concept of mere practical rationality in general. But this could be the case if it were right to regard the concept of a good person as a determinable to which we could only gain access by attuning ourselves to one of its determinations like the concept of a good human being.

One way of understanding this suggestion is that there is a third test that any maxim must pass, in addition to the contradiction in conception test and the contradiction in the will test proposed by Kant, in order for us to know that its agent could act on it with a good will. The contradiction in conception test rules out any maxim that somehow contradicts itself; the contradiction in the will test rules out any maxim that somehow contradicts something that *any* rational being necessarily wills. It is no accident that the tests already proposed by Kant are valid for *any* rational being. But the third test that I propose—call it the "contradiction in the

21

determination test"—rules out any maxim that somehow contradicts something necessarily willed by any rational being that bears, not just *any*, but the *same* determination of the concept of a rational being borne by the agent whose maxim it would be.

I do not think that this proposal—which is an attempt to marry Kantian meta-ethics to an Aristotelian sensitivity to the nuances of human life—necessarily undermines the claim of ethics to be an *a priori* discipline. To know something *a priori* is to know it without any contribution from experience, so it is to know it exclusively on the basis of the activity of one's faculty of judgment. So the set of things that a person can know *a priori* depends entirely on the nature of that person's faculty of judgment. So the faculty of judgment possessed by human beings can generate *a priori* knowledge that differs in content from the knowledge that other rational beings can generate *a priori*, if their faculties of judgment are relevantly different from ours. And it is plausible to think that their faculties of judgment may be relevantly different from ours in their practical employments because there may be specific standards of practical excellence for a human being that differ from those that apply to a rational being of some other kind; that is what it means to say that the concept of excellence in mere practical rationality is a determinable concept with potentially many determinations.

If this is so, then we can attune ourselves *a priori* to the concept of human life at its best, or of human happiness, and it is only thereby that we can attune ourselves to the concept of a good person. Then it is on the basis of such an attunement that we can generate the good will in ourselves. The more thoroughly we attune ourselves to the concept of happiness with which we confront the world, the less often we will find ourselves subject to the ought-ought gap. For us, becoming good people therefore requires us to devote ourselves to strengthening this attunement to the possibility of happiness by meditating on that concept in an exploratory way. That is,

finally, the way to obviate the ought-ought gap and become good people is by open-minded *daydreaming*, an activity in which we deploy our practical imaginations in order to elaborate to ourselves not only what happiness can be like for us, but also in which we learn to be confident that we are capable of elaborating this concept more fully in response to novel situations—that is, we learn through daydreaming that we have happiness (goodness) in ourselves as an indefinitely extendable possibility with which we are capable of overcoming difficulties and realizing the good as it is for us. I do not claim that the way out of every difficulty is to start daydreaming; I claim that the activity of daydreaming is necessary to *prevent* certain difficulties—characteristically, the ought-ought gap—from arising.

I have engaged so intensely with Kant in this part of my dissertation in part because I think I can generate a Kantian argument for this conclusion. Kant certainly maintains that all knowledge of an object requires a faculty of intuition that attunes us to its form, so as to make possible the synthesis that such knowledge consists in. In the case of theoretical knowledge, the object of knowledge is the world, the form of this knowledge is space and time, and the faculty of intuition that attunes us to the form of this kind of knowledge is the faculty of imagination. But what about the case of practical knowledge? It is clear that the *object* of practical knowledge is the highest good. It is not exactly clear what Kant thinks the form of this knowledge is supposed to be. But whatever its form is, we are entitled to ask what the practical faculty of intuition is that attunes us to this form, and I suggest that the faculty of open-minded daydreaming (the practical imagination) is exactly the faculty that stands to *a priori* practical reason in the relation in which the theoretical faculty of imagination stands to *a priori* theoretical reason.

The body of the dissertation I have written therefore concludes with this affirmation of the role I have just suggested we must ascribe to daydreaming in an ethical life. A clearer understanding of this role will enhance our conception of an unfairly neglected and crucial aspect of ethical conduct; it will lead to a more humane and respectful sort of ethical philosophy that does not merely command people to do the things that any good person would do; it will help those who suffer from alienating conditions like the ought-ought gap to find their way to living better lives; and it will make it possible for ethical philosophy to escape for good the apparent inevitability of the debate between consequentialism and intentionalism.

2.0   THE ARGUMENT FROM FAILURE

2.1   INTRODUCTION: RECONCEIVING THE DEBATE BETWEEN CONSEQUENTIALISM AND DEONTOLOGY

Ethical philosophy of a certain sort tends to concern itself with the question of which actions would be ethically right, good, permissible, obligatory, or perhaps something else, to perform. It is not controversial that such questions about what to do are at least some of the questions that ethical philosophy should aspire to answer. But I think that attention to these questions tends to obscure the importance of questions of another sort—let us call them questions about motivation—that ethical philosophy should also aspire to answer. For reasons that will become clear in my final chapter, I think that these questions can only be answered by investigating the role of the imagination in an ethical life. But first it is important to generate a sense for their importance by pointing to the way in which certain answers to even more basic questions about the metaphysics of agency tend to be presupposed by those who only try to settle questions about what to do. I will try to do this by providing a novel characterization of the debate between consequentialism and deontology.

Now, I find myself in the uncomfortable position of thinking both that the deepest axis of disagreement between consequentialism and deontology is very different from how it is commonly conceived, and also that the idea (which is as widespread as it is unrecognized) that a

disagreement can turn on this axis can fairly easily be shown to be a consequence of a confused

metaphysics of action. If what I have to say is to interest an audience of consequentialists and

deontologists, therefore, I must first convince them to recognize as deeply their own, shortly

before refuting it, a view that will seem initially quite alien to their way of thinking. To do this I

will show first that the way in which this debate is customarily conceived—namely, around the

question whether, and in what way, morality involves an element of agent-relativity—cannot do

it justice, so that the right way to conceive it must be otherwise. Having done that, I will try to

show that the spirit of the disagreement that the concept of agent-relativity was developed to

clarify in fact turns on what I think is a deeper issue concerning what I will provisionally call the

original locus of ethical value. And I hope that these preliminaries will generate sufficient

interest in my argument that both consequentialism and deontology are inspired by a confusion

about what sort of issue this really is.

So let us consider the question whether, and in what way, morality involves an element of

agent-relativity. When matters are framed thus, deontology finds itself charged with a kind of

inconsistency deriving from its supposed commitment to the claim that morality furnishes us

with what are called agent-centered restrictions. An agent-centered restriction is, in Scheffler's

canonical formulation, "a restriction which it is at least sometimes impermissible to violate in

circumstances where a violation would serve to minimize overall total violations of the *very*

same restriction, and would have no other morally relevant consequences."[13] It is plain that lines

of defense against this charge available to deontologists will fall into two categories:

deontologists can claim that they are not in fact committed to the claim that morality furnishes us

with agent-centered restrictions, or else they can claim that this commitment does not render

---

[13] Samuel Scheffler, "Agent-Centered Restrictions, Rationality and the Virtues," Mind 94 (1985): 243 (my
emphasis).

their position inconsistent. I think that the former line of defense is more promising: deontology is not, contrary to certain appearances, committed to the claim that morality furnishes us with agent-centered restrictions. Developing this line of defense requires showing how those appearances can be understood differently.

The sort of case in which it most vividly appears that deontology is committed to the claim that morality furnishes us with agent-centered restrictions is exemplified by a moral dilemma in which an agent must choose between two courses of action. Her first option is to kill one person; her second option is to refrain from killing anyone and thereby cause or make it more likely for five other people who would otherwise have been left unmolested to be killed by another agent. Let us give these people names. Our heroine is Olga; the person she herself will kill if she chooses her first option is Penelope; Quisling is the one who will kill five others if Olga does not kill Penelope; and the five people who will be killed if Olga does not kill Penelope are Abner, Baxter, Cuthbert, Dexter and Edgar. Now let us consider all of the deontological theories committed to the claim that in this situation morality forbids Olga to kill Penelope—and thereby, according to opponents of these theories, to the claim that morality furnishes us with agent-centered restrictions. It should be clear that different deontological theories will give different accounts of this moral restriction. Whether a version of deontology is committed to the existence of agent-centered restrictions will depend on whether it is impossible to provide an account of this moral restriction on Olga's actions according to which this restriction is not an agent-centered one, in the sense defined by Scheffler's canonical formulation. And I claim that there is a plausible account of this moral restriction on which it is not agent-centered. For there is a plausible account of this moral restriction on which the moral restriction that Quisling would

violate in killing, say, Abner is not, in Scheffler's words, the "*very same*" moral restriction that Olga would violate if she were to kill Penelope.

This is plausible if we think, first, that morality forbids each person to kill each other person *because* every pair of people determines a two-place moral relation that is itself the source of moral restrictions on the actions of each of its members insofar as they bear on the other, and, consequently, that these moral restrictions are *distinct* from the moral restrictions derived from any other such two-place moral relation. (Notice that to reverse the direction of this *because* relation is to endorse a version of platonism.) Olga's obligation not to kill Penelope is not the very same moral restriction as Quisling's obligation not to kill Abner. And so Olga would not be minimizing violations of the very same restriction she would violate by killing Penelope; she would be minimizing violations of a set of other restrictions all of which merely have *something in common* (person-killing) with the restriction she would be violating.

It is an old criticism of Kant's ethical theory that any proposed course of action can have its maxim gerrymandered to pass his contradiction tests; the thought I here suggest is that, in every case in which a deontological theory of ethics appears to be committed to the claim that morality furnishes us with an agent-centered restrictions, the maxims of the various actions possibly to be performed in the envisioned scenario can be formulated to ensure that the restrictions that would be violated in those actions would not be the very same. Whether or not these formulations count as gerrymanderings will depend upon the details of the deontological theory in question. For the attribution of agent-centered restrictions to an ethical theory must always among other things be the attribution to that theory of criteria of identity for moral restrictions. Since every restriction has *something* in common with every other restriction, the question is whether the acts possibly to be performed in an envisioned scenario would have

28

enough in common to constitute violations of the very same restriction, and that is what depends on the details of the deontological theory in question. I have given an example of a plausible deontological theory that can endorse the claim that killing a person is always wrong without committing itself to the claim that every killing is a violation of the very same restriction.

Not only is this theory plausible, I think it is at least plausible to attribute it to Kant. For Kant's formulation of the categorical imperative as the claim that we ought always to act in such a way that we treat humanity as an end and never merely as a means[14] can plausibly be interpreted as involving the claim that each person is herself the source of the various moral restrictions there may be on our conduct insofar as it bears on her. So it seems quite plausible that any moral restriction Kantian theory may reveal on Olga's killing Penelope will not be the very same restriction as any restriction it may reveal on Quisling's killing Abner. Morality forbids us to kill Penelope because that would violate the restriction on treating *Penelope* merely as a means, whereas morality forbids us to kill Abner because that would violate the quite different restriction on treating *Abner* merely as a means.

At least one strategy for providing criteria of identity for moral restrictions is to derive them from the reasons for which good people act in accord with those restrictions. Every deontological account of ethics generates, or at least purports to provide an algorithm for generating, a system of general restrictions or imperatives that supposedly manifest *what* it is right, obligatory, morally worthy, or something like that, to do. That is, it provides a picture of *what* good people do. Many critics of deontology—and, to be sure, many deontologists—are misled by this fact into thinking that a deontological account of ethics is obliged to conceive the moral grounds for forbidding us in each case to act contrary to these imperatives at the same

---

[14] *Groundwork*, 429.

29

level of generality. But this is a mistake. The actual thought expressed by these generalizations can be something like: *in each case, there will be* a *moral restriction on the agent's doing A*. And it simply does not follow from a thought of this sort that the moral restriction on the agent's doing A in one case will be the very same moral restriction as the moral restriction on the agent's doing A in another case. For a picture of *what* good people do—we may call it a "command picture of morality"—does not have to be interpreted as a depiction of *why* good people do the things it depicts them as doing. And if the criteria of identity for moral restrictions are linked to the reasons for which good people act in accordance with them, then we may say that *we do not yet have a moral restriction fully in view when we merely have in view* what *morality obliges or forbids us to do*.

The debate about agent-centered restrictions therefore does not get to the heart of the disagreement between deontology and its opponents. The illusion that this debate gets to the heart of that disagreement has had an unnaturally long life in large part because the command picture of morality has misled many deontologists themselves into thinking that they are essentially committed to the view that each good person acts in the way that she does on the basis of reasons (typically, reasons of *obedience*) at the same level of generality as the characterizations of their conduct provided by that picture. But general characterizations of morally excellent conduct in imperatival form are not yet necessarily characterizations of identity-conditions for moral restrictions, and thereby of the conditions for what counts as living up to moral restrictions. Deontologists who *take* them to be so are making an additional philosophical move not forced by deontological theory as such.

2.2   BIFURCATIONISM

So much, I hope, for the idea that we would settle the debate between consequentialism and deontology if we were to settle the debate about agent-relativity. I would now like to juxtapose quotations from two ethicists who disagree in a way that we need not characterize in terms of agent-relativity. Kant the deontologist:

> A good will is good not because of what it effects or accomplishes, nor because of its fitness to attain some proposed end; it is good only through its willing, i.e., it is good in itself…. Even if, by some especially unfortunate fate or by the niggardly provision of stepmotherly nature, this will should be wholly lacking in the power to accomplish its purpose; if with the greatest effort it should yet achieve nothing, and only the good will should remain (not, to be sure, as a mere wish but as the summoning of all the means in our power), yet would it, like a jewel, still shine by its own light as something which has its full value in itself. Its usefulness or fruitlessness can neither augment nor diminish this value. Its usefulness would be, as it were, only the setting to enable us to handle it in ordinary dealings or to attract to it the attention of those who are not yet experts, but not to recommend it to real experts or to determine its value.[15]

And Mill the consequentialist:

> The creed which accepts as the foundation of morals "utility" or the "greatest happiness principle" holds that actions are right in proportion as they tend to promote happiness; wrong as they tend to promote the reverse of happiness.[16]

In these remarks we catch sight of an opposition between two different conceptions of the original locus of ethical value, an opposition in which Mill maintains while Kant vigorously denies that this locus is to be found among the consequences that actions "promote." In order to make this opposition precise, I will work with a conception of the difference between

---

[15] *Groundwork*, Ak. 394.

[16] *Utilitarianism*, p. 12. In an interesting article ("Consequentialism in Modern Moral Philosophy and 'Modern Moral Philosophy'", in Oderberg and Laing (eds.), *Human Lives: Critical Essays on Consequentialist Bioethics*), Cora Diamond points out that it is possible to read Mill's use of "tend" in this formulation in such a way as to exonerate him from counting as a consequentialist, at least according to the usage for "consequentialist" coined by Anscombe in "Modern Moral Philosophy," and also that Anscombe herself did *not* regard Mill as a consequentialist. I cannot get into details about these points here, except to say that even if it is true that Mill does not count as a consequentialist according to Anscombe's usage, that usage will then label a feature that some philosophers have in virtue of their particular ways of working out positions animated by the same thought as Mill on the question concerning the original locus of ethical value, which is the fundamental issue on which Mill and Kant differ in these quotations.

consequentialism and deontology on which consequentialism is approximately the view that the ethical value of an action is logically posterior to (because it is explained by, inherited from, or something like that) some sort of value possessed by its consequences. And the proponents of deontology, in contrast, maintain that an action possesses at least some of its ethical value in virtue of what it is, all by itself. In particular, such an action has some of its ethical value in virtue of what it *is* and not in virtue of what it *brings about*.

It may be objected that this characterization of deontology is unfair, since—as we have just seen—Kant himself maintains that the good will would lose none of its value "even if… with the greatest effort it should yet achieve nothing," so that what we should say instead is that a deontologist locates the original locus of ethical value in (say) the intention with which an action is performed rather than in the action itself. All I can say in response to this is that carving up the exercise of agency in this way would make the action itself into a consequence, one of the external elements of the exercise of agency that derives ethical value from its internal elements, according to the deontologist. And it does not matter—at least for the purposes of my argument—where we draw this boundary between internal and external elements, so long as we recognize that the disagreement between consequentialism and deontology can be understood as a disagreement about the side of this boundary on which we may find the original locus of ethical value. Where this sort of boundary should be drawn hangs on whether we ought to draw one, and it is that prior question that I answer in the negative. So for the remainder of this section I will continue to speak of actions as the internal element opposed to consequences, but this is merely a convention: the main thing is the idea of a boundary between the inner and the outer. Indeed, because the word "intention" wears its internality much more clearly on its sleeve,

it will suit our purposes to use the word "intentionalism" henceforth as a more precise term than "deontology" for the position that takes itself to be opposed in this way to consequentialism.

Now, our way of putting the contrast between these two possible accounts of the value of death prevention should raise philosophical eyebrows. It is not clear how to get a grip on a disagreement that is supposed to be about whether an action has some of its ethical value in virtue of what it *is* rather than in virtue of what it *brings about*. An action that is a preventing of someone's death, for example, is often thought by deontologists to have ethical value because it *is* a preventing of someone's death, rather than because it *prevents* someone's death. But what is there to choose between an action's *being* a preventing of someone's death and its *preventing* someone's death? It appears that the intentionalist and the deontologist who squabble over this agree that appropriate tokens of the type *death prevention* confer ethical value upon the exercises of agency that involve them while disagreeing about whether exercises of agency involve these tokens by *being* them or by *causing* them. That is, their disagreement appears really to be about the metaphysics of exercises of agency; but I think it is clear that this dispute about one issue in the metaphysics of agency necessarily takes place in the context of a more fundamental agreement about another.

To see this, consider first the consequentialist view. Since this view maintains that the ethical value of an action is in some important sense logically posterior to the value of its consequences, it must maintain that all of the facts (and they can often be quite complicated) that determine this value must be facts about its consequences. It follows that it is a standing option for consequentialists to say, of any value-determining fact we might propose as a counterexample to consequentialism, that it is actually a fact about the consequences of the relevant action rather than a fact about the action itself; and if facts about the action itself help to

33

determine its value, then that is only because we can consider the action itself to be (one of) its consequences. Thus, in their zeal to defend consequentialism, consequentialists may expand the concept of an action's consequences to the point where (they think) it cannot but include every fact that could be important to the determination of the worth of the action. And the reverse expansion is liable to occur among intentionalists, who expand their concept of an action (or, if you prefer, of an intention) to the point where (they think) it cannot but include all of these same important facts within itself, including the consequences that have a bearing on the ethical value of an action.

The important point here is that in order for the debate between intentionalists and consequentialists to be coherent, its parties must agree that at least one of these expansive processes is out of order. For otherwise each party to this debate would end up claiming the same territory under a different flag—and as long as the territory is the same (as long as each party to this debate agrees about which facts are relevant to a determination of the ethical value of an action) it is unclear what difference in content is marked by the different words stitched on their flags. As long as we have a suitable grip on the concept of the totality of the facts relevant to determining the ethical value of an exercise of agency it does not seem to matter whether we use the word "action" or "consequence" when we squabble about what that concept includes or about how those facts determine. In any case, it is at least not *yet* clear why this difference in terminology should be thought to matter. There will always be substantive disagreements about which facts are relevant to ethical evaluation and how; but in the absence of a rigorous account of the metaphysics of exercises of agency, and in particular of the precise significance of the distinction between actions and consequences, it can serve no helpful purpose to say of some fact that it can have no bearing *because* it is merely a fact about the action, or about its consequences.

I have developed this dialectic between consequentialism and intentionalism in order to point out that the debate between these two philosophical traditions tends to involve a presupposition that we would do well to question. For although they do disagree about its direction, consequentialists and intentionalists agree that there is some priority relation between the ethical value of an action (or of an intention) and the value of its consequences. So philosophers of both persuasions are committed to the metaphysical claim that the distinction between actions (or intentions) and consequences is robust enough to underwrite a priority relation between the sorts of value that these different elements of an exercise of agency can have. I refer to this claim as "bifurcationism," since it is the claim that there must be a sort of fundamental bifurcation between different parts, or perhaps aspects, of any exercise of agency. There would be nothing left for consequentialists and intentionalists to disagree about if they were to reject bifurcationism. And since I myself deny bifurcationism, I find that there is indeed nothing to choose between these two camps. But what I have said so far merely gives us reason to be suspicious of bifurcationism; the remainder of the present chapter is dedicated to conclusively exposing it as an error.

2.3   THE ARGUMENT FROM FAILURE

Bifurcationism is a doctrine that comes in many different forms, but all of them have in common a structure according to which every intentional action possesses an "inner" component (a trying, an intending, a wanting, a pro-attitude, or something such) that is within our power, and an "outer" component (a doing, an accomplishing, an achieving, or something such) not really up to us, and fundamentally distinct from the "inner" component, although perhaps deriving from it in

35

various ways. (We must learn what "fundamental" means, if anything, as we go along.) I believe that the covert motivation for this widely and covertly held doctrine can be brought to the philosophical surface by considering the merits of a practical analogue of the Argument from Illusion, an argument that I shall refer to as the "Argument from Failure." I will lead up to this argument by making some remarks about the facts that enable the Argument from Illusion to get its foot in the door of the philosopher's mind, and by noting how analogous facts obtain in the practical realm.

Prominent among these is the fact that many of our belief-forming practices seem to involve an element of luck. For it is possible, as we all know (how do we all know this?), for it to perceptually appear to us that p on an occasion when it is not really the case that p, and this makes it seem as if the occasions on which it is really the case that p are occasions on which we have had a bit of good luck in the fact that p is true. And it is not just when we perceive that we seem to be subject in this way to an element of luck. For it is possible for us to be told that p, or to seem to remember that p, or so on, on an occasion when it is not really the case that p; and that is all that is really needed for the Argument from Illusion, which has its historical home in the philosophy of perception, to establish colonies in the philosophy of testimony, or in the philosophy of memory, or anywhere else. It therefore seems appropriate to isolate the following as the first premise of the Argument from Illusion from which I shall try to develop a practical analogue:

I1. There is slack between its perceptually appearing to us that p, all on its own, and the fact that p.

where "slack" is just a fancy word for the seeming fact that it is always possible for the appearance that *p* to fail to guarantee the truth of *p*—the seeming fact, in other words, that any such alignment of appearance and reality requires an element of luck.

It should be plain that an element of luck is as characteristic of occasions of action as it is of any other sort of occasion on which the Argument from Illusion can get a grip. For it is possible for us to try to do A without ever actually getting A done, and this is true not just in cases where A is something we can do no better than try to do, but also in cases where A is something we have the greatest confidence in our ability to accomplish. For an accident can always divert our efforts in such a way that our trying to do A fails to result in, or blossom into, a successful doing of A. In the gap between trying to do A and successfully doing A, we therefore seem to find a sort of slack that needs to be taken up by luck. This gives us the first premise (I will sometimes call it the "slack premise") of the Argument from Failure:

F1. There is slack between our trying to do A, all on its own, and our getting A done.

Now, the Argument from Illusion, classically, is an argument about *knowledge*, about whether it is possible for us to know certain things perceptually. So the colonies that the Argument from Illusion sends into other realms that appear to possess the appropriate element of luck tend to retain this epistemic character. The version of the Argument from Illusion that pertains to testimony, for example, concerns whether we may *know* something on the basis of having someone tell it to us; the version that pertains to memory is about whether we may *know* something on the basis of seeming to remember it. In all of these cases, the Argument from

Illusion insists that the answer is no, and that is because its second premise, in the version that concerns perception, is the following:

> I2. We cannot know that *p* on any basis such that there is slack between it and the fact that *p*.

(And from the two premises I1 and I2, the dreaded conclusion follows:

> IC. We cannot know that *p* on the basis of its perceptually appearing to us that *p*, all on its own.

So all we can ever know on the basis of its perceptually appearing to us that *p* is that it perceptually appears to us that *p*.)

Notice that the second premise of the Argument from Illusion picks no favorites among the various ways in which knowledge can be thought to have a basis, so it will serve equally well for any version of that argument we choose to examine. Therefore, if we liked, we could incorporate this second premise without modification into the Argument from Failure to yield the conclusion that we cannot know that we are doing *A* on the basis of our trying to do *A*, all on its own.

I would like very much to write a chapter about this epistemic version of the Argument from Failure, because I follow Anscombe's famous doctrine that an action is intentional under a certain description only if its agent knows "without observation" that she is performing an action

that description truly describes.[17] Anyone who follows Anscombe in this doctrine is very likely to find the basis of this non-observational knowledge, as I do, in the agent's trying, deciding, wanting, intending, doing, practically committing to, or something like that, an action under the relevant description.   So this is at least one tradition in the philosophy of action whose representatives will be intensely interested in the merits of an Argument from Failure that denies that any such thing can furnish any such epistemic basis.

But not everyone will be interested in such an essay, since not everyone follows Anscombe's famous doctrine. Therefore, in order to keep them interested, we will make two adjustments for those who demur from it. First, we will make do here with the more modest claim that an agent is *trying* to do A only if she knows that she is *trying* to do A. Because it is more modest, this claim is more difficult to deny than Anscombe's famous doctrine, and it is clearly consistent with the main source of opposition to Anscombe's doctrine, which is a counterexample famously proposed by Davidson: "A man may be making ten carbon copies as he writes, and this may be intentional; yet he may not know that he is; all he knows is that he is trying."[18]   As the clinching remark "all he knows is that he is trying" explicitly registers, Davidson's canonical counterexample is not a counterexample for our more modest claim.

The second adjustment we will make for those who reject Anscombe's famous doctrine will be to consider an entirely de-epistemologized version of the Argument from Failure, a version we might call the "efficacious" Argument from Failure, since its second premise concerns only whether we ever really get done what we try to do, rather than whether we ever know that we are getting it done. The second premise of this argument will be as follows:

---

[17] *Intention*, §8.
[18] "Agency," *Essays on Action and Events*, p. 50

F2. We cannot do A by a certain means if there is slack between that means and our getting A done.[19]

And from premises F1 and F2, the dreaded conclusion follows:

FC. We cannot do A by trying to do A, all on its own.

This is the argument that I will call the Argument from Failure, and whose merits I propose to examine in the remainder of this chapter. It is valid, and it leads to global skepticism about our practical capacities, since the argument tells us that trying—which is all we can ever do on our own—is never enough to get anything done.[20] To invert Yoda's charming aphorism: "try or do not try, there is no do."[21]

2.4   AN ACCOMMODATION?

Descartes, among many others, is sometimes characterized as having tried to reach an accommodation with the Argument from Illusion, an uneasy truce that grants its soundness, but that nonetheless retains hope in the possibility that we can know more or less everything we think we can know on the basis of perceptual experience. This hope derives from the prospect of

---

[19] It may be worthwhile to wonder whether the Argument from Illusion can be de-epistemologized as well.
[20] As if that weren't bad enough, it should be clear that repeated meditation on this demoralizing result will eventually allow us to count as "trying" only the most basic of actions, if there are such things, and only mere acts of the will if there are not. I will write as if the Argument from Failure somehow still allows us to speak of trying, but my argument will be just as good if the reader sees fit to replace every occurrence of "trying" with an occurrence of (say) "intending."
[21] What Yoda says in *The Empire Strikes Back*, of course, is "do or do not do, there is no try."

working around the "slack" premise of the Argument from Illusion by devising fail-safe methods for forming beliefs on the basis of perceptual appearances, methods the fail-safe reliability of which we can certify without making any unwarranted assumptions or rationally dubious moves. For Descartes, this strategy took the specific form in his Third Meditation of trusting only those perceptions that are "clear and distinct."[22] We cannot know that *p* on the basis of its perceptually appearing to us that *p, all on its own,* but we can know that *p* on the basis of its perceptually appearing to us that *p*, if that appearance survives a diligent application of this fail-safe method. For there is, according to this Cartesian line of thought, no slack between a perceptual appearing to us that *p that survives this method* and the fact that *p*.

I do not aspire to determine whether this is indeed what Descartes was up to; the fact that it is at least something like a received myth concerning Cartesian doctrine is itself enough to show that this strategy for addressing the Argument from Illusion is a standing temptation worthy of discussion.   Accordingly, a version of this strategy, which I call "Cartesian accommodation," will be a standing temptation for anyone who wrestles with the Argument from Failure. Although trying to do A, *all on its own,* may be insufficient to guarantee that we get A done, we can nonetheless devise fail-safe methods for ensuring the success of our attempts to do A. We can make a list, check it twice, and so on—and if we apply these fail-safe methods of diligence to our trying to do A, then our trying to do A *in accord with these fail-safe methods* can guarantee that we are getting A done. ("Guaranteeing" is taking up all the slack.) In this way, we may perhaps be able to grant the soundness of the Argument from Failure without losing hope in our ability to do anything other than try.

---

[22] *Meditations on First Philosophy*, in *The Philosophical Writings of Descartes*, J. Cottingham, R. Stoothoff and D. Murdoch (eds.), p. 24.

I said that these strategies of Cartesian accommodation are ways of working around the slack premises, but it is in fact quite a delicate question whether we ought instead really to say that they are attempts to deny them. To say that a perceptual appearance that *p*, all on its own, is insufficient to guarantee that *p*, but that a perceptual appearance that *p* that survives a sufficiently diligent method *is* sufficient to guarantee *p*, seems to involve a kind of verbal quibble about what is meant by the phrase "all on its own," and the same goes, *mutatis mutandis* and *a fortiori*, for the case of trying to do A. For what counts as trying to do A, for someone who understands which methods are sufficient to guarantee the success of her trying and who really wants to guarantee that success, surely *just is* using those fail-safe methods, if such there are, so that it seems inaccurate to say that the slack premise is true of someone who has grasped those methods. We here glimpse the fact that what counts for an agent as trying to do A will vary greatly depending upon what she happens to know about how to do A; the same goes for what counts for a perceiver as basing a belief that *p* on its perceptually appearing to her that *p*. The strategy of Cartesian accommodation simply asserts that, at least for some of us, it can vary into something fail-safe, something strong enough to falsify the slack premise. There is something important for our investigation about the fact that what counts as trying to do A or basing a belief on a perceptual appearance can vary, depending on other things that one happens to know; but I will argue that it is a mistake to hope and search for a way for what so counts for us to vary *into* methods impervious to their respective slack premises. The trick is to see that the slack premises are false for us as we already are, not as we can someday hope to become.

2.5   MISTAKES AND ACCIDENTS

This is a good place to remark, before returning to the main thread of our argument, that one source of our temptation to accept the slack premise of the Argument from Failure is the seemingly obvious thought that an accident can "always" divert our efforts in such a way that our trying to do A fails to result in, or blossom into, a successful doing of A. Even if we are firmly resolved to see the project to completion and have a whole afternoon to make our ten carbon-copies, something could go wrong: the carbon paper could burst into flames, our fingers could go numb, or so on. Here I shall try to explain why the fact at which we gesture with these anxious remarks should not dissuade us from denying the slack premise.

It is important that we have used the word "accident" in describing this anxiety, for thereby we mark our respect for a distinction, suggested by J.L. Austin, between accidents and mistakes; and it will serve us well to examine this distinction more closely. What Austin said was this:

> You have a donkey, so have I, and they graze in the same field. The day comes when I conceive a dislike for mine. I go to shoot it, draw a bead on it, fire: the brute falls in its tracks. I inspect the victim, and find to my horror that it is *your* donkey. I appear on your doorstep with the remains and say—what? 'I say, old sport, I'm awfully sorry, &c., I've shot your donkey *by accident*'? or '*by mistake*'? Then again, I go to shoot my donkey as before, draw a bead on it, fire—but as I do so the beasts move, and to my horror yours falls. Again the scene on the doorstep—what do I say? 'By mistake'? Or 'by accident'?[23]

Obviously we are supposed to regard the first case as a mistake and the second as an accident. Austin coyly refrains from offering any explicit account of the distinction between mistakes and accidents, so we are left to our own devices. I therefore suggest the following way of conceiving the distinction.

---

[23] J.L. Austin, "A Plea for Excuses," p. 183n1.

Every agent has various ideas—let us call them "beliefs" for now—about how to do a great many of the things she might see fit to do. These beliefs together constitute the (instrumental) standards of practical conduct that she recognizes and accepts, and to which she therefore regards herself as answerable. Then we may say that a mistake is what occurs when an agent's conduct is contrary to her own standards of practical conduct. For when Austin's agent makes her mistake, she is presumed to have generally good beliefs about how to select her own donkey from the available donkeys. But she does not act in accord with these beliefs when she draws a bead on the donkey in question. Here the fault is in the trying, as Anscombe might have said. An accident, on the other hand, occurs when, even though the agent conducts herself exactly as her own standards of practical conduct recommend, she nonetheless fails to achieve her end because these standards are inadequate to her situation. The agent's beliefs about how to shoot donkeys do indeed constitute good standards of practical conduct for someone inclined to shoot donkeys, but these standards do not cover her particular case, since they apply only when the donkeys do not make sudden movements. In a mistake, the agent violates her own standards of practical conduct; in an accident, the agent does not violate her own standards, but they are themselves inadequate for her situation. Understood in this way, what motivates the Argument from Failure is the idea that failures are always accidents; accidents are always possible.

Notice, though, that the agent can learn a lesson from her accident—to be on the lookout for sudden movements when she is about to fire—and that by learning this lesson, she can refine the standards of practical conduct that she herself recognizes. When an accident does not render her goal unattainable, this refinement of the agent's standards of practical conduct takes the form of a process of checking and correcting her efforts within a single occasion of trying. But even if an accident does render her goal unattainable, the agent can still revise her standards of practical

44

conduct in some appropriate way for the sake of her future endeavors. This fact will become

crucial as we develop our argument.

2.6   DENYING THE SECOND PREMISE

I will get back to that. But first we should consider the only other option, aside from global

skepticism about our practical capacities, for responding to the Argument from Failure. This

would be to deny its second premise, which asserts that we cannot do A by a certain means if

there is slack between that means and our getting A done. The only motivation we cited above

for advancing it was simply that it is the de-epistemologized analogue we needed of the second

premise of the Argument from Illusion, and surely that is not enough to compel anyone's assent.

Indeed, it is thought by many that the second premise of the Argument from Illusion is its

weakest link,[24] so surely we cannot go wrong by taking seriously an attempt to deny its analogue

in the Argument from Failure.

Let us suppose, then, that the second premise of the Argument from Failure is false.

Then there are at least some occasions on which we can indeed do A by trying to do A even

though, as the slack premise maintains, the cooperation of the element of luck is always required

if A is to get done on those occasions. A successful doing of A would therefore have to be a

composite of two independent factors only one of which is within our power, or (as I shall say,

following vocabulary used by the Stoics)[25] "up to us." The factor that is up to us is our trying to

---

[24] This is the view implicit in externalist accounts of knowledge.
[25] Here is Epictetus (*Handbook*, 1):

do A; the factor that is not up to us is the cooperation of the element of luck on which we must inevitably rely to get A done. If this element of luck cooperates (perhaps it would have to do this by *drawing* a getting done of A out of our trying to do A) then we will have successfully done A—but there is no other, more secure way to do A. The denial of the second premise of the Argument from Failure is therefore equivalent to the doctrine introduced above under the label "bifurcationism," since the denial of this premise entails the existence of a fundamental bifurcation between these two independent factors. [26]

Now, it would be easy to dismiss bifurcationism straightaway by observing that this way of thinking likens successfully doing A to winning a lottery that one enters by trying to do A, and that this does violence to common intuitions about what agents may and may not say about what they are (intentionally) doing in lottery cases. For unless she rigs it, no one can intentionally win a lottery, and this is precisely because an element of luck would be required for her to win it— the best she can actually *do* is purchase a ticket. The only sense in which one can intentionally win a lottery requires the elimination of this element of luck. Indeed, ordinary lottery cases are some of the very special cases in which we would ordinarily grant that the Argument from Failure *does* apply, precisely because these are cases in which the slack premise is clearly true. If the truth of the slack premise is the odd condition that must obtain for the Argument from Failure to kick in, it seems more than likely that the second premise of the Argument from Failure is not its weakest link. So there may very well be good reasons to accept the second premise of the Argument from Failure. But at this point our only grip on these reasons comes in

---

> Some things are up to us and some are not up to us. Our opinions are up to us, and our impulses, desires, aversions—in short, whatever is our own doing. Our bodies are not up to us, nor are our possessions, our reputations, or our public offices, or, that is, whatever is not our own doing.

[26] Bifurcationism is a practical analogue of what McDowell calls the "hybrid view" in his discussion of the Argument from Illusion in "Knowledge and the Internal."

the form of an intuition that denying the second premise has unpalatable consequences for how we should describe cases in which agents participate in lotteries. Even if, as seems unlikely, every philosopher shares it, this intuition at best reveals *that* this strategy for addressing the Argument from Failure is a mistake, but we should not cease our investigation of this strategy until we understand *why* that is so.

2.7 THE DISTINCTION BETWEEN SUCCESS AND FAILURE

We can get closer to this understanding by devoting our attention to what the bifurcationist wants to call "trying to do A." We already have part of this account laid out: trying to do A is a way of conducting ourselves within the realm of what is entirely up to us, and it is insufficient to guarantee a getting done of A. But if that is *all* it is, the question certainly arises why we call a way of conducting ourselves within that realm a trying to do A rather than a trying to do something else, say B. So far, the bifurcationist has provided no answer to that question.

A bifurcationist might try to say that she does not need an answer to that question—that a trying is simply a way of conducting ourselves within the realm of what is up to us that bears no relation to what it is a trying to do other than the connections perhaps cooperatively established by the element of luck—but it can be shown fairly clearly that that would be an error. For if the element of luck draws a getting done of B-and-not-A out of a trying so conceived, then it is plain that there can be no criterion with which to determine whether it was a trying to do A that failed or a trying to do B that succeeded. We may put this point by saying that every trying has a kind of "content," in the special sense that *something* about each trying provides for a distinction, in advance of its issue, between what counts as its success and what counts as its failure; and the

47

bifurcationist can only refuse to account for this "contentfulness" of tryings if she is prepared to say that the distinction between success and failure does not really apply to them. But if she does say that, then she will have undermined her own aspiration to describe what counts as a successful doing of A as a combination of a trying to do A with the cooperation of the element of luck, for she will have undermined the distinction between luck's cooperating and its disrupting our efforts. So the bifurcationist cannot shrug off our demand for an account of the contentfulness of tryings.

Now, in addressing this demand, a bifurcationist must take care to account for the "modest claim" we put in place earlier, that an agent is trying to do A only if she knows that she is trying to do A. If a bifurcationist account of the contentfulness of tryings leaves open the possibility that an agent could be trying to do A without knowing that she is trying to do A, then it must be rejected. And what I claim is that no account of the contentfulness of tryings available to a bifurcationist can avoid this consequence. If it is not up to us to guarantee a successful getting done of A, it is not possible for us to know that a certain way of conducting ourselves within the realm of what is up to us is a way to try to do A.

2.8   SUFFICIENT LIKELIHOOD AND KNOWLEDGE THEREOF

A bifurcationist can try to respond to this charge by maintaining that, while it is not up to us to guarantee a successful getting done of A, we can know that certain ways of conducting ourselves within the realm of what is entirely up to us are *more likely* than others to result in a successful getting done of A. Trying to do A, then, is conducting oneself within this realm in a way that we know to be (let us say) "sufficiently likely" to result in a successful getting done of A. Since the

likely issue of a way of conducting oneself can differ from its actual issue, this position seems as if it can provide for the distinction between the success and the failure of tryings; and since it builds in a requirement that one know the sufficiently likely results of a way of conducting oneself that counts as a trying, this position seems to make sufficient provision for our "modest claim."

There are various ways to raise difficulties for this line of thought. We might wonder, for example, what to say about a bit of conduct within the realm of what is up to us that we know to be sufficiently likely to have more than one result, A and B. Is this bit of conduct necessarily a trying both to do A and also to do B, or only of the one (if there is only one) that we want to get done? But I will leave this and many other problems to one side in order to focus on what I think is really the central question, namely: *how* does the agent know that the way she is conducting herself is sufficiently likely to result in a getting done of A?

This might not look like a hard question, but I think that it is impossible for a bifurcationist to give it a satisfactory answer. For it is only possible to *knowledgeably* assess the likelihood of success of ways of trying to do A by comparing one's efforts with what one *knows* about their results. So in order to answer our question, the bifurcationist will have to grant that there is some way—by sight, by sound, by smell, or something like that—in which an agent can at least sometimes know whether some bit of her conduct in which she was trying to do A successfully resulted in A's getting done. But to grant this is in effect to secure the possibility that doing A can fall entirely within the realm of what is up to us, and thereby to give up the bifurcationist game. For an agent who can knowledgeably check whether her trying to do A has successfully resulted in A's getting done can check the results and correct how she is conducting herself within that realm until A gets done; and that self-critical process of conducting oneself in

a certain way, checking the results, and correcting one's conduct is itself a way of trying to do A that can be sufficient to guarantee success.

Consider, for example, Davidson's carbon-copyist, who aspires to make ten carbon-copies but who does not know whether he is succeeding—"all he knows is that he is trying." Much depends here on what we are to count as the trying, for many things can count. If for some reason this agent only has one chance to make as many carbon-copies as possible, then the agent's way of trying will have to be *writing as hard as possible once*. But if he has a whole afternoon to make his ten copies, and if he has a way of knowledgeably checking the success of his efforts, then he *also* has the following available to him as a way to try to make ten carbon-copies: *writing as hard as possible once, knowledgeably checking how many copies have been made, removing the completed sheets, writing as hard as possible again, and so on until ten carbon-copies have been made*. Anyone who can try in *this* way to make ten carbon-copies can thereby, at least sometimes, guarantee the success of his efforts.[27] And in general anyone with a way of knowledgeably checking the success of a way of trying to do A can, at least sometimes, thereby try in a way that guarantees the getting done of A.

It may be thought that I have pulled a fast one in this argument by envisioning, under the label of "knowledge," a way in which the agent can check the results of her efforts that is immune to the difficulties raised by the Argument from Illusion. For if she can see, hear or smell how successful she has been in getting A done, then it is plain that she is not limited in these various realms, as the Argument from Illusion would limit her, merely to conclusions about how things visually, audibly or odoriferously appear to her. The solution of the Argument from

---

[27] These points about Davidson's carbon-copyist are inspired by a lecture entitled "Knowing What's Up: The Unity of its Empirical and Practical Forms" delivered by Michael Thompson in the spring of 2009 at a conference in honor of Anscombe.

Failure by covert stipulation of a solution to the Argument from Illusion will obviously win no converts. And this would indeed be a damning objection to the argument I have provided if an analogue of it did not apply with equal force to the visual, audible and odoriferous versions of the Argument from Illusion. For the Argument from Illusion is an argument *from illusion*, which means that no response to it that undermines the distinction between illusory and veridical appearances can be a stable position. Cartesian accommodation does not work, and so the only response to the Argument from Illusion available to anyone who cannot stomach skepticism and who prefers to retain the slack premise is to deny its second premise. To deny its second premise is to assert that the best thing there can be to call "knowing that p on the basis of (say) an odoriferous appearance that p" will be a composite of two independent ingredients—an odoriferous appearance that p and the cooperation of the element of luck that makes it the case that p—only the first of which we can "really" know. But here we may ask, as we asked in our analogous discussion of tryings, how we can know that a certain odoriferous appearance has the content p rather than some other content (say) q. The answer to this question must be that we have ways of "really" knowledgeably comparing occurrences of this sort of appearance with how things really are in order "really" knowledgeably to assess the reliability of this sort of appearance in representing various kinds of facts. And that requires us, at least sometimes, to have "real" knowledge-constituting access to the facts represented by appearances, access that is not subject to the whims of the element of luck. So I have not covertly stipulated a solution to the Argument from Illusion in my treatment of the Argument from Failure—it would be more accurate to say that I am offering a solution that works equally well for both Arguments.[28]

---

[28] I must emphasize that the solution I am offering is not mine, for this paragraph summarizes what I take to be McDowell's treatment of the Argument from Illusion in "Knowledge and the Internal," an article to which this chapter is pervasively indebted.

It may also be thought that I have pulled a fast one in my argument against bifurcationism by considering a case in which what counts as trying is or includes a way of conducting oneself—*writing as hard as possible*—that a bifurcationist would not concede to fall within the realm of what is up to us. But the argument does not turn on this point, since it will apply with equal force to any way of conducting oneself that the bifurcationist would concede to fall within that realm. As long as an agent can knowledgeably check the success of a way of trying to do A (whatever realm it may inhabit) and adjust her conduct appropriately, it will be possible for her to guarantee, at least sometimes, the getting done of A.

I urge the reader not to be misled, as I suspect Davidson and many other bifurcationists are misled, by the fact that the variable A can take certain values that falsify the antecedent of this conditional. This is the case when the variable A takes as its value what Davidson seems to have envisioned—*making ten carbon-copies in a single attempt*. In *that* case, we are to suppose that the agent can do no better than try. But that is because the requirement that the ten carbon-copies be made in a single attempt does not allow the agent (time) to adjust her conduct appropriately. Some values for the variable A do allow the agent to adjust her conduct appropriately, however, and that is what matters.

The values that the variable A can take, for a given agent, will fall into three classes—things that she can do, things that she can only try to do, and things that she cannot do or try to do at all. Among other things, bifurcationism is the claim that the second of these classes is prior to the first, since a successful doing of something is just a trying to do something that receives the cooperation of the element of luck. What I have argued is that if there is such a thing as trying to do something, and if one always knows what one is trying to do, then the first of these classes must have members that are not properly understood as mere combinations of a trying

with the cooperation of the element of luck. And that is a way of saying that a trying to do A does not always require the cooperation of the element of luck to get A done. In other words, it is a way of denying the slack premise, which is the strategy we should really adopt in response to the Argument from Failure.

2.9   MISTAKES AND ACCIDENTS AGAIN

I take the considerations I have advanced to demonstrate that, contrary to appearances, an accident cannot "always" divert our efforts in such a way that our trying to do A fails to result in, or blossom into, a successful doing of A. That will depend on whether we happen to be in a situation for which our standards of practical conduct are adequate. If we are in such a situation, then an accident is impossible. What is possible in those situations is that we will make a mistake—but it is entirely *up to us* to guarantee that a mistake does not occur. So the slack premise is false, since there are occasions on which we do not require the cooperation of the element of luck for a trying to do A to result in a getting done of A. All that is required on these occasions is our own diligence in acting in accord with the standards of practical conduct that we know to be adequate to our situation—trying in the way we know we should.

I just used the word "know" because I think it should be clear by now that these standards constitute a body of know-how, rather than merely a body of what we might awkwardly have to call "believe-how." For one always knows what one is trying to do, and if our treatment of the Argument from Failure is correct, then it is impossible to know that much if one does not *know* that conducting oneself in a certain way is what we called "sufficiently likely" to result in (say) A's getting done. And we have seen that the conditions for the possibility of *this* knowledge

require us to know, for a range of cases, how we can conduct ourselves in a way that is sufficient to guarantee that A gets done. So this know-how is a kind of expertise with a certain *generality*—it can be applied again and again on different particular occasions. What goes wrong in a mistake is that the agent does not apply her know-how properly on a particular occasion; what goes wrong in an accident is that she enters a situation outside the range of cases about which this expertise counts as knowledge.

There are parallels in the various realms of appearance. We have a certain expertise that counts as knowledge, for a wide range of cases, about which sorts of visual appearances are appearances of (say) bent sticks. We can apply this expertise lazily, and get a false conclusion on an occasion within that range of cases—that is an epistemic *mistake*. But we can also apply this expertise with all the diligence it might require on an occasion when the stick is partially immersed in water, a case (say) outside the range of cases for which this expertise might count for one of us as knowledge—that is an epistemic *accident*. (Illusions are epistemic accidents.) And while it is true that one can be in a situation outside this range of cases without knowing that one is outside of it, that just does not imply that one can "always" suffer from an illusion or that one "never" knows whether one is in a situation within this range of cases. When one is in a situation within the range of cases for which this expertise about sticky appearances counts as knowledge, one knows how to tell whether the stick is bent; when one is not, one doesn't, although one might very well think one does. So we should not think that one can "always" suffer from an accident or an illusion. On the contrary, if there can be such a thing as an accident or an illusion at all, then there must be such a thing as occasions on which we are immune to them.

I have just in effect endorsed a doctrine that often goes by the name "disjunctivism"—the claim that, when things appear to us to be a certain way in accord with our standards for grasping the significance of appearances, then *either* that appearance is an illusion *or else* it really is a sufficient basis for a bit of knowledge that things are that way. If this chapter had been dedicated to an epistemic version of the Argument from Failure, its punchline would have been disjunctivism about the potential for our trying to do A to provide a sufficient basis for a bit of knowledge that we getting A done. Since we have concerned ourselves instead with an efficacious version of the Argument from Failure, our punchline is efficacious, rather than epistemic, disjunctivism: when we try to do A in accord with our own practical standards, then *either* that trying will meet with an accident, *or else* it really is sufficient to guarantee that A gets done.

2.10   THE THING DONE

I take this clarification of the distinction between mistakes and accidents to dispel one source of the temptation to accept the slack premise. But there is at least one other source of temptation that I should mention before closing this discussion of the Argument from Failure. For rejecting the slack premise requires us to see a sort of unity between our attempts to do A and our successful doings of A—trying to do A in accord with our own practical standards can be sufficient to guarantee that A gets done—and the very idea of such a unity can appear mysterious. So it is necessary to grasp the concept of something an agent can do in a way that dispels this sense of mystery. I believe that this sense of mystery arises from the fact that there are at least two demonstrably different senses in which, in ordinary language, we say that

someone is doing something, and that philosophers of action have tended to investigate one of these senses at the expense of neglecting, even forgetting, the other. For we often say that an agent is doing more than one thing at the same time, or in other words that she is *multitasking*, while we also often say that an agent is doing one thing *in order to* do another, and in these two sorts of cases, we use the word "thing" in two different ways. This is clear from the fact that there are many situations in which an agent does one thing in order to do another, doing them both at the same time (as when she moves her finger in order to flip a switch), but in which we would not say that she is multitasking. So we are entitled to conclude that someone is multitasking only when we know she is doing more than one thing at the same time in a sense of the word "thing" that is distinct from the sense of the word "thing" that is usually operative when we say that she is doing one thing in order to do another.

Lack of clarity about this distinction tempts us to accept the slack premise in the following way. Say that an agent is trying to do A in order to do B. Then it looks as if there are two things she is doing or at least trying to do—A and B. It looks this way even though it is plainly written on the surface of the given fact that the agent is also doing or at least trying to do something else—A-in-order-to-do-B. But this third thing, which is the task the agent is performing, is frequently forgotten. Now let us suppose that, with respect to A, the agent is not in circumstances in which she can suffer from an accident, but that with respect to A-in-order-to-do-B, she is in circumstances in which she can suffer from an accident. Then the agent's trying to do A in accord with her practical standards will be sufficient for her to get A done, but it will not be sufficient to get A-in-order-to-do-B done. And this is so even though, importantly, there is a sense in which the agent's trying to do A *is* her trying to do A-in-order-to-do-B. This configuration of possibilities leads us to say that *trying* in accord with our practical standards,

even in situations where the trying cannot itself suffer from an accident (since the agent cannot fail to do A if she tries in accord with her practical standards), is insufficient to guarantee the getting done of what it is a trying to do. But this conclusion turns on an equivocation on the phrase "what it is a trying to do." Almost every trying is a trying to do many things, and in many cases, it is only in an accident-free range with respect to some of those things. But the general principle does stand: if an agent is in a situation within the range free from the possibility of accident with respect to a certain thing she can do, and she tries to do *that very thing* in accord with her own practical standards, then her trying to do that very thing will be sufficient to get it done.

In order to insure ourselves against this source of temptation, I therefore propose to subject this distinction between the two different sorts of thing that an agent can do to further philosophical scrutiny.

3.0   MULTITASKING

3.1   FIRST SECTION

Very roughly speaking, a Copernican revolution is a revolution in our thinking in which we come to understand something better by understanding our experience of it as the *work* of our own cognitive faculties, rather than as something that the world independent of us somehow does all the work of *putting* into our minds or else as something somehow *innate* to our minds that we inherit like a fortune without doing any work at all. Very roughly speaking, the first of these alternative pictures is the favored image of "pre-critical" empiricism, and the second is the favored image of "pre-critical" rationalism.

I have begun this chapter with these very rough characterizations because I think that much current ethical theory derives substantial motivation from what one might call a pre-critically *empiricist* conception of what it is for something to be a practical option, and is therefore in need of a Copernican revolution with respect to this concept. This is a form of empiricism according to which practical options are *given* to practical reason—*then*, and *only* then, does practical reason begin to be active when we try to choose one of them. With this notion of pre-critical empiricism about practical options I am supplying what I hope is a helpful label for a view that, for example, Cora Diamond diagnoses in William Frankena's reading of the *Crito*:

Socrates says that his escaping would be breaking an agreement. If that is a premise in the argument, and it is not a moral principle, it must be a statement of fact—so *that* cannot be where any moral thinking is. That is how Frankena sees the case. And this is despite the fact that the moral originality of the description of facts is underlined by Plato. When Socrates asks Crito whether, if he escapes without persuading the State to let him go, he would be treating someone badly, and not *just* treating *someone* badly but treating badly those whom least of all he should, when Socrates follows that with the question whether he must stand by his agreements or no, Crito has no idea how to answer; he does not understand the questions, does not know how to bring the terms of the questions into connection with the case before him. Socrates then by an exercise of moral imagination involving the personification of the Laws enables Crito to see the situation differently. All of which is regarded by Frankena as nothing to do with moral thinking. Facts are facts; describe them, and then comes the moral work: apply your principles.[29]

This way of thinking about moral problems encourages and is encouraged by the widespread use in moral philosophy of kill-one-to-save-five-style thought experiments like the Trolley Problem. For in order to be useful at all a thought experiment must present at least certain features of our practical options as given, as features that we take in with an infinitesimal effort—"you can see the situation at a glance"[30]—so that inquiry into the relevant moral principles can begin.

Is there anyone who occupies the logical space of a pre-critical rationalism to oppose the empiricism with which this concept is widely treated? That would be a rationalism on which practical options are entirely *innate* to us rather than *given* to us. The best example of such a rationalism that I can offer is a conception of our practical options in the State of Nature that sometimes seems to motivate the social contract theory about the legitimacy of the authority of the state. On this conception, there *is* no story to be told about how it comes to pass in the State of Nature that we have various options; that is just the starting point. And this is not an empiricism, because we are not up and running as practical beings before these options—to cooperate with one another in various projects, to make war against one another, and so on—are

---

[29] Diamond, "Missing the Adventure: Reply to Martha Nussbaum," in *The Realistic Spirit*, pp. 310-311. Diamond refers to Frankena's discussion in *Ethics* (Englewood Cliffs, N.J., 1963), pp. 1-3. Notice the affinity almost explicitly marked here by Diamond between empiricism about practical options and a kind of scheme/content dualism.

[30] Judith Jarvis Thomson, "The Trolley Problem," p. 1397, in Yale Law Journal Vol: 94 Issue: 6, pp. 1395-1415.

given to us. Rather, what it is for us to be up and running as practical beings at all just is for us to have these options. So the State of Nature is simply a state in which we *have* certain practical options but not others, and from which, therefore, it is in our interest to make certain sorts of agreements with one another. That is a kind of pre-critical rationalism about practical options. But such a rationalism is just as misleading as its empiricist antipode—although it misleads in a different direction.

In setting up pre-critical empiricism and pre-critical rationalism in this dialectic as positions to be overcome, I do not mean to deny that practical thought can take place, as it were, in an atmosphere in which it is somehow settled what one's practical options are, and in which it is simply our job to choose one. Nor do I mean to deny that there can be value in investigating how we ought to think practically when we find ourselves in such an atmosphere. The point is just to emphasize that when we stipulate that certain practical options are available and others aren't, we stipulate that a certain kind of *work* has been done. For practical thought can take place in other atmospheres of thought as well, atmospheres in which this kind of work has not *yet* been done, and which are therefore in important sense prior to those in which pre-critical empiricism and pre-critical rationalism place us. And our conception of what it is to do the job of living an ethical life goes wrong if we miss this fact. For just as Kant showed (for those who follow him) that we cannot get clear on what it is to experience an object except by recognizing that it requires the active participation of our own intellects in combination with receptivity, so too we must show that it is impossible to get clear on what it is to have a practical option except by recognizing that it also requires the active participation of our intellects in combination with receptivity.

I will try to generate an argument for this assertion by attending to the mere notion of the sort of thing that can be done. For it should be clear that that is all we should need to clarify in order to clarify the notion of an option to do that sort of thing. In particular, if our aim is to pursue a Copernican revolution in our thinking about the concept of a practical option, we should clarify the sort of *unity* characteristic of a thing that can be done—for just as Kant showed (for those who follow him) that we begin to see that experience of an object must involve the participation of our intellects when we begin to see (with help from Hume) that it has a kind of unity that something merely given could not supply, so too we begin to see that having a practical option must involve the participation of our intellects when we begin to see that it has a kind of unity that something merely given could not supply.

We can start to think about this unity by attending to the concept of a thing that can be done that is implicit in the notion of doing *more than one* thing at a time. For if there is some philosophically workable concept of a thing that one can do according to which one can do more than one of those things at a time, then attending to that notion will teach us how some of whatever one is doing at a time is unified into one of them while the rest is unified into the other. And so we will learn about at least one sort of unity characteristic of a thing that can be done, to which what I have called "pre-critical" conceptions of practical options cannot do justice.

I hope it is clear that although its topic is quite different, this project bears a strong resemblance to Frege's project in *The Foundations of Arithmetic*. Frege sought to understand *what a number is*, and he took it that at least one workable criterion of adequacy for such a search would be to arrive at "a means of arriving at a determinate number and of recognizing it again as the same"[31]—that is, to arrive at criteria of identity for numbers sufficient to enable us

---

[31] *Foundations of Arithmetic*, §73.

to say whether it is true of two number-designating expressions *a* and *b* that they designate the same number. As he says, "If we are to use the symbol *a* to signify an object, we must have a criterion for deciding in all cases whether *b* is the same as *a*, even if it is not always in our power to apply this criterion."[32] Just so, we seek to understand *what a task is*, and we take it that at least one workable criterion of adequacy for such a search would be to arrive at criteria of identity for tasks sufficient to enable us to say—as in our everyday talk of multitasking we *do*—whether it is true of two task-designating expressions that they designate the same task. The obligation to find such criteria is not specific to any particular theory of agency, but is quite general, and it is no less pressing than the Fregean scruples that have imposed analogous obligations on philosophers who work on many other topics.

The first philosopher whom I know to have posed the problem of multitasking was Castañeda.[33] The problem, roughly, is that one may have, on the one hand, a reason to do A and also a reason to do B and thus a reason to do A while one also does B, while at the same time, on the other hand, one has a reason *not* to do A-and-B. Bratman discusses this difficulty as it is articulated in the context of Castañeda's theory of practitions in an extended footnote in "Castañeda's Theory of Thought and Action." The punch line of that discussion is as follows:

> The upshot is that the Legitimacy of certain practitions will depend on whether we think of the relevant action as simple, or as a compound. In a later discussion… Castañeda grants this point and replies that this shows that contexts of Legitimacy must include a specification of what actions are to count as "prime actions." Doubtless, this is technically possible. But when we do this which context will be the "absolute context" involved in specifying the designated value of

---

[32] *Foundations of Arithmetic*, §73. I think that Frege should here have said "in at least a broad range of cases" rather than "in all cases," for otherwise it is not clear how to sidestep the "Julius Caesar problem." This is the problem that we confront when we are asked how we may be in a position to know that the number seven is not Julius Caesar.
[33] See, for example, Castañeda, H. N., "The Theory of Questions, Epistemic Powers, and the Indexical Theory of Knowledge," in P. French et al., eds., *Midwest Studies in Philosophy V* (Minneapolis: University of Minnesota Press, 1980): 193-237.

practitions? Whichever choice we make may seem to have a kind of arbitrariness which threatens Castañeda's claim to have articulated the "psychologically preferred value" of practitions. [34]

In other words, Bratman accuses Castañeda of lacking a satisfactory criterion on the basis of which to determine whether an agent is single-tasking or multitasking in doing A and also B, so that Castañeda leaves every agent in a position to gerrymander descriptions of her actions so as to disavow responsibility for at least a great deal in them that is not to her credit. [35] But it is not just Castañeda's idiosyncratic theory of practitions that confronts this difficulty—the provision of such a criterion is obligatory for *any* account of agency. So while I cannot claim to be the first philosopher to notice that there is something here to philosophize about, nonetheless I think it is a problem to which insufficient philosophical attention has been given. And if the account of the concept of multitasking that I develop in this chapter is correct, then the cogency of this concept can serve as the premise of a novel refutation of an important source of motivation for consequentialism and of intentionalism that I will try to sketch in its closing sections.

3.2   SECOND SECTION

There is a joke currently in circulation about some sort of guru who teaches his students that they should only ever do one thing at a time. "Whenever you are drinking coffee," he instructs them, "*only* drink coffee, and whenever you are reading the newspaper, *only* read the newspaper." But

---

[34] "Castañeda's Theory of Thought and Action," p. 234n10, in *Faces of Intention* (New York: Cambridge University Press, 1999).

[35] Bratman also discusses various issues relating to the topic of multitasking under the label of "intention agglomeration" in "Acting with an Intention," pp.134-138, in *Intention, Plans and Practical Reason* (Cambridge: Harvard University Press, 1987).

one morning a group of his students finds him in his garden reading the newspaper and drinking coffee at the same time.

"Master," they say, "you told us *only* to drink coffee whenever we drink coffee, and *only* to read the newspaper whenever we read the newspaper, but here we find you drinking coffee and reading the newspaper at the same time. Please explain your behavior."

"Well," he responds, "whenever you are drinking coffee, you should *only* drink coffee, and whenever you are reading the newspaper, you should *only* read the newspaper. But neither of these principles applies to me right now, since what I am presently doing is drinking coffee *and* reading the newspaper. And I say that when you are drinking coffee *and* reading the newspaper, you should *only* drink coffee and read the newspaper."

This joke is supposed to be funny thanks to its play on the concept of doing more than one thing at a time, or of multitasking, as it has recently come to be called. The guru's justification for what he does is supposed to be laughable because, if it were satisfactory, then the concept of multitasking would capture nothing. It would capture nothing because anyone accused of multitasking by doing A and doing B at the same time could respond as the guru does, by saying that in fact he is only doing the one thing A-and-B. So if the guru's justification for what he does were satisfactory, then the guru's doctrine that one should not do more than one thing at a time would be an empty doctrine, since there would be no way to act that it ruled out.

This points to a suggestive philosophical lesson, namely, that we must be cautious in our attribution of conjunctive doings to agents, for it is not always correct to infer from the fact that an agent is doing A and doing B at the same time to the conclusion that she is doing A-and-B at that time. It would be more modest to say that we can endorse this conjunctive pattern of inference generally only at the cost of making the ordinary concept of multitasking unintelligible.

But it would be quite embarrassing to have to renounce the ordinary concept of multitasking, since we (take ourselves to) multitask whenever we simultaneously drink and drive, walk and chew gum, or speak on the phone and scratch an itch. This ubiquity of the phenomenon alone provides sufficient motivation for us to prefer accounts of agency that allow for its possibility, and that therefore reject this conjunctive pattern of inference. But there is also reason to hope that a correct understanding of the concept of multitasking will enable us to determine what sorts of items are fit to be substituted for the variables A and B, since one can tell whether someone is multitasking only insofar as one can count the things she does, and since one can count a sort of item only insofar as one knows what sort of item it is.

To be clear, it is not my ambition here to offer another refutation of (what I call) consequentialism or intentionalism as such, for there are possible versions of those views that do not rely upon or derive motivation from the empiricist conception of practical options that is my target here. But I believe that consequentialism or intentionalism are inevitable for anyone who does have an empiricist conception of practical options, so that the prevalence of this conception accounts for the appearance of inevitability that these positions enjoy. To the degree to which this is so, a refutation of the empiricist conception of practical options should help to dispel that appearance.

My plan in this chapter is to develop a minimally committal account of what multitasking has to be in order for our ordinary speech about it to make sense, and then to defend against various objections the claim that multitasking so construed is possible. In defending multitasking against these objections, we will repeatedly find opposition in the guru's idea that there is something *wrong* with multitasking—that multitasking can never be practically justified. By addressing this source of opposition, therefore, we will have established by the end of this

chapter not only that multitasking can take place, but also *that* it can be practically justified, or in other words that an agent can be practically justified in multitasking. *How* multitasking can be practically justified, however, is a topic we shall address in the next chapter.

3.3   THIRD SECTION

Speaking loosely, multitasking is doing more than one thing at the same time. So it might seem as if we can reconstruct the concept of multitasking from Anscombe's celebrated distinction between an intentional action and the descriptions of it under which it is intentional, a distinction subsequently appropriated and relabeled by Davidson as the distinction between "doings" and "things done." In the example he gives to illustrate the distinction, Davidson moves his finger, flips the switch, turns on the lamp, and illuminates the room.[36] Here we are supposed to see that four things are done (the finger moves, the switch is flipped, and so on) even though only one doing occurs. The doing is some sort of particular; the things done in it are (perhaps some special subset of) the many properties that it has, action-concepts it instantiates, event-kinds it brings about, propositions it makes true, or something like that. This is a fine distinction; and using the terms "doing" and "thing done" in the way Davidson suggests would give us something, at least, to mean by saying that someone is doing more than one thing at the same time. But it should be clear that this is not what we mean when we say that someone is multitasking. For since every particular doing will have many properties, instantiate many

---

[36] See Anscombe, *Intention*, §22, and Davidson (1963), p. 4. In his original case, Davidson also *un*intentionally alerts the prowler—a complication we shall ignore here. Henceforth, whenever we speak of something's being "done" or "performed" by an agent, the reader should understand us to mean "*intentionally* done" or "*intentionally* performed."

action-concepts, bring about many kinds of events, make many propositions true, or something like that, an agent cannot act at all without doing more than one thing at the same time, according to Davidsonian usage. But an agent can act without doing more than one thing at the same time, in the sense that would make her count as multitasking. Since agents sometimes act without multitasking, the sense in which one does more than one thing at the same time *only when one multitasks* is not the sense in which, as Davidsonian usage would entail, one *always* does many things in any doing.

So it is clear that Davidsonian usage for the term "things done" can lead us into a great deal of confusion—even though the distinction between doings and things done itself is a fine one—since our goal is to clarify a sense of the phrase "S does more than one thing at the same time" that is different from the sense that Davidsonian usage would give to it. We could avoid this confusion by reverting to the idiom of Anscombe in which the distinguished items are referred to as intentional actions and the descriptions of those actions under which they are intentional, but this is a bit of a mouthful. We shall therefore adopt the slightly awkward convention of using the term "doable" to refer to what Davidsonian usage would call a "thing" that can be done and of using the term "task" to refer to the kind of thing such that an agent multitasks when she is the agent of more than one of them at the same time. To be the agent of a doable, according to this convention, is to "do" it, and to be the agent of a task is to "perform" it. And for the sake of brevity, we shall sometimes refer to an agent's doing of a doable A as her "A-ing."

Now, since the object of our investigation is the performance of tasks, we shall presuppose nothing about what sorts of items tasks are, or about what it is for someone to perform them. But we shall also presuppose nothing about what sorts of items doables are, or

67

about what it is for someone to do them. This is a topic about which there has been rich and heated debate. But we give the reader permission, while reading our argument, to project onto the term "doable" her own view about whether doables are action-concepts, propositions, states-of-affairs, event-kinds or perhaps something else, and onto the terms "do," "does" and "doing" her own view about whether what it is for an agent to do a doable is for her to instantiate it, to make it true, to constitute it, to bring it about, or perhaps something else. No doubt the reader performs much the same feat whenever she discerns her own views about these topics lurking under the surface grammar of ordinary speech. But if this projection becomes impossible for the reader in the light of whatever we discover about tasks, then we will have refuted her view by showing, in effect, that it cannot plausibly be discerned under the surface grammar of a core aspect of ordinary speech.

The fact we aspire to explain is simply that one does some doables (whatever those are) if and only if one performs at least one task (whatever that is). For an explanation of this fact will help us, by illuminating the connection between doing doables and performing tasks, to understand the concept of a task, and thus the concept of multitasking. We can begin our account by pressing a suggestive intuition about how Davidson's case differs from that of the guru.

3.4   FOURTH SECTION

It would be natural for a proponent of the concept of multitasking to reach for language more or less like this: "The reason why Davidson does not multitask when he does so many doables (moving his finger, flipping the switch, and so on) is that these doables all *have something to do*

with one another in Davidson's doing, whereas the guru multitasks when he does his two doables (drinking coffee and reading the newspaper) because these doables *have nothing to do with one another* in his doing." The idea here is that there is some *relation* that obtains among Davidson's doables but that does not obtain among the guru's doables, a relation that somehow glues all of Davidson's doables together into a single task to which they all belong. And we can give an initial characterization of the relation in question by saying that two doables A and B *have something to do with one another* in the agent's doing, in the relevant sense, if there is some doable the doing of which her A-ing and her B-ing both promote. This can happen if one of them promotes the other, or if her A-ing and her B-ing both promote her doing of some third doable C.

Those who regard doables as event-kinds will no doubt be inclined to propose a causal analysis of this "promotion" relation. On this analysis, two doables A and B have something to do with another in an agent's doing, in the sense that entails that A belongs to at least one task the agent then performs that also includes B, if her A-ing causes her B-ing, if her B-ing causes her A-ing, or if her A-ing and her B-ing are both causally upstream of her doing of some third doable C. Perhaps we would have to add, in order to head off problems about deviant causal chains, that causal connections among doings can only glue them together into performances of the same task if the causal connections are of some special kind that it is the obligation of causal theorists to worry about, explain or explain away. However that may be, we can sum up this proposal with the slogan that the tasks an agent performs correspond to the termini of whatever complete causal chains exist among the doables she does. For if her B-ing causes her to do both C and D without its coming to pass that either of these doings causes the other—that is, if the causal consequences of her B-ing *branch*—then B (and, assuming that the relevant kind of causal

connection is transitive, every doable a doing of which causes B in the right way) belongs to a task the agent performs that includes C, and also to a task the agent performs that includes D, even though C does not belong to a task the agent performs that also includes D. It follows that a doable B belongs to more than one task performed by an agent whenever her doing of it has branching causal consequences that do not rejoin among the agent's other doings, for then the agent performs *distinct* tasks by doing this one doable B. Complete causal chains, like highways, can in this way overlap and then split apart, and that is how we multitask.

It is remarkably easy to generalize this view into a form fit for adaptation into other accounts of the nature of doables, an issue concerning which we ourselves are officially agnostic here. For it is pretty universally agreed, even among those who disagree about what sorts of items doables are, that an agent can sometimes do one doable *in order to* do another. Different theories of the nature of doings and doables will perhaps be inclined to different theories of the conditions under which this is possible, as for example the event-kind theory of doables inclines its proponents to the view that an agent can only do one doable in order to do another in circumstances in which her doing of the first would *cause* her doing of the second. We cannot resist remarking that there seems no more compelling reason to require that doables only ever cause one another than there is to require that doables only ever constitute one another, or make one another true, or anything else—surely agents do their doables by whatever means are available. But that is not an issue we aspire to settle here; we merely use the term "promote" as a kind of dummy term on a par with "doing" and "doable," so that the reader is free, as with those terms, to project onto the expression "A-ing would promote B-ing" whatever view she takes of the relation between the doables A and B that makes it possible for an agent to do A in order to do B. (Notice that we can therefore remain agnostic about whether a doing of A that promotes a

doing of B *is* this doing of B that it promotes or is merely some kind of condition for its occurrence.) This use of the term "promote" enables us to generalize the causal theorists' proposal with the slogan that the tasks an agent performs correspond to the termini of whatever complete promotional chains exist among the doables she does. Since these termini are doables that the agent does, but not in order to do any further doables, we shall call them her "ends." Thus we can explain the fact we set out to explain: that one does some doables (whatever those are) if and only if one performs at least one task (whatever that is). And it follows from this account that an agent multitasks when what she does has more than one end.

3.5   FIFTH SECTION

Intuitively, then, one multitasks when what one does has more than one end, since then these ends are doables that have "nothing to do with one another" in one's doing, which we have analyzed in terms of the absence of a relation that the proposal under consideration aspires to explain: the two doables have "nothing to do with one another" in the relevant sense if and only if neither of them promotes the other and there is no third doable that they both promote. But even doables that have "nothing to do with one another" in this special sense must nonetheless have *something* to do with one another if they are to figure in a genuine case of multitasking. For it is manifest that no one would have been guilty of multitasking in the guru's garden if, say, the guru was merely drinking coffee while at the same time his neighbor, another agent, was reading the newspaper by the guru's side. For in this case, even though two different tasks are simultaneously performed in the garden, they are performed by two different agents, so that no one is multitasking in such a case since there is no one who is performing more than one task at

the same time. Multitasking only occurs when a *single* agent performs multiply many tasks. (We may put this point somewhat suggestively by saying that it must be possible for the "I do" to accompany all of my doables.) And this requirement that the various activities that constitute a case of multitasking have a single agent excludes all cases in which the doing of A has its source in the agency of one agent while the doing of B has its source in the agency of another, no matter how intimately these two agents rub shoulders. And we can picture a constellation of doings occurring on a particular occasion as the doings of a *single* agent only if this constellation has the structure of a tree, each of whose promotional branches stems from a common promotional trunk, which is the agent herself.[37]

Now, it would be a stretch of our dummy vocabulary to say that an agent "promotes" any of her doables, since promotion is a way in which doables can be related to one another, and an agent is not herself a doable. So the relation between the agent-trunk of the tree we have envisioned and its branches cannot simply be assimilated to the promotion relation. We could perhaps call it the "source" relation. And here again we give the reader permission to interpret this term in whatever way she finds agreeable, as long as she understands that no account of agency can be adequate if it cannot countenance and explicate facts to the effect that *this* agent, rather than someone else, is the promotional source of a particular doable A—for there must be such facts if multitasking is to be possible.

---

[37] Many philosophers believe that multiply many agents can sometimes join together to constitute a single agent. This view is held in common, for example, by both Aristotelian and Hobbesian theories of the nature of the state: the state is itself a kind of agent, constituted in various ways by its citizens or subjects, that engages in all sorts of different activities at the same time. If *that* is what the guru and his neighbor do, and if it is *as* this single agent that they both constitute that each of them conducts himself in the way he does, then perhaps this single agent multitasks through their various actions. We do not here mean to deny (or affirm) that multiple agents can synthesize themselves into a single agent; we only mean to say that it is possible for it *not* to happen when two agents act in close proximity.

3.6   SIXTH SECTION

If multitasking is possible, then, this admittedly modest and trivial result allows us to infer from the fact that an agent is doing A and doing B to a conclusion about how things stand promotionally "*downtree*" (upstream)[38] of these two doings: in such a situation, these doables must have this single agent as their common promotional source. We shall call this the "Common Trunk Inference." But this is not how the guru wanted to infer. For the guru wants to justify himself to his students by inferring—from two distinct facts, the first that he is doing the doable *drinking coffee* and the second that he is doing the doable *reading the newspaper*—that he is only performing one task, or that all of his doables have a common end, in the doable *drinking coffee and reading the newspaper*. So the conclusion he wants to draw is about how things stand promotionally "*uptree*" (downstream) of his two doables, about the *end* or *ends* to which these doables are related in his doing as means. The guru therefore subscribes to what we might call the "Common Branch Inference," which moves from the fact that an agent is doing A and doing B in the same doing to the conclusion that both A and B have a common end. This formulation of the "Common Branch Inference" is a more exact statement of the claim that no one ever multitasks, since no one ever multitasks if and only if each member of every pair of doables an agent does has an end in common with the other. This occurs whenever one of A and B promotes the other, or if they both promote a third, for then both A and B have an end in common.[39] So far we have found no basis for such an inference.

---

[38] I have been told that the Ancient Egyptian word for "south" was homonymous with the word for "upstream," with the result that early Egyptian explorers of Mesopotamia found themselves awkwardly obliged to send home mad-sounding reports that they had found rivers that flowed upstream. I here coin the words "downtree" and "uptree" in order to spare the reader a similarly disorienting mixture of metaphors.

[39] Simply in order to avoid excessive verbiage, I stipulate that an end *has itself* as an end. This allows me to use the concise phrase "A and B have an end in common" for what the Common Branch Inference would allow us to infer.

It may perhaps be thought that the possibility of multitasking is ruled out by an Aristotelian conception of rational agency, since Aristotle sometimes seems to have held that there is only one end (*eudaemonia*) that any action ever promotes, or which any action ever seeks.[40] Whether or not that is plausible as a reading of Aristotle, we will sidestep the issues produced by theories of this form, according to which there is some unique and special task that everyone is always performing in any doing of any doable, by stipulating that in the context of such a theory "multitasking" refers to situations in which the tree that exhibits the promotional structure of one's doables *would have* multiply many ends if this unique and special task were removed from it. For even if Aristotle (on this reading) were right, it seems as if our ordinary notion of multitasking would continue to serve its function by singling out cases in which one's conduct conformed to this pattern.

Now, it is worth remarking that both the Common Trunk Inference and the Common Branch Inference are inferences about how agents *actually* do doables and perform tasks, rather than about how they *ought* to do doables and perform tasks. We have found a basis for the Common Trunk Inference in our account of what it is for an agent to do or perform anything at all—it is simply impossible for a single agent to be doing A and to be doing B unless *she* is the promotional source of these two doables, for otherwise she would not be a single agent. It might be thought, however, that the Common Branch Inference is justified by considerations about what it is to do doables and perform tasks *well*, rather than exclusively by considerations about what it is to do or perform anything at all. According to this line of defense, the Common Branch Inference is justified because it is an excellent pattern of practical reasoning for an agent,

Otherwise I would always have to render the upshot of this Inference more or less as that "either one of them promotes the other or else they both promote a third," which is unpleasantly awkward.

[40] This claim is suggested by the first sentence of the *Nicomachean Ethics*, for example.

whenever she wants to do two distinct doables A and B, to ensure that they both have a common end, either by promotionally subordinating one of them to the other, or by promotionally subordinating both of them to a third doable C, which in limiting cases such as the guru's would simply be the conjunctive doable A-and-B. (Notice that this norm of practical reasoning is the very doctrine that the guru propounds to his students.)

We have three objections to this move, which we might call the "appeal to excellence." First, as is well known, "ought" does not imply "is." The fact, even granting that it is one, that whenever one wants to do A and also wants to do B, one *ought* (in some substantive sense in need of clarification) to promotionally subordinate them both to a common end, does not entail that one *does* structure what one does in that way whenever one wants to do A and also wants to do B. One would also have to add the bold metaphysical claim that the observance of this particular "ought" is of such crucial importance for the constitution of an agent that no agent can do anything at all without observing it. We have hopefully made this bold metaphysical claim look implausible, or at least in need of defense.

More fundamentally, second, we simply should not grant that one ought (in some substantive sense in need of clarification) to ensure that both of the doables A and B have a common end whenever one wants to do A and also wants to do B. Indeed, observance of such a norm of practical reasoning would be madness, if only because not all of the things one wants are compatible with one another. There are two ways that a defender of the Common Branch Inference might try to dispel that problem. First, she might maintain that this norm is but one of the many norms of practical reasoning, and that it can be defeated in some circumstances, as for example in circumstances in which two of the things one wants are incompatible with one another. But if she says this, then she will have undercut the universality of this norm that was

supposed to underwrite the Common Branch Inference in the first place. Her second option would be to concede that we do indeed have many idle wishes that are incompatible with one another, but that it is not just any kind of "wanting" to do A and to do B that obliges an agent to ensure that they both have a common end, but "wanting of the will," so to speak—or in other words, positive doings by the agent of A and of B. The principle to which she would appeal, then, is that one ought (in some substantive sense in need of clarification) to ensure that both A and B have a common end whenever one is doing A and also is doing B. It is not clear whether this solves the problem of incompatibility, since it certainly seems possible for an agent to be fully engaged in the prosecution of incompatible doables at the same time, and to realize the incompatibility only after her efforts frustrate one another. One could even say that whatever guarantee of compatibility among an agent's doables there may be results not from the logic or the metaphysics of doing, but from what the agent herself does in harmonizing her various doables, a harmony that she sometimes establishes precisely by ceasing to do one of these doables when it is impossible to create harmony otherwise. But we shall let that pass, and move to our third remark.

Third, and most importantly, the Common Branch Inference can rest on this normative principle only if what it requires is that the agent do a certain doable as an *end*, promotionally uptree of at least one of the doables A and B. For only then can the possibility of multitasking be denied, which was the point of the Common Branch Inference. But it is very difficult to imagine a rule of conduct that applies to an agent only under the hypothesis that she is doing a certain constellation of doables, but that does not bear on how to promote those doables. It is not a hypothetical imperative, for while a hypothetical imperative does indeed kick in for an agent only under the hypothesis that she contingently happens to be doing some doable or other, what

76

it requires of her is the adoption of a *means*, a doable that is promotionally downtree of the doable-hypothesis that made the hypothetical imperative kick in. But it is not a categorical imperative either, if such there be, for while a categorical imperative does indeed require of an agent that she adopt an end, a doable not promotionally downtree of any other, it applies to all agents categorically, and not only under some contingent practical hypothesis. The norm on which we are currently imagining the Common Branch Inference resting, however, straddles the distinction between hypothetical and categorical imperatives, for it supposedly requires the agent to adopt an *end*, even though this norm kicks in for her only under a contingent practical hypothesis. The guru would be required by this norm to do something (*drinking coffee and reading the newspaper*) not in order to do anything he is already doing, but nonetheless only because of some of the things he is already doing (*drinking coffee* and *reading the newspaper*). Such a norm would not belong to any already recognized class of imperatives, and it is for that reason difficult to imagine how it could be a norm at all.

For these reasons, it seems best to reject the appeal to excellence.

3.7   SEVENTH SECTION

We will consider one more attempt to defend the Common Branch Inference. The upshot of the previous section was that, if the Common Branch Inference is justified, then that is only because the true account of what it is to do or perform anything at all obliges us to posit a common end for A and B whenever an agent is doing A and doing B at all, rather than because that is how things go when they go well. But it might be thought that we must ascribe to the guru a common end for A and B, simply because otherwise we would be unable to *explain* (in perhaps some

special sense) the fact that he does both of these doables. This view is perhaps most likely to be propounded by those who think that there is some conceptual distance between the agent's *intending* of A (whatever that is) and her doing of A, that it is impossible to explain the latter except as caused or produced (in perhaps some special way) by the former, which must therefore be posited, but that her intending of A and her intending of B are together insufficient to explain why the agent ends up doing A and doing B. But we *can* explain the fact that the agent does A and does B, according to this "explanatory intentionalist" view, if we can attribute to the agent an intending of A and of B in order to do some third doable C, or of A in order to do B, or of B in order to do A. For if one of these three attributions were correct, then the doing caused or produced by the agent's intending would have to be either of A and B separately in order to do the common end C, of A in order to do the common end B, or of B in order to do the common end A, so that there would no longer be any mystery about how it happens that the agent does A and does B. This result would justify the Common Branch Inference, since it would entail that no one ever multitasks, and it would cover the guru's case, assuming that the doable C could simply be the doable A-and-B—that is, assuming that A-and-B is a doable.

If we grant for the sake of argument that a doing of A can only be explained (in perhaps some special sense) by positing an intending of A (whatever that is), the obvious challenge for this explanatory intentionalist view is to defend her claim that the agent's intending of A and her intending of B are together insufficient to explain why the agent ends up doing A and doing B. If we can be forgiven for saying so, this claim has a whiff of that species of superstition that sometimes drives people to seek explanations for surprising coincidences: "How improbable that we would both go to Saskatchewan on the same weekend!—there must be some deep explanation." We cannot deny that this drive to explain surprising coincidences has frequently

led to valuable insights, but our point is that there is positive reason to doubt that this is such a case, simply because the coincidence in question does not have to seem surprising at all. The so-called "coincidence" that the agent does A and does B only seems surprising and in need of (perhaps some special sort of) explanation if the explanatory intentionalist resolutely refuses to accept that in this case the agent's doing of A is satisfactorily explained by her intending of A and that her doing of B is satisfactorily explained by her intending to B. It is unclear what grounds there could be for such a refusal, given that the explanatory intentionalist regards doings of A as adequately explained by intendings of A in cases where the prospect of multitasking does not arise. Why should this explanatory force be undercut in cases where the agent is also keen on doing B?

But perhaps the explanatory intentionalist can convince us to be puzzled about this point. There will remain the question of how and why the explanatory intentionalist thinks that the intending of a doing with a common end can generally be counted on to arise. Say the agent starts out intending to A and intending to B, and ends up doing A and doing B. Since it would be unacceptably mysterious, according to the explanatory intentionalist, for this agent to have just done A and done B without any further ado, the explanatory intentionalist must attribute to the agent, between her initial stage of intending and her final stage of doing, an intermediate stage of intending to do A and to do B in order to do C, intending to do A in order to do B, or intending to do B in order to do A. But how and why does the agent undergo this transition to this intermediate stage? Unless it is itself no more than a very surprising (in other words, *ad hoc*) coincidence, the agent undergoes this transition because only thereby can she achieve the ends A and B with which she started out.

Our response to this suggestion is to eliminate, one by one, various plausible understandings of this claim. First, the claim might mean that the agent's undergoing of this transition is something that she *does* (in our sense that is coordinated with the term "doable"), *in order to* do the distinct ends A and B with which she began. The first thing to note about this account of the transition is that it would then appear to be a case of "intending at will," the possibility of which is vigorously disputed.[41] But more decisively, if the agent's undergoing of this transition is something she does in order to do the distinct ends A and B, then it will itself be a case of multitasking, and a counterexample to the explanatory intentionalist's claim that multitasking is impossible. For it will be something that she does in order to do A and also in order to do B. It follows that this transition is of a sort that can generally be counted on to happen in the life of an agent, either because she can somehow generally be counted on to do it (but not *in order to* do the distinct ends A and B with which she began), or as a result of some other generally reliable process.

Now, the explanatory intentionalist might respond to this problem by dropping the supposition that undergoing this transition is something the agent *does in order to* do the distinct ends A and B, and by retreating to the mere claim that undergoing this transition is something the agent simply *does* (in our sense) *because* undergoing this transition is (in some substantive sense in need of clarification) an excellent way to conduct herself as an intender. Note again that this account, like any account on which the transition is something the agent does, would commit the explanatory intentionalist to the possibility of "intending at will." But more decisively, the explanatory intentionalist who takes this line will merely have modulated the appeal to excellence into an intentionalist key, and her view will be vulnerable to intentionalist

---

[41] For a vivid presentation of this issue, see Kavka (1983). The question here is whether an agent can intend to *intend* to do A.

modulations of the three objections that we posed for the original version of the appeal to excellence: "ought" does not imply "is," the proposed "ought" is an insane one, and anyway it could not even have the reassuring status of either a categorical or a hypothetical imperative.

The explanatory intentionalist's final retreat will be to drop all talk of the transition as something the agent does at all. She simply holds on to a normative account of the transition that does not require it to be something the agent does, but rather something that can be counted on to occur in the agent because it is excellent for such transitions to occur. But our three objections will apply with equal force even then, since ascribing this transition to some normatively responsive but non-agentive process or faculty will not change the fact that the relevant "ought" cannot imply "is," that the proposed "ought" is insane, and that the norm would fit in to no known category. [42] So the explanatory intentionalist is stuck with the claim that an intending of A and of B in order to do C, of A in order to do B, or of B in order to do A, always intervenes between intendings of distinct ends A and B and the doings of A and of B, if there are any, that they explain; but she is precluded from accounting for these interventions by claiming that they are done (in our sense) by the agent herself in order to achieve her ends or that they either are done or merely occur in fulfillment of some other norm. And we will close our discussion of the explanatory intentionalist's view by pointing out her obligation to give a non-normative explanation of these interventions is simply a modulation into an intentionalist key of the obligation we have already ascribed to defenders of the Common Branch Inference generally: to find a justification for the claim that no one ever multitasks in the true account of what it is to (intend to) do or perform anything at all, rather than in an account of what it is to (intend to) do

---

[42] I admit that this last point might lose at least some of its punch, however, since (very arguably) not all norms are imperatives.

or perform anything well. Whatever may be the attractions of explanatory intentionalism, therefore, an improvement in our prospects for finding such a justification is not among them.

3.8   EIGHTH SECTION

Having defended the possibility of practically justified multitasking against these objections, we are now in a position to return to the topic of practical options that inspired us to begin this investigation. A practical option is one of the things that an agent can do. The empiricist position against which we aspire to argue is one on which practical options are merely *given*, in the sense that an agent's having the practical options she has is something that does not yet require any work on her part. To bring this conception a bit closer to home, we may say that this form of empiricism tries to model practical options on the agentive situation in which we find ourselves when we are shopping in a tie shop, or when we are making bets in a casino. There are various *self-standing* options available and the agent may *select* one, or some. And *what* practical options they are is always worn on their sleeves—or in other words, is always given.

Let us now suppose of some practical option that it answers to this characterization. To make things definite, let us suppose that a certain dance—call it the gadfly—has taken the nation by storm, a dance that consists in simultaneously raising both of one's arms while stepping forward. (Let it be stipulated that "dance facts" supervene on "limb facts.") It should then be plain that, at least in very many cases, when one has the practical option to dance the gadfly, one also has at least a few others— the practical option to raise both of one's arms and the practical option to step forward—and that these are compossible, meaning that the agent can multitask in performing each of them as an end.

A question that now arises is how it comes to pass that someone who is singletasking in dancing the gadfly *is* dancing the gadfly rather than merely multitasking in raising both her arms and, at the same time, stepping forward. What is the conceptual glue that in one case but not in the other glues these doables together into a single doable? Naturally, our account of multitasking enables us to say that in the case of singletasking, the agent does these two doables (raising both her arms and stepping forward) *in order to* do the further doable *dancing the gadfly*, and therefore that the answer to our question turns on the right way to understand the promotion relation—the right account of the semantics of this notion that we have so far treated as quasi-syntactic. But that is a way of reformulating the problem rather than solving it.

More to the point, however, we may ask how a pre-critical empiricist about practical options can maintain that the practical option to dance the gadfly is given to an agent. For it should be plain that an agent to whom it is given as a practical option to raise both of her arms and also as another practical option to step forward is not thereby an agent to whom it is given as a practical option to dance the gadfly—she must also possess, or at least be somehow appropriately in tune with, the concept of the gadfly.

This is the point at which the inevitable affinity between pre-critical empiricism and a certain foundationalism or scheme/content dualism is likely to manifest itself in the inquiry into practical options. For an empiricist is likely here to say that one's capacity to recognize *complicated* practical options such as the gadfly in one's manifold of practical options depends upon which complicated doable concepts one has acquired; but that nonetheless there are certain *simple* or *basic* practical options such as raising one's arms and stepping forward such that one's capacity to recognize them in this same manifold does not so depend. According to this view, we have stacked the deck against pre-critical empiricism by choosing a rather complicated

doable concept. But there are at least *some* kinds of practical options that can be given, in the sense that the agent need apply no concepts in order to recognize them as the practical options that they are: they wear on their sleeves their contents as practical options. These are the "raw opt-data," so to speak.

The problem with raw opt-data, like any other kind of raw data, is that we can only get a grip on the idea that any kind of item is a kind of *data* in the first place by recognizing its data-content as some sort of bearing in what Sellars calls "the logical space of reasons, of justifying and being able to justify what one says", or, as we may add, what one does.[43] For example, we can only get a grip on the idea that a certain practical option has as its content *stepping forward* if we can position this content in justificatory relations. Since the justificatory relations among doables (conceived as practical options rather than as items about which we entertain beliefs) are articulated by the promotion relation, this entails that we must be able to position the content *stepping forward* in means-end relations of the sort that the promotion relation articulates. We must in some sense of "able" be able to do *the thing called justifying* with the concept *stepping forward*—we must at least sometimes know things such as how to justify our stepping forward by showing that we are doing it *in order to* dance the gadfly (or vice versa). Otherwise *stepping forward* could not be an option for us.

I should emphasize here that I do not mean to deny that there are basic actions any more than Sellars denies that empirical science rests on a foundation.[44] It may, for all I have argued,

---

[43] Sellars, EPM, §36. I am here relying on an internalist account of the content of action defended in the first chapter.

[44] Sellars himself is an empiricist in the following sense:

> If I reject the framework of traditional empiricism, it is not because I want to say that empirical knowledge has *no* foundation…. There is clearly *some* point to the picture of human knowledge as resting on a level of propositions—observation reports—which do not rest on other propositions in the same way as other propositions rest upon them. On the other hand, I do wish to insist that

be the case that whenever an agent acts, there are at least some doables that she does on that occasion without doing any other doables in order to do them. The point, however, is that in order for one of these doables to be a *doable* that she *does* rather than merely an occurrence in her life that happens to contribute toward her agentive success, the agent must know how to situate that doable in relations of practical justification. And this is the fact in which the refutation of pre-critical empiricism about practical options consists.

In order for it to have content and thereby be a bit of opt-data for us, we must know how to situate a practical option in relations of practical justification. But just knowing how to do this with that doable is *already* recognizing something of its justificatory valence. And this fact goes to refute pre-critical empiricism about practical options because recognizing some justificatory valence is *already* doing something—forging the practical option according to our own conception of the promotional relations among doables and thereby taking responsibility for a justificatory relation between one thing we can do (or say) and another thing we can do (or say). Even though (perhaps) there are occasions on which we can act in a certain way "for no particular reason," it would not be *acting* in that way (as opposed to some other), as opposed to merely having something of that sort (as opposed to some other) happen in our lives, unless we were practically deploying the concept of acting in that way; and having the concept of acting in that way ready for deployment *is* knowing how to connect the concept of acting in that way instrumentally to concepts of acting in other ways. So our recognition of these orienting possible promotional relations must be written into the practical options themselves, and since the

the metaphor of 'foundation' is misleading in that it keeps us from seeing that if there is a logical dimension in which other empirical propositions rest on observation reports, there is another logical dimension in which the latter rest on the former. (EPM, §38)

recognition of these promotional relations is *ours* and not the *world's*, it must be our work and not that of the world that writes it into them.

Why does it matter to refute pre-critical empiricism about practical options? I think it matters because we learn from this refutation that the kill-one-to-save-five-style conundrums with which much ethical philosophy is concerned are conundrums in which an essential aspect of the agentive (ethical) life is left out—the work of *ascertaining what the options really are*. It can, for various purposes, be worthwhile to abstract away from that work, or to consider situations in which it has already been done. But if the Sellarsian argument I have sketched here is correct, ascertaining what the options really are is not a job that can be done for us in real life, is a part of the job of living an ethical life that takes place *prior to* the selection of options, and is therefore a job we had best start figuring out how to do if we are also going to have a philosophical literature of ethical option-selecting. That is my ambition in the remainder of this dissertation.

4.0   MULTIJUSTIFICATION

4.1   FIRST SECTION

In the preceding chapter we used a collection of dummy concepts—*doing*, *doable*, *promotion* and *promotional source*—to establish that any correct theory of agency must make room for the possibility of practically justified multitasking. Multitasking, on this minimally committal account, occurs when there is more than one end (that is, more than one doable that does not promote another) among the collection of doables an agent is doing at a particular time. Our only restriction on the possibility of multitasking *simpliciter*, which is summed up in the Common Trunk Inference, is that the tree that exhibits the structure of the promotion relations among this agent's doables must have a common promotional source, the single agent, who functions as the trunk from which all of her doables branch. So, if our argument is right, no account of agency can be right that does not leave this possibility open when it supplies our dummy concepts *doing*, *doable*, *promotion* and *promotional source* with content. It is the *promotion* and *promotional source* relations that especially matter in this connection, since the criterion of adequacy for accounts of agency that we have derived from the possibility of multitasking requires these relations to be such that an agent can be the promotional source of multiply many doables that are not promotionally related. And notice that this criterion of adequacy would be conveniently satisfied if the promotion relation were such that a single

doable can promote multiply many doables that are not promotionally related, as happens when the guru walks into the garden in order to drink coffee *and also* in order to read the newspaper. For the tree whose structure exhibits the promotion relations among an agent's doables can branch, and the ends of the different branches—the doables that the agent does not do in order to do any further doables—correspond to the different tasks she performs in each doable downtree of those branch-ends.

I now propose to add another dummy concept to our account—the concept of a *reason to do something*, or of a *practical reason*, and the coordinate concept of *practical justification*, or *justificatory status*. The relevance of these concepts to our account should be fairly clear, since it is generally agreed that whenever an agent does B in order to do C, her doing of B derives at least some of whatever practical justification it enjoys from the promotion relation it bears to her doing the doable C. The structure of what is called "instrumental" practical justification covers these cases: as long as the agent is doing, or perhaps merely intends to do, the doable C, a doing of B will be instrumentally practically justified in virtue of the fact that doing B would promote doing C. The first point we shall make with this dummy concept *practical reason*, then, is that the tree that exhibits the promotional relations among the agent's doables—thus allowing us to count the tasks she is performing by counting the ends of its branches—is also a tree that exhibits the structure of the instrumental practical reasons on which she acts when she does the doables represented in that tree. And it follows, given that instrumental practical justification is transitive, that every doable downtree of a particular branch-end in this tree derives instrumental practical justification from that branch-end, with the result that some doables—those closer to the trunk—may well derive instrumental practical justification from *more than one* end. We can say, therefore, that these doables are multi-justified.

There are some philosophers (quintessentially, Kant) who maintain that, at least some of the time, this tree will provide an incomplete picture of the practical justification that an agent's doables enjoy. For, according to these philosophers, there are practical reasons that provide practical justification in ways not captured by the structure of instrumental practical justification, and therefore "off the tree" that merely represents the promotional relations among the agent's doables. These practical reasons are attached to the tree like Christmas ornaments—not part of what the agent does, or at least, not in the same sense, but contributing, perhaps very greatly, to the justificatory status of her doing. Unlike ornaments, however, of which trees have no special need, these practical reasons are generally thought to be indispensible in such a way that although they are themselves no part of what the agent *does*, they are nonetheless the sources of whatever practical justification is enjoyed by any of her doables. Those who believe in these "ornaments" therefore tend to believe that these non-instrumental practical reasons can only be attached to the tree of an agent's doables at her ends, and that whatever practical justificatory status is instrumentally transmitted downtree by promotion relations can only ever be initially conferred on its source or sources by these non-instrumental reasons.

Now, one point we want to make with the dummy concept *practical reason* is that the possibility of multitasking obliges theorists of agency and of practical justification—including both those who do and also those who do not believe in non-instrumental practical reasons, or "ornaments"—to supply this dummy concept with content in a way that is compatible with the possibility of multi-justification.[45] But we have only demonstrated that multi-justification is

---

[45] Nietzsche, of all people, makes it clear in the following quite beautiful passage (*Beyond Good & Evil*, §215) that he believes that it is possible for an action to be multi-justified:

possible in cases of multitasking, in which an agent's doing of a doable A instrumentally derives practical justificatory status from multiply many ends uptree of it. It is therefore natural to wonder whether an agent's doing of a doable A can be multi-justified even in cases of single-tasking, where there is only one end uptree of A. Now, those who do not believe in non-instrumental practical reasons are obliged to maintain that multi-justification is impossible except in cases of multitasking, since they do not believe in any kind of justification for a doable except that provided by an end. But it is an open question whether those who do believe in non-instrumental practical reasons can or must regard it as possible for a doable to be multi-justified even when there is only one end uptree of it. Since we shall assume, for simplicity, that non-instrumental practical reasons cannot be attached to a doable that is downtree of an end, our question would be settled if we knew whether a single end can be practically justified by more than one non-instrumental practical reason. Notice that this question about multi-justification is structurally akin to our initial question—whether and how multitasking is possible. But we may also pose a question structurally unlike that initial question: whether a single end can be practically justified by a non-instrumental practical reason that *also* practically justifies another end of the same tree. And just as our answer to our initial question helped us to understand what sorts of things tasks are, there is reason to hope that our answers to these new questions—the first about multi-justification of ends, and the second about multiple points of justificatory contact between a non-instrumental practical reason and a tree that exhibits the promotional

As in the realm of stars the orbit of a planet is in some cases determined by two suns; as in certain cases suns of different colors shine near a single planet, sometimes with red light, sometimes with green light, and then occasionally illuminating the planet at the same time and flooding it with colors—so we modern men are determined, thanks to the complicated mechanics of our "starry sky," by *different* moralities; our actions shine alternately in different colors, they are rarely univocal—and there are cases enough in which we perform actions *of many colors*.

structure of an agent's doables—will put us in a position to say what non-instrumental practical reasons are, if there are any, since it will put us in a better position to count them.

4.2   SECOND SECTION

Our response to the first of these questions—whether a single end can be multi-justified by non-instrumental practical reasons—is simply that it would be very difficult to deny that multi-justification of ends frequently occurs in the lives of ordinary agents. It would be bizarre, for example, to insist that no agent ever has more than one non-instrumental practical reason for voting for one candidate rather than another in an election. Surely an agent can vote for her candidate because she supports that candidate's policies foreign *and* domestic. And we can add further heterogeneity to the agent's collection of non-instrumental practical reasons by noticing that she might equally well at the same time vote for her candidate because she promised her friends that she would, because her chosen candidate has many impressive credentials, and because the rival candidate insulted her family ten years ago in a private setting. So we shall take the possibility of such (at least apparent) heterogeneity of non-instrumental reasons as a datum, and try to see whether a skeptic about the multi-justification of ends can explain what happens in these cases in a satisfying way.

Now, the skeptic might maintain that, contrary to appearances, these considerations— foreign policy, domestic policy, promises, credentials and revenge—do not, or do not all, provide

distinct non-instrumental practical justification for what the voter does.[46] For there are many

different theories about the nature of non-instrumental practical justification, and some of these

theories rule out for one or another of these considerations the possibility that it can genuinely

provide non-instrumental practical justification, or that it can provide non-instrumental practical

justification that is importantly distinct from what is provided by the other considerations. It is

common, for example, for those who believe in non-instrumental practical reasons to hold that

for each doable A there is at most one non-instrumental practical reason to do it. A contentious

reading of Kant (which we do not mean to endorse, and which is almost certainly wrong) will

make this point vivid. On this contentious reading, Kant's claims that there is no practical reason

that is not either a hypothetical or a categorical imperative, and that while there are many

hypothetical imperatives there is only one categorical imperative, together entail that to each

doable A there corresponds exactly one non-instrumental practical reason that could (if it

obtained) justify doing A as an end: the fact, if it were one, that *doing A is my duty*. So, on this

reading of Kant, if any two of the considerations we have attributed to our voter count as non-

instrumental practical reasons for her to vote as she does, they do so only by having some sort of

bearing on what the voter's duty consists in. But since the only non-instrumental practical

reason that can *really* and directly justify an agent's end, on this view, is the fact that adopting

this end is her duty, other considerations can only ever indirectly prompt agents to adopt ends by

helping to constitute the agent's duty.

This seemingly Kantian position exhibits what we shall call the "vector sum structure,"

since it maintains that the only way any consideration other than that *A is her duty* can contribute

---

[46] This will of course be the line taken by those who do not believe in non-instrumental practical reasons at all, according to whom all of these various considerations are merely misleading manifestations of considerations that are really just instrumental reasons. This argument is not addressed to such people.

to the non-instrumental justification of doing A is by somehow entailing or constituting the fact that *A is her duty*, which is therefore a vector sum of all the other considerations. And any view on which there is a special property Φ such that the only non-instrumental practical reason an agent can have for doing a doable A is that *A is* (or would be) *Φ* will exhibit this vector sum structure, regardless of whether this single property is supposed to be that of *being to the agent's advantage*, *being such as to promote world utility*, or perhaps something else. For on any of these views, a bit of non-instrumental practical justification can transfer to an agent's end only by being resolved into a reason of the form *A is Φ*.

It should go without saying that a view of the sort we mean to indicate cannot regard the property Φ as a disjunction of two or several more basic properties. For views of this sort aspire to find a property Φ that is common to and that explains all genuine cases of non-instrumental practical justification, whereas to say that this property Φ is a disjunction of more basic properties is really to say that there is no such property—or in other words, it is to give up the attempted explanation.

Distinctions between different views of this sort are most often characterized as distinctions between different candidate-properties for the role of this property Φ. But we may equally well say that the proponents of different views exhibiting this structure are merely inclined to refer to this single thing, the property Φ, by different names, such as the agent's "duty," her "advantage," her purchase on the world's "utility," or perhaps something else. On this way of viewing matters, the point of each of these names is merely to suggest the primacy of a certain method of resolving into a single non-instrumental practical reason the diversity of the considerations that bear on what to do. So we can trade in different candidates for the property Φ for different candidates for what is supposed to be the only practically rationally excellent

method of resolving considerations into non-instrumental practical reasons. This is not a substantive move, for all we learn if we learn that (for example) the property Φ is *the agent's duty* is that the method we must use to determine what she has non-instrumental reason to do is *finding out what her duty is*.

So every view on which there is only ever at most one non-instrumental reason to do a doable A—the fact, if it is one, that *A is Φ*—is committed to the claim that there is only one practically rationally excellent method for resolving considerations into non-instrumental practical reasons. Naturally, any attempt at such a resolution will involve various "submethods"—*in one's head*, *on paper*, *by reading the* Groundwork, *by consulting Wikipedia*, and so on. But according to this kind of view there can only be one "supermethod"—only one way of finding non-instrumental reasons that is not itself a submethod of some more architectonic method of finding non-instrumental reasons. For if there were more than one practically rationally excellent supermethod for resolving considerations into non-instrumental practical reasons, then it would be possible, contrary to the basic idea of this approach to non-instrumental justification, for a doable to be justified by more than one non-instrumental reason—first, *that A is the result of Method M_1*, and second, *that A is the result of Method M_2*. (We shall use "method" to mean "supermethod" henceforth.)

Now, the vector sum structure has entered our dialectic as an attempt to deny that ends can be multi-justified. We shall assess the success of this attempt shortly. But it will help us if we first consider whether views exhibiting the vector sum structure are even consistent with the possibility of (practically justified) multitasking, or whether instead a commitment to the vector sum structure entails the guru's doctrine (which we have already refuted) that multitasking is never practically justified. For the point of saying that a doable A is Φ is that the totality of the

practically rationally relevant considerations resolve themselves in favor of doing A when we evaluate them with the *only* practically rationally excellent method for generating non-instrumental reasons; and if that is so, then it would seem that the agent lacks justification for doing another doable B as an end of the same tree of doables that has A as an end justified by the Vector Sum. For the doable B would then not be the Vector Sum.

As it stands, this argument is unconvincing. For it presumes that the totality of the practically rationally relevant considerations can only resolve themselves in favor of a *single* doable, whereas the Vector Sum theorist might well try to maintain that the totality of practically rationally relevant considerations can issue more than one recommendation, each of which would then count as *a* Vector Sum. It may be, therefore, that this totality recommends that the agent adopt A as an end, and also that she adopt B as an end. And in this way the Vector Sum theorist might try to defend her view from the charge that it entails the guru's doctrine.

We therefore do not argue that ascribing a vector sum structure to non-instrumental reasons is in general incompatible with the possibility of practically justified multitasking. That will depend entirely on the property Φ that a given Vector Sum theorist proposes to explain the concept of a non-instrumental reason. A Vector Sum theorist who thinks that the property Φ is, say, the property of *being required by God*, and who thinks that many different doables are required by God, can maintain that the guru's behavior in the garden is practically justified if God requires both the doable *reading the newspaper* and also the doable *drinking coffee*. And no argument against this sort of view can be worked up from the idea that, since a Vector Sum theorist can only recognize one method for determining which non-instrumental reasons an agent has, she must think, implausibly, that an agent can only identify one of her non-instrumentally justified doables by abstracting it from the result of a single application of this method. For the

agent may simply have applied this method well enough to know some of the doables that God

requires without thereby having put herself in a position to know all of them: she may have read

the commandments in Exodus without yet having read any of the laws in Leviticus.

In this way, many of the views that exhibit the vector sum structure can admit the

possibility of practically justified multitasking: there can be what we may call two different

"constellations" of considerations—E and L, say—one of which is *sufficient*, according to the

standards of *the* method for determining facts about the property Φ, to recommend the doable A,

and the other of which is sufficient, according to those standards, to recommend the doable B.

We have no objection to this "constellation" move; indeed, given a sufficiently bland candidate

property Φ, a view having this structure and making this move may well be true, although it

probably does not have as much explanatory power as its proponents would like. But we may

return to the main thread of our argument by pointing out that a Vector Sum theorist who makes

the constellation move will be ill-equipped to persist in her denial of the multi-justification of

ends, which was supposed to be the point of the vector sum structure for our own dialectic. All

that a Vector Sum theorist denies is that considerations other than the consideration, if it obtains,

that *A is Φ* can have *direct* non-instrumental practical rational bearing on an agent's end. But a

constellation of other considerations can have *indirect* non-instrumental practical rational bearing

on an agent's end A. And furthermore, *more than one* constellation of other considerations can

have indirect non-instrumental practical rational bearing on the *same* end A. Our agent may

learn not only from Exodus but also from Leviticus that God requires her to rest on the Sabbath.

Or if we return to our original example, our agent's voting for her preferred candidate, which

seemed to be justified by so many distinct non-instrumental practical reasons, can *still* derive

separate sufficient practical justificatory status from each of them, if each of them represents a

distinct constellation. For a constellation is a distinct bearer of justificatory force that flows through the conduit of the fact that it entails, the fact that (say) voting for her preferred candidate is her duty, or would promote world utility, or perhaps something else.

To sum up: our criticisms of the "appeal to excellence" in the previous chapter constitute an argument that practically justified multitasking is possible; but a Vector Sum theorist can only countenance the possibility of practically justified multitasking by making the "constellation" move; and making the constellation move undercuts her aspiration to claim that ends cannot be multi-justified. This is not to say that the dispute between advocates and opponents of the vector sum structure is *entirely* verbal. It is not, for its advocates maintain, while its opponents deny, that no consideration ever confers non-instrumental justificatory status on a doable A except through the conduit furnished by the consideration, if it obtains, that *A is Φ*, where the property Φ is sufficiently contentful to *explain why* doing A is justified instead of merely saying that it is. Our point is just that this dispute has limited significance for the question of whether multi-justification of ends is possible, since as long as they think that practically justified multitasking is possible, advocates and opponents of the vector sum structure will have to agree that *indirect* multi-justification of ends is possible, by *distinct* constellations of considerations each of which confers non-instrumental practical justificatory force *sufficient* to constitute a non-instrumental reason to do an end, and that is good enough for us.

Our final remark about views that exhibit the vector sum structure is that their proponents are obliged to answer our second question about non-instrumental reasons—whether the same non-instrumental reason can justify more than one end in the agent's tree of doables—in the negative. For it is plain that if the only direct non-instrumental reason there can be to do A is that *A is Φ*, whereas the only direct non-instrumental reason there can be to do B is that *B is Φ*, it

97

cannot happen that two ends A and B are directly justified by the same non-instrumental reason, simply because the consideration that *B is Φ* is not the same consideration as the consideration that *A is Φ*. Or if they are the same consideration, it is only because A is the same doable as B—which would mean that they could not be distinct ends of the agent's tree of doables. Of course, the consideration that *A is Φ* might perhaps function as an *indirect* non-instrumental reason to do B, if somehow the fact that *A is Φ* enters as one among many other facts into a constellation that entails that *B is Φ* as well.

Now, we might sum up this remark by saying that a single constellation of non-instrumental reasons—where by "constellation" we mean not just the reasons themselves, but also a structure of practical rational connections among them, insofar as these two aspects of a constellation can be distinguished—cannot, according to a view with the vector sum structure, directly non-instrumentally justify doing more than one doable as an end. We can close our discussion of vector sums by pointing out that this is not a mere quirk of views with the vector sum structure, but is quite general: whatever a *single* constellation of non-instrumental reasons justifies must itself be a *single* end. This is not to deny that a single constellation of non-instrumental reasons can leave certain features of the end it recommends indeterminate—"shall I take the red umbrella or the purple umbrella?"—so that what it recommends is that one do A *or* B. For in these cases, the constellation of non-instrumental reasons recommends that one adopt a *single* end—*one* of A or B. Our point is rather that whenever it looks as if a constellation of non-instrumental reasons recommends that one adopt *more* than one end—A *and* B—then what this constellation really recommends is that one adopt the conjunctive end A-and-B. Otherwise we could separate from this single constellation those aspects of it that recommend A from those that recommend B, and in that case it would not be a *single* constellation, but rather a pair of

constellations that had somehow gotten tangled together. For, and this is the heart of the matter, we may derive the unity of a single non-instrumental practical reason (if such a thing exists) from the unity of a set of doables into a single task that it justifies. Let us call this the "coordination thesis."

4.3   THIRD SECTION

The question we propose to consider in defense of this coordination thesis is whether it is always possible to "divide" cleanly without remainder the non-instrumental practical justificatory force directed at a certain doable into further independent components, each of which would then have to be understood as conferring justification on independent components of the doable. Suppose, for example, that our agent has a non-instrumental practical reason to dance the gadfly, which consists in simultaneously raising one's arms and stepping forward. It follows that, for this agent, there is non-instrumental practical justificatory force directed at the doable *dancing the gadfly*. Our question, then, is whether this is true *if and only if* the agent has a non-instrumental practical reason R to raise her arms and also a non-instrumental practical reason S to step forward, or, in other words, *if and only if* there is non-instrumental practical justificatory force directed at her raising her arms, and also non-instrumental practical justificatory force directed at her stepping forward, the two of which simply "add up" to the non-instrumental practical justificatory force directed at her dancing the gadfly. And our response to this question is that an affirmative answer leads to an unacceptable regress. For if practical justificatory force only ever manifests itself as a kind of divisible rational quantum, then the same will be true of the doables to which that practical justificatory force is attached. The non-instrumental justificatory force

directed at the agent's raising of her arms, for example, will itself divide cleanly without remainder into some non-instrumental justificatory force directed at raising her left arm, and some other non-instrumental justificatory force directed at raising her right arm—and so on.

We take it to follow from these considerations, therefore, that non-instrumental practical justificatory force manifests itself indivisibly—or, if you like, organically, in such a way that no mere list of its components "adds up" to the practical justificatory force of which they are components. We may draw an analogy with Frege's context principle:[47] just as it is only in the nexus of a sentence—a potential bit of knowledge, if it is a satisfactory response to theoretical reasons—that words have meaning, it is only in the nexus of a potentially satisfactory response to practical reasons that doables have the significance for the agent of something she does. For the practical justificatory force that attaches to a complete task is prior to the practical justificatory force that attaches to each of the doables that promote that task, and explains their justificatory status, rather than the other way around. We shall say that such indivisible or organic manifestations of non-instrumental practical justificatory force are "unified" bits of non-instrumental practical justificatory force.

Now, we have no wish to compel the reader always and everywhere to use the term "non-instrumental practical reason" exclusively for the considerations that confer unified non-instrumental justificatory force. We cannot help remarking, however, that one point in favor of adopting this usage is that the word "reason" is a count noun, which obliges those who use it either to assimilate this feature of the word to the misleading contingencies of surface grammar, or else to have or seek a principle of individuation for the individual reasons to which it refers— a principle of individuation which we have suggested a way to discern. But philosophers have

---

[47] Expressed, for example, in *The Foundations of Arithmetic* §60: "Only in a proposition have the words really a meaning."

found it helpful to use the term "practical reason" in many different ways, and have derived valuable insights into the nature of agency by doing so, so we shall not insist. As Wittgenstein says with admirable calm, "Say what you choose, so long as it does not prevent you from seeing the facts."[48] The fact presently facing us is simply that there is something in need of a label—namely, a non-instrumental consideration that confers *unified* justificatory force. We shall reserve the term "unified reason" for this sort of consideration.

Our coordination thesis, then, is vindicated by the fact that we must recognize a sort of unity characteristic of such non-instrumental practical considerations. The practical justificatory force conferred by a unified non-instrumental practical reason is itself unified. So this force cannot be divided cleanly without remainder into distinct components, one of which non-instrumentally justifies the task A, while another non-instrumentally justifies the task B. There must be something that remains, something that leaves the mere conjunction of these two tasks incomplete as a response to the unified non-instrumental practical reason. This remainder is a necessary connection between the two tasks, since what the non-instrumental practical reason required was the performance of both of them *together*. This requirement entails as a minimum that the agent does the doable A-and-B as a single doable, as for example the *gadfly*—the raising of the arm and the stepping forward—is a single doable promoted by every doable downtree of it. And if this is so, then it cannot be right to characterize what the non-instrumental practical reason justifies as two independent tasks; rather, what it justifies is a single task with these two parts that have a necessary connection to one another, merged together into a single concept like the *gadfly*.

---

[48] Ludwig Wittgenstein, *Philosophical Investigations* §79.

This is in general how it happens that two doables A and B merge into a single doable A-and-B, which we can now more helpfully render as the doable A-*while*-B. And this merging is prompted by the non-instrumental reason (whatever it may be) for the agent to do this. For if we have non-instrumental practical reason to do A only because we have a unified non-instrumental practical reason to do A-while-B, then it is clear that what a unified non-instrumental practical reason to do A-while-B is a reason to do is a single end. After this non-instrumental practical reason gets up and running, we may even give this single end a name, as we give names to different kinds of dances, like "gadfly," "tango" or "foxtrot." And it is only an accident of linguistic history that, say, chewing-gum-while-walking has not been given a name, since it is like a kind of dance that people can endeavor to do for a single reason—typically, to prove that they can do it. For a conjunctive doable that can be done as an end is no *mere* conjunction of two distinct doables, but a coordinated one, where the principle of their coordination is supplied by the unified non-instrumental reason that there is to do them both together. This non-instrumental practical reason therefore supplies a kind of conceptual glue that makes a single doable out of what was once a pair of distinct doables.

Now, this way of viewing matters should not be taken to entail that two doables can merge together into a single doable *only* if a non-instrumental practical reason prompts the agent to do them both together. Once the tango enters the agent's repertoire of doables, she can do this doable on a whim, without a non-instrumental practical reason to justify it. (This is not to say that she can be justified in doing the tango in this way.) So we do not claim that two doables A and B are only ever merged into one in an agent's doing if they are prompted by a reason to do A-while-B. Nor, pulling back to a higher level of generality, do we claim that two doables A and B only start to constitute a single doable in the agent's repertoire of doables that she *can* do after

102

she has once had a non-instrumental practical reason to do A-while-B. Our claim is rather that the fact that an agent does both A and B in order to cope with a single non-instrumental practical reason is *sufficient* to make it the case that these two doables have merged into a single doable, which is promoted by all of her doables that promote A and also by all of her doables that promote B. And at the level of the agent's repertoire of doables, our claim is that the sorts of things that there are in this repertoire are themselves unified, and that this repertoire expands if and only if the collection of doables that the agent can have a practical reason to do expands in kind.

The most pressing question that remains is how we are to understand the way in which a non-instrumental practical reason is fitted to a doable that it justifies an agent in doing as an end *when all goes well*, that is, when the agent does that doable as an end for that non-instrumental practical reason. An agent may be thought of metaphorically as a kind of membrane across which practical reasons and potential actions have various justificatory relations, a membrane that is in a condition of health when those relations—and not *just* the actions that extant reasons justify—are realized in her activity. My final chapter will investigate what this sort of health consists in by considering what goes wrong when an agent suffers from the sort of sickness to which it corresponds.

5.0   IN DEFENSE OF THE PRACTICAL IMAGINATION

5.1   INTRODUCTION AND DIAGNOSIS OF THE OUGHT-OUGHT GAP

My goal in this chapter is to understand how to close what we might refer to as the "ought-ought gap." For now we may simply think of this as the gap expressed when someone says, "Of course, such-and-such is the right thing to do—but why should I do it?" This gap opens up when we misinterpret the significance of thinking of morality as a system of commands, constraints, requirements, or something of the kind. Let us refer to this as the "command picture" of morality.

The command picture of morality derives its plausibility from our ordinary phenomenon of issuing and obeying commands. For only on the basis of our familiarity with this phenomenon could we illuminate the concept of moral obligation by saying that morality is like a system of commands. So it will serve us well to consider this ordinary phenomenon of issuing and obeying commands, in order to grasp what sort of understanding it can give us of morality.

A command is a quite complicated thing. For one thing, the word "command" is ing/ed ambiguous: it can refer either to the act of commanding someone to do something, or to the thing that that person is commanded to do. With the word "command," we therefore refer to a complicated relationship between two people and a thing that can be done. But a commander abuses this relationship if she commands her subordinate to do A without having good reason

why A should be done. So let us say that a command is "exemplary" when the person who commands her subordinate to do A has some good reason R why A should be done. An *exemplary* command is therefore a complicated relationship between two people, a reason, and a thing that can be done. The possibility of a bad or abusive command—in which the commander has no good reason why A should be done—derives from the fact that most commands are exemplary, much as the possibility of a lying promise derives from the fact that most promises are kept. The relationship between a commander and her subordinate is therefore a kind of division of rational labor, for in this relationship, the phenomenon of responding to a reason R is split up among two people, the first of whom is in charge of appreciating the practical rational significance of this reason and the second of whom is in charge of doing what counts as a response to it.

We may now puzzle about why a subordinate ought to obey a command. For if it is right to think of commanding as involving a division of rational labor, then when the subordinate does A because it is commanded by a commander in response to the reason R, the subordinate is doing A in response to R. That is because a doing of A in response to R is constituted by the interchange of command and obedience between the commander and the subordinate, for the commander functions as the conduit of the rational (justificatory) force of the reason R. But the subordinate need not herself appreciate this fact, since that isn't *her* job according to the way that rational labor is divided. All that the subordinate seems to require from her own point of view is that her commander commands it. She conceives her activity as justified *because* it is obedient, even though this is nothing like the whole truth or the most important truth, which is that she does A because of the reason R. The subordinate therefore thinks: "*I shall do A* because that is the way to be obedient," while her commander thinks "she shall do A in obedience to me

105

because of R." We may say that the principle of the subordinate's action is the commander's command, but the principle of the command, and therefore of the entire process—though the subordinate can only ever have accidental access to it—is the reason R.

Now, there is nothing wrong with dividing rational labor in this way in circumstances that call for it. Military personnel, for example, among whom commands are so often exchanged and obeyed, have a practice of obeying commands not because that is in itself an excellent way to live, but because the function of the military is to address situations of a sort in which those who know best what needs to be done (generals and strategists) are seldom those best in a position to do it (low-ranking soldiers), and in which there is seldom sufficient time for the former to explain the relevant considerations to the latter so that the necessary process of responding to reasons can achieve its characteristic unity in the cognition of a single agent. That is why, in this particular realm of life, there is a certain division of rational labor in which the rational role of appreciating the reasons there are for acting is separated from the rational role of acting appropriately. But this is a kind of emergency measure to deal with special circumstances. And even in the military, no one is immune to criticism, even from the point of view of her subordinates (considered as voters, say), in contexts in which these special circumstances no longer constrain what it is possible to think. We may always wonder whether General S was right to give a certain order, and we can only determine that she was or wasn't right by referring to her reasons for giving it. This is why it is quite wrong to think that we could have the motivations of obedient subordinates globally, throughout our lives, rather than locally, in the special circumstances that call for them.

For if we take the command picture of morality to justify thinking of ourselves as *always* or somehow *essentially* on the subordinate side of this relation between commander and

subordinate, we will think of ourselves as not essentially in contact with the reasons there are to do the things that morality commands. We will begin to think, like subordinates, that we ought to be moral not because of the various reasons that there are to do the moral things (reasons that there must be if the commands to do those things are to count as exemplary) but simply because we are commanded to—for that is good enough in cases where relations of subordination are indeed justified. This is a confusion; it mistakes for the core phenomenon of acting well what is in fact a derivative case of it in which various essential elements of this phenomenon (appreciating the reason, and acting on it) are divided among different people, and it thinks that all that is required to actualize this phenomenon is to actualize one half of it.

Someone who has grown accustomed to this picture of morality finds it very hard to know what to do when she is told that she ought not only to do the right things, but also that she ought to do them for the right reasons. She has all along left responsibility for the reasons to the source of the commands, thinking that there is one somewhere to whom she can defer this responsibility. Such an unfortunately oriented agent thinks: "I know that morality requires me to do A, and yet I do not want to do A; or perhaps I do want to do A, but only for some ulterior motive, or because I enjoy something accidentally connected to it, or in order to appease my vanity about my moral stature."[49] If one finds oneself in such a fix, how can one bring oneself to do A for the right reason? It can sound like yet another odd trick one is supposed to pull, out of obedience, as if what one is *commanded* to do is to do the right thing for the right reason—as if the thing one were to *do* were to *do A for reason R*, while the reason why one should do *this* is still merely that it is commanded. It can be very frustrating to try to do something of this form, since it is impossible to gerrymander one's reasons after decision but prior to action. This would

---

[49] These are various cases that Kant distinguishes in which an agent can act *in accord with duty* but not *from duty*. (*Groundwork*, 397; for the point about vanity, see Kant's discussion of the "dear self" at around *Groundwork*, 407.)

quite clearly be to treat rationality as a means only: if one tries, out of obedience, to get one's reasons to accord with what one has already decided, one conceives obedience as one's end and rationality as a means to it. This makes it clear that to think of the thought (which is correct) that one ought to do the right things for the right reasons as yet another command with *do A for reason R* as its content is just to preserve the division of rational labor that attentiveness to reasons was supposed to overcome. For it is not as if there is some further reason R_2 for us to *do A for reason R*, as there would have to be if *do A for reason R* were just another thing we were commanded to do—a regress threatens, and wrongly, for it is plain that there is no such extra reason.

It follows that what we mean when we say that morality commands us to do A is better expressed by saying that there are reasons to do A that any good person is appropriately responsive to. That is, the logical form of a command is here entirely at the level of surface grammar, since what we are really doing is giving an existentially quantified clarification of the concept of a good person: *there is* some reason such that any good person would respond to it by doing A. This is the kind of thought that lies underneath the "ought-ought" gap; it has the same form as the thought: "I know that *there is* some number between 1 and 100 that maximizes this function—but what number is it?" Recognizing that existentially quantified facts like these lie under the grammatical surface of morality's commands will help us to escape the confused idea that one can always be good simply by being obedient in ignorance of what makes one's actions good. Quite the reverse: typically one can only be good by responding to the right reason by doing A on the basis of one's own appreciation of its rational significance. For one must respond to A on the basis of *someone's* appreciation of its rational significance, and since the relation of commander to subordinate is not the core case of moral excellence, the core case of moral

108

excellence must be that in which one responds to R on one's own authority, and not on the authority of something outside of us that commands—freely and not out of obedience.

It is worth pointing out, though this is merely an aside, that the bad connotations of the command picture of morality are exactly what consequentialism endorses. Let us understand consequentialism roughly as the claim that the moral value of an action is logically posterior to (because it is explained by, derived from, or something like that) some sort of value possessed by its consequences. Understood thus, consequentialism maintains that an action has its moral value independently of how we are motivated to perform it—so long as it is what Kant would call "in accord with" what morality commands, an action is as good as it can be as an action.[50] Consequentialism is therefore equivalent to the endorsement of the sort of global division of rational labor implicit in the command picture of morality. Since this picture requires no motivation from the agents who act well other than obedience, it alienates rational action from its rational ground. Consequentialism is tempting nowadays because nowadays rational labor is everywhere thoroughly divided up: each particular realm of life has its community of experts.[51] But the ubiquity of this phenomenon should not blind us to the essence of consequentialism as a temptation for how to cope with the phenomenon of the ought-ought gap. Consequentialism tells us not to worry about the "ought" closer to ourselves, that the moral worth of our conduct does

---

[50] Naturally, there may be some versions of consequentialism that violate only the spirit and not also the letter of this point, but Mill is certainly guilty. As he says (demonstrating, by the way, an astonishing ignorance of Kant's system), "It is the business of ethics to tell us what are our duties, or by what test we may know them; but no system of ethics requires that the sole motive of all we do shall be a feeling of duty…. [U]tilitarian moralists have gone beyond almost all others in affirming that the motive has nothing to do with the morality of the action, though much with the worth of the agent" (*Utilitarianism*, p. 18). Mill does not absolve himself of this guilt in his use, in a footnote subsequently appended to this passage, of a distinction between motive and intention, for he there continues to affirm: "The morality of the action depends entirely upon the intention—that is, upon what the agent *wills to do*. But the motive, that is the feeling which makes him will so to do, if it makes no difference in the act, makes none in the morality" (*Utilitarianism*, 18n2).

[51] The claim that a division of rational labor accompanies the sort of alienation manifested in the ought-ought gap obviously bears a strong affinity to a central thesis of Marx's critique of capitalism.

not depend upon it, even though it is the "ought" that is closer to ourselves that makes us human (or rational)—we are human (rational) because we can do something because we see the point of it (on the strength of reasons), rather than because it is required by some system of evaluation of which we need not have any glimpse in order for it to apply. And since the temptation of consequentialism will be with us for as long as the ought-ought gap opens up in our lives, someone who sets herself the goal of overcoming consequentialism will have not only to sketch persuasive demonstrations of its falsity, which is easy enough, but also to develop a more satisfactory strategy for dealing with the ought-ought gap—a strategy, that is, that does not constitute submission to the alienating command picture of morality, but which instead connects the person in whose life the ought-ought gap has opened up to the real reasons for action, or (even better) which obviates the ought-ought gap entirely. Something about us tempts us to think—though it is easy enough to see that this is false—that the command picture of morality is right after all, and we will not have closed the ought-ought gap for good until we address the source of this temptation.

5.2   THE ESSENTIAL INDEXICAL AND THE TORTOISE REGRESS

Now, the "ought-ought gap" is our fanciful term for a mental ailment in which one knows that there is some reason such that any good person would respond to it by doing A, but in which one does not know what that reason is. But this way of formulating the problem obscures a great many issues, for if we consider any particular case in which someone wrestles with the ought-ought gap, it is clear that she may know a great deal *about* the reason to do A and nonetheless fail to understand why she ought to do it. Anyone who says, for example, something like "I

know I should return the money—but if I don't, no one will ever know I acquired it, and I won't get into any trouble—so why should I do it?" is not entirely in the dark about the reasons that there are to return the money. We can even imagine her able to supply herself with a demonstration, grounded in secure *a priori* principles about (say) the concept of property, that morality requires her to return this money, and nonetheless unable to understand why she should return it. This shows that what someone in such a fix really needs is not merely to appreciate that a good person would return the money because of such-and-such a reason. For this can appear to her merely as an alien fact having nothing to do with her. What is really needed, therefore, is for her to recognize these facts as having to do with her. That is, she must relate herself to the concept of a good person in such a way that, for her, recognition that a good person returns the money for such-and-such a reason *is the same thing as* practical recognition of the rational force of this reason that manifests itself in her returning the money. In other words, not only must she learn to refine her concept of a good person to the point where it is adequate to her situation, but also she must think of *herself* as a good person; she must learn to think of herself as primarily or essentially what she learns about when she learns about good people.

For this it is not enough merely to hold the form of words "I am a good person" before one's mind in an approving manner. That is easy enough, and it is something one can do merely because one knows one ought to although one inadequately understands why that is so—in other words, it is something one can do "consequentialistically." When one tells oneself that one is a good person in this manner, it is as if one is giving oneself a moral pep talk, exhorting oneself to see oneself in a certain way, but without providing (because one does not appreciate them) any reasons why that is the right way to see oneself. Here the command picture of morality still has all of its misleading force. It follows that what is necessary, and also more difficult, in order for

111

an agent to close an ought-ought gap that has opened up in her life, is for her to think any form of words at all in which the thought (perhaps among others) that she herself is a good person is affirmed, in an appreciation of its truth. One way to think about the upshot of that claim is that the ought-ought gap seems to point us to a specifically moral case—and really it would be better to call it a specifically moral *aspect*—of Perry's "problem of the essential indexical."[52]

I say that this is a moral aspect of Perry's problem of the essential indexical because the "change in behavior" (PEI, 3) that this problem was a problem about how to explain—the change that consists in Perry's rearranging of the torn sack of sugar in his cart when he realizes that he himself is making a mess—is implicitly understood throughout his exposition as following that realization straightaway. That is, the absence of an ought-ought gap in the psychology of the agent is at least a necessary condition for this change in behavior for which Perry says we need an explanation. Perry avoids drawing attention to the necessity of this condition by only ever considering agents for whom it is met: he tells us to "consider all the *good-hearted* people who have ever been in a supermarket, noticed sugar on the floor, and been ready to say 'I am making a mess'", or "all the *responsible* professors who have ever uttered 'The department meeting is starting now.'"[53] These good-hearted, responsible folk are people for whom there is no gap between realizing that one is making a mess and rearranging the torn sack in one's cart,[54] or between realizing that the department meeting starts *now* and moving towards the room where it is held.[55] Perry conceives the problem of the essential indexical as a problem about how to characterize its meaning, or failing that, about how to characterize what it is to have a belief the

---

[52] I will quote from John Perry, "The Problem of the Essential Indexical," *Noûs*, Vol. 13, No. 1 (Mar., 1979), pp. 3-21.
[53] PEI, 17; my italics.
[54] PEI, 3.
[55] PEI, 4.

linguistic expression of which would necessarily (hence essentially) contain an indexical. But what the essential indexical is essential *to* is an explanation of certain bits of behavior that could not be explained in the right way without it. And only certain sorts of people—people who affirm essentially of themselves that they are good-hearted, responsible, or whatever else is relevant—can mean or believe that to which an indexical is essential in this sense, since only such people would behave in the way Perry aspires to explain. Precisely because they are not subject to an ought-ought gap relevant to their situations, the question does not arise for them whether or not to do what their situations require of them; they are little heroes of the sort that an agent who is subject to an ought-ought gap wants, but does not know how, to be. What she wants is to act as they do—to have the facts *determine* her, say, to return the money. She knows how to explain that the facts ought so to determine her, but her difficulty is that those facts nonetheless do not, and that is because she fails in some relevant way to affirm first-personally that she is a heroine of the right sort (that is, the sort of person described by the explanations she can give) on the basis of a genuine appreciation of the truth of this claim.

I just said that our unfortunate agent's fundamental problem is that she fails "in some relevant way" to affirm her own goodness first-personally. But it would be helpful to clarify the nature of this relevant way of affirming in which she fails to affirm her own goodness. Now, we have already noticed that it is a failure to draw a certain practical conclusion on the basis of reasons that she herself recognizes as obliging it. This is perhaps a practical analogue of a difficulty people often have when they do logical or geometric proofs, in which they recognize *in some sense* that each step follows from the previous ones, but fail somehow to feel appropriately confident of the conclusions. And something like this theoretical version of the ought-ought gap is the difficulty that Achilles faces in Lewis Carroll's famous dialogue between that hero and the

Tortoise.[56] Achilles recognizes that *p* and *if p then q* justify the conclusion *q*, but he is prevented from affirming *q* on their basis. This is because the Tortoise dupes him into thinking that he has always to write down, as another premise, his understanding of the justificatory relation between the premises he has already affirmed and the conclusion *q*, before he can conclude *q*. But no rational transitions (theoretical or practical) from reasons to a conclusion could ever occur if someone's understanding of such relations could only ever manifest itself in the form of a premise, a thought particularly affirmed. What an enormous and encyclopedic brain! It is of course true that the sort of understanding the Tortoise and Achilles indefinitely articulate must somehow be operative in a rational transition—but not as any collection of premises, for as premises this understanding will be no help. What is needed is therefore a way of affirming that *p* and *if p then q* together entail *q* that is enough to move you to conclude *q*. And that is the moral of Lewis Carroll's dialogue: that any account of rational transitions must make room for a way for an agent to affirm certain of her commitments concerning the relations between reasons and what they justify that is *sufficient* to move her to judge or act in accord with those commitments, but that does not involve merely adding the contents of those commitments to her catalog of established premises that count as reasons. There must always be reasons affirmed in the usual way and a way of taking them up *affirmed in another way* if there are to be reasons at all in someone's life.

It should be clear that this other way of affirming is exactly what an agent lacks when an ought-ought gap opens up in her life.[57] She may be able to affirm, Tortoise-style, and in (almost?) all relevant respects, that she is a good person, but this affirmation fails to move her.

---

[56] Carroll, Lewis. "What the Tortoise Said to Achilles". Mind, n.s., 4 (1895), pp. 278–80.
[57] I will not pursue the point in detail, but I think this is the correct diagnosis of weakness of the will. The problem of the ought-ought gap is at root the same as the problem of weakness of the will.

She can indeed affirm it as a premise—but that helps her no more than affirming new premises helped Achilles. This is why moral pep-talks are no help. What she needs is to *attune* herself to the concept of a good person, for she needs to be *moved* the way good people are moved. We can put the point that way because the concept of a good person is, perhaps among other things, the concept of a way of taking up the justificatory significance of reasons, recognizing the right reasons to do the right things. This is exactly parallel to how Achilles needs to attune himself to the justificatory significance of *modus ponens* in such a way that he is moved the way good people (good thinkers) are moved, by sufficient reasons for concluding *q* to concluding *q*. So the "relevant way" in which our agent must affirm the thought that she herself is a good person is in this other way, the way of affirming that Lewis Carroll pointed to, which we shall call "attuning" herself to the thought that she herself is a good person.

Being attuned to a thought is having it structure your thinking, in the way that the thought that *p* and *if p then q* entail *q* should have determined the thinking of Achilles. Some (and perhaps all?) of the thoughts to which one can attune oneself require an essential indexical: *I* am a good person; *I* am a good thinker. When one of these thoughts contains an essential indexical in the subject place, I will sometimes speak simply of attuning oneself to the *concept* that figures as its predicate.[58] Thus being attuned to the thought that one is a good person is being attuned to the concept of a good person, and being attuned to the thought that one is a good thinker is being attuned to the concept of a good thinker.

---

[58] I here refrain from any interesting speculations about why this usage comes naturally. But David Lewis is probably on to something relevant in his "Attitudes *de dicto* and *de se*."

5.3   OBJECTIONS AND RESPONSES

I therefore assert that failing to attune oneself to the essentially first-personal thought that one is a good person in a genuine appreciation of its truth is precisely what it consists in to be subject to the ought-ought gap, and this assertion is equivalent to the claim that obviating the ought-ought gap consists precisely in such an attuned affirmation. The goal of this chapter is to characterize how an agent who is subject to the ought-ought gap can bring herself to that sort of affirmation. But before searching for such a characterization, I can anticipate three related objections to the characterization I have already offered. All of these objections grant (perhaps for the sake of argument) that there is some thought that an agent who is subject to the ought-ought gap needs to attune herself to in order to bridge it. But the first of these objections questions whether the content of this thought is that *she herself is a good person*. It seems that many other thoughts that she could attune herself to would do just as well—for example, something like the thought that *returning the money would make her herself more happy than acting otherwise*. This proposal is only one of the potentially many rivals there may be to our claim that she can only bridge the ought-ought gap by attuning herself to the thought that she herself is a good person in a genuine appreciation of its truth, but it is a plausible one, for we do not have to bend over backwards to imagine that beliefs about how she can attain her own happiness could engage her agentive faculties. And indeed it can sometimes seem mysterious how her genuine appreciation of the truth about her own moral status could engage those faculties directly. Let us call this the "objection to goodness."

The second objection takes exception to my use of the notion of *truth* in the formulation I have provided. That is, this objection will dispute my claim, implicit in my language of genuine appreciation of the truth, that our troubled agent is indeed a good person. It does not seem to

116

have to be true that an agent who is subject to an ought-ought gap is a good person—quite the contrary, since she is precisely not determined to do what she ought to do by the reasons there are to do it. If anything, that is what a bad person looks like, or at least a person who is not good. So why should we think it is true at all, rather than simply something it would be helpful for her to think? Let us call this the "objection to truth."

The third objection I anticipate resists my use of the notion of *genuine appreciation*. I say that our agent cannot bridge the ought-ought gap unless she attunes herself to the thought that she herself is a good person not only *sincerely*, but also in a *genuine appreciation* of its truth. This is a wordy way of saying that her attuned affirmation of this thought must be a bit of knowledge. But even if the objection to truth could somehow be dispelled, it might be claimed that sincere affirmation of this claim is all that bridging the ought-ought gap really requires. Sincere attuned affirmation of the thought that she herself is a good person seems to be something less than genuine appreciation (knowledge) of a truth—something less, moreover, that would do the trick without entailing (as does the claim that she can know it) the controversial claim that she is indeed a good person. Let us call this the "objection to knowledge."

While I respond to these three objections, I ask the reader to hold it fixed in her mind that all of them concern the nature of the affirmation, whatever it might be, that would constitute bridging the ought-ought gap. The question posed by the objection to goodness concerns the content such an affirmation must have; the question posed by the objection to truth is whether (the content of) such an affirmation must be true; and the question posed by the objection to knowledge is whether such an affirmation must count as a bit of knowledge. The sort of necessity marked by the word "must" in these three questions is to be understood in terms of what is necessary in order for the agent to bridge the ought-ought gap.

117

I shall tease my response to the objection to goodness out of some rather freewheeling remarks about Kant's conception of morality. Let us take a central claim from Kant—that an action has moral worth only if it is performed from respect for the moral law—as nothing but his way of affirming the claim we have already embraced, that an action has moral worth only if the relation in which its agent stands to the reasons on which she acts is inconsistent with the possibility of her being in the ought-ought gap with respect to them. We can pull this interpretive sleight-of-hand if we loosely understand "the moral law" to refer to the force of the reasons that there are to do what one ought to do (for example, to return the money). Now, Kant glosses respect for the moral law as a "feeling" that is "self-produced by means of a rational concept," namely "the representation of a worth that thwarts my self-love."[59] And this is clearly just to say that respect for the moral law involves giving no power to determine what one does to any personal inclination one might have to act otherwise than as the relevant reasons are reasons to act. This will mean that if our agent were to act from respect for the moral law, she would represent a worth that thwarts any desire she might have not to return the money, and would be moved by that representation (and without the help of anything else) not to act on any such desire. The question, now, is whether this way of relating to the various reasons there are to act in various ways involves affirming that one is a good person. In an obvious way, the answer is no: one's *self*-love must be thwarted in order to have bridged the ought-ought gap. But in another way, the answer is yes, since the representation of a worth that thwarts one's self-love is also the representation of oneself as the sort of being that is capable of being moved to act in a way that is not inevitably determined by one's self-love—that is, it is the representation of oneself as someone who is moved by the reasons that there are to do what any good person

[59] *Groundwork*, 401n1.

would do simply because they are reasons to do what any good person would do. It is in this sense that *any* way of conceiving one's relation to one's reasons in which it consists that one is free of the ought-ought gap is a way, perhaps among other things, of affirming that one is a good person.

Kant happens to (seem to) maintain that there is only one such appropriate way of being related to the reasons to which a good person is responsive, which consists in being motivated by mere respect for the moral law. I prefer to say (not here prejudging whether Kant was right to insist, if he really did, that there is only one such motivation) that any way of being moved to act as a reason requires must be respect*ful* of that reason, rather than that it be *mere respect itself*. This leaves open the possibility that *respect* is a determinable having many determinations (*courage*, *benevolence*, *justice*, and so on), and can never be a property of an agent except insofar as one of its determinations is a property of her—just as *being colored* is a determinable having many determinations (*being crimson*, *being scarlet*, *being chartreuse*, and so on), and can never be a property of a visible thing except insofar as one of its determinations is a property of that thing.

My response to the objection to truth is considerably simpler than this excursion into a Kantian framework. For if the considerations raised by that excursion are correct, then the thought in the attuned affirmation of which it consists that an agent bridges the ought-ought gap is the thought that she herself is a good person. But to have bridged the ought-ought gap *is* to be a good person, since it is to be the sort of person who is moved in the right way by the right reasons to do the right things. So anyone who affirms that she herself is a good person in a way that constitutes bridging the ought-ought gap (that is, in an attuned way) thereby must affirm something true.

119

And this points us in the direction of a response to the objection to knowledge. For we have just seen that anyone who affirms that she is a good person in a way that constitutes bridging the ought-ought gap thereby *must* affirm something true. This is therefore a kind of affirmation that achieves its own truth, since affirming that one is a good person achieves the bridging of the ought-ought gap in which being a good person consists. It therefore *seems* to possess the formal marks of the phenomenon discovered by Anscombe and taken up by subsequent philosophy under the label "knowledge in intention"—it is a kind of knowledge that is what accounts for (causes) the truth of what it knows.[60] I hesitate to draw this conclusion, however, since an attuned affirmation that one is a good person is or involves knowledge that accounts for the truth of what it knows in a noteworthy and special sense. For anyone who does whatever an agent has to do with her concepts merely to count as intentionally going about town shopping for butter (let us call it *efficaciously affirming* her own shopping for butter)[61] knows about her *shopping for butter* in a way that generally counts as knowledge about some butter-shopping. But she *might* be mistaken, since in some particular case her efforts might misfire, and in such a case her claim to have known about some butter-shopping in her efficacious affirmation will have evaporated. That is a wordy way of saying that in the realm of efficacious

---

[60] See *Intention* §32. This knowledge must account for this truth *in the right way* in order to be what Anscombe discovered. For imagine: I know that several white roses are about to be painted red, but I don't know how—perhaps a friend of mine has told me that these roses are likely to be painted red in some way or other, but that if they are not red by 5:00, then she will paint them red herself. For reasons of my own, I would very much like these roses to remain white, and I would never intentionally participate in their reddening, but I am resigned to what I know. Then, via some deviant causal chain of which I am completely unaware, my knowledge that the roses are about to be painted red causes them to be painted red. This Gettier case can probably be used to illustrate many distinct points, but the point I wish to draw from it is that the efficacy of the kind of knowledge Anscombe discovered has to rest not just on its being a bit of knowledge that has appropriate causal effects, but on its being a *special kind of knowledge*. The effects that make it knowledge of this special kind have to be effects that it has *in a certain non-accidental way*, a way in which other kinds of (attempted) knowledge simply do not have effects.

[61] "Efficacious affirmation" here means an affirmation such that, if it counted as knowledge, it would count as a bit of the special kind of knowledge Anscombe isolated. When an efficacious affirmation fails to count as knowledge, we may think of it as whatever an agent would have to have done with her concepts in order to count as unsuccessfully trying.

affirmation there is something like false appearance—merely trying that seems like confident doing—and *efficacious affirmation* is the genus that includes both cases of knowledge and also cases of this sort of false appearance. Someone who attunes herself to the thought that she herself is a good person in the way we have discerned as necessary for bridging the ought-ought gap, however, *cannot* be mistaken in this way. We can put this by saying that *attuning oneself to the thought that one is oneself a good person* in the way required to obviate the ought-ought gap (that is, in an attuned way) is not a process with completion conditions, as *efficaciously affirming to do A* usually is, and this is because its content is already completely actualized as soon as it is efficaciously affirmed.[62] In this way, the attuned affirmation that one is oneself a good person is a kind of practical analogue of the *cogito*, or perhaps merely a way of specifying some of its practical content.[63]

5.4 AVOIDING A KANTIAN ERROR

We can advance our argument at this point by considering whether it follows that the thought that one is oneself a good person is a kind of *tautology*. Perhaps it is, for any attuned affirmation of it will be an affirmation of something true, and it is affirmed in an attuned way in every good piece of practical thought—or in other words, in every piece of practical thought in which a good person can as such engage. Its attuned affirmation must therefore be a bit of knowledge *a priori*,

---

[62] We are here very close to Kant's point that the good will does not derive its value from anything outside of itself. "[I]f with the greatest effort it should yet achieve nothing, and only the good will should remain (not, to be sure, as a mere wish but as the summoning of all the means in our power), yet would it, like a jewel, still shine by its own light as something which has full value in itself' (*Groundwork*, Ak. 394).

[63] Descartes, Meditation II. "[T]he proposition *I am, I exist* is necessarily true whenever it is put forward by me or conceived in my mind."

since its credentials as knowledge do not have to rest on any empirically given information, but can be derived entirely from the agent's own knowledge that she is someone capable of engaging in good practical thought. But we must wonder whether it has another of the marks many philosophers regard as essential for a tautology—that of having *no content*.[64] And there are at least two things to mean by the claim that a proposition has no content. First, one might mean that the relation it affirms between the concepts it involves is analytic—true simply in virtue of something like the essences of the two concepts *I myself* and *good person*. And it should be clear that the thought that one is oneself a good person is not analytic, since it is true not only in virtue of some relation between these concepts—some relation that obtains whether or not one affirms it—but necessarily also in virtue of one's attuned affirmation of it. This is roughly why Kant asserts that the categorical imperative is synthetic.[65] (I will come back to this point.)

But Kant seems to think that the thought that one is oneself a good person is devoid of content in another sense, since he thinks that the resources one needs to attune oneself to the concept of a good person are in principle available to anyone as long as they have the concept of mere practical rationality in general. And so Kant thinks that this concept of mere practical rationality in general is what we must attune ourselves to in order to obviate the ought-ought gap. Kant derives this conclusion from the fact that we can know *a priori* not only of humans, but also of any practically rational beings at all, that they are morally obliged (say) not to tell a lie.[66] He takes this fact to entail that all of our knowledge of morality derives from the mere concept of practical rationality in general; but this conclusion does not follow. What follows is that the access one has to that concept the attunement to which constitutes obviating the ought-ought gap

---

[64] I do not mean to imply that I have gone through a complete catalog of all the criteria for counting as a tautology.
[65] *Groundwork*, Ak. 420.
[66] *Groundwork*, Ak. 408.

has to be access, *perhaps among other things*, to the concept of mere practical rationality in general. But this could be the case if it were right to regard the concept of a good person as a determinable to which we could only gain access by attuning ourselves to one of its determinations—much as we may know of the concept *colored object* that *anything* it captures must be an extended object only because of our familiarity with various of its determinations. As far as I know, Kant does not cite good grounds for omitting the "perhaps among other things" qualification that makes room for this possibility. He thinks that the concept of a good person attunement to which constitutes obviating the ought-ought gap must be available for us to attune to it in the necessary way in isolation from our access to any of its species, and he is simply dead-set against the idea that we might only be able to arrive at a concept of *mere* practical reason, a concept of something that we can share with other possible creatures, by abstracting it away from a more robust concept to which we must first be attuned and of which it is but an essential aspect.

This prejudice of Kant's makes itself felt in his discussion of his claim, which we do not here dispute, that the only maxims that are not allowed by the moral law (which means: the only ways of acting such that there are no reasons to which a good person would respond by acting in one of those ways) are those that somehow involve a contradiction. For there are only two sorts of contradiction that Kant distinguishes, from which it follows that there are only two tests that any maxim we propose to ourselves must pass in order for us to know that it is something a good person could do. The first is what has come to be called the "contradiction in conception" test, in which the proposed maxim somehow contradicts itself. The second is what has come to be

called the "contradiction in the will" test, in which the proposed maxim does not contradict *itself*, but rather contradicts something else that any practically rational being necessarily wills.[67]

We are not interested in contesting Kant's claim that no maxim can be consistent with the will of a good person unless it passes each of these tests. But we suggest the possibility of yet a third test, let us call it the "contradiction in determination" test, which a maxim that an agent proposes to herself fails if it contradicts not *itself*, nor something else necessarily willed by *any* rational being, but something else necessarily willed by any rational being that bears *her* determination of the determinable concept *practical rationality*. Kant cannot countenance the possibility of a contradiction in determination test because he does not countenance the possibility that the concept attunement to which constitutes the good will might always have to be a determinable of the concept *good person*, rather than the concept *good person* itself.

Accepting this amendment to Kant's view seems to be one (and perhaps the only?) way to synthesize broadly Kantian and broadly Aristotelian theories of ethics. For if this proposed amendment is correct, then obviating the ought-ought gap requires attuning oneself not merely to the concept of a good person, but to some particular determination of this concept, and naturally the first such determination for which human philosophers like us are likely to reach is the concept of a *good human being*. And since what is essentially practical about the specifically human form of rationality is contained in the concept of the kind of happiness of which human beings in particular are essentially capable—for the concept *humanity* is, perhaps among other

---

[67] I hope that these proposed emendations of Kant are in line with something that Philippa Foot says (*Natural Goodness*, p. 14):

> Kant was perfectly right in saying that moral goodness was goodness of the will; the idea of practical rationality is throughout a concept of this kind. He seems to have gone wrong, however, in thinking that an abstract idea of practical reason applicable to rational beings as such could take us all the way to anything like our own moral code. For the evaluation of human action depends also on essential features of specifically human life.

things, the concept of a sort of success to which the concept of human happiness is essentially bound—a human has a good will, on this view, if she acts on the basis of an attuned affirmation of this sort of happiness as a possibility for her.[68] For this sort of happiness supplies content to the notion of practical excellence relevant to us: when a good human acts well for the right reasons, she does so because that is the way a happy human acts.

We may well wonder whether this mere notion of a good, or happy, human being is the only determination of the concept of a good person to which we can attune ourselves, or whether we can only ever be attuned to this concept by being attuned to one of *its* determinations. Here, for example, we may think of an opposed pair of determinations of this concept such as the pair *good (human) man* and *good (human) woman*. Perhaps it is possible to attune oneself to one of these concepts, and to count as attuned to the concept *good human* only thereby—just as (if we are right to oppose Kant on this point) it is possible to attune oneself to the concept *good human*, and to count as attuned to the concept *good person* only thereby.

Having gone this far, we might wonder further about such opposed pairs of concepts as *good Athenian* and *good Spartan*, species of the concept *good human* that are not, at least not obviously or in the same way, determinations of it. The question to ask about these concepts is simply whether they are concepts to which one can be attuned. For if one can be attuned to the concept *good Athenian*, then that concept can play the sort of role Lewis Carroll demonstrated that *some* concept or other must play in any rational transition from premises to conclusion. From the premise *Athens will be destroyed if this ship sails*, together with the premise *this ship will sail if and only if I raise this flag*, an Athenian and a Spartan will draw very different

---

[68] There cannot be a kind of rational being such that there is nothing that counts for them as happiness, as long as we understand "S is happy" to mean "S actualizes the particular determination of the concept of a good person to which she is such as to be attuned."

practical conclusions. We *can* analyze what happens in such cases by attributing a few more premises to each—to the Athenian that *I want to save Athens* and to the Spartan that *I want to destroy Athens*—but as Lewis Carroll demonstrated, the attribution of premises will never be enough to explain the fact that a rational transition (the drawing of a conclusion) occurs. In the case of practical rational transitions, we will also have to attribute some appropriate self-conception to which the agent *is* attuned.[69] There does not seem to be any good reason to deny that *good Athenian* or *good Spartan* can play this role.

It may be objected that allowing concepts like *good Athenian* or *good Spartan* to play the role of a species of the concept *good person* attunement to which enables someone to be attuned to that more general concept will undermine the claim of ethics to be an *a priori* discipline. But this objection merely presupposes what we have claimed is Kant's prejudice against the possibility that access to the concept of a good person can only be acquired through attunement to one of its more specific determinations, or even (as we now imagine) to one of its species that is not even a determination of it. As Kripke pointed out, *a priori* is a way of knowing things, not a self-standing class of knowable things.[70] Usually this way of knowing is characterized negatively—one knows *a priori* what one knows not on the basis of experience—but a positive characterization of the basis on which one does know what one knows *a priori* is in the offing if we imagine that one knows *a priori* what one knows exclusively on the basis of the activity of one's faculty of judgment. If this is right, then it is at least possible for the class of things that one rational being can know *a priori* to be different from the class of things that another rational

---

[69] This point is very likely the point Anscombe was making when she said that "[t]he role of 'wanting' in the practical syllogism is quite different from that of a premise. It is that whatever is described in the proposition that is the starting-point of the argument must be wanted in order for the reasoning to lead to any action" (*Intention*, §35). To want A in the relevant sense is to conceive yourself as (the sort of person who is) seeking A in such a way as to be moved by facts bearing on how to achieve A; that is, it is to be attuned to the concept of a person who wants A.

[70] *Naming and Necessity*, p 35ff.

being can know *a priori*—just as long as the two rational beings have relevantly different faculties of knowledge. Kant emphasizes the aprioricity of ethics in order to warn against the disillusionment that comes from seeking the highest good in experience.[71] But the point suggested here is not, and does not entail, that our idea of the highest good (or indeed, of any good) derives from experience, which can only ever provide us with premises. Kant himself says that "although all our knowledge begins with experience, it does not follow that it arises from experience."[72] We suggest that particular sorts of experience can condition—but not as premises—the very essence of the faculty of knowledge a particular rational being has, so that what she can know *a priori* on the basis of this faculty of knowledge in abstraction from any particular experience can include facts about the particular sort of life that she has been brought up to live. This is not to say that an Athenian *cannot* change the sort of goodness to which she is attuned, if she moves to Sparta; but this sort of change is not the same as the sort of change involved when one revises a belief.[73] And it is a mistake to think that this means we should look to experience for the concept of the highest good by which to steer our lives. Precisely not: we must cultivate our *a priori* attunement to those species of goodness we have in ourselves as essential possibilities. Only thus can we acquire attunement to the concept of a good person in general.

---

[71] *Groundwork*, Ak. 407.
[72] *Critique of Pure Reason*, B1.
[73] As Wittgenstein says in his discussion of river-bed truths, "But if someone were to say 'So logic too is an empirical science' he would be wrong. Yet this is right: the same proposition may get treated at one time as something to test by experience, at another as a rule of testing." (*On Certainty*, 98).

5.5   ON THE PRACTICAL IMAGINATION

But we shall not press these extremely controversial points about concepts like *good Athenian* and *good Spartan*, and instead shall revert to the less controversial claim that we can attune ourselves to the concept of a good human, or of human happiness, and that only thereby can we attune ourselves to the concept of a good person. If this is so, then it is only on the basis of such an attunement that we can generate the good will in ourselves. The more thoroughly we attune ourselves to the concept of happiness with which we confront the world, the less often we will find ourselves subject to the ought-ought gap. Obviating the ought-ought gap therefore requires us to devote ourselves to strengthening this attunement to the possibility of happiness by meditating on that concept in an exploratory way. Now, we have already seen that the connection between the concept *I myself* and the concept *good* (or as we can now say, *happy*) *person* in the attuned affirmation of which obviating the ought-ought gap consists is not analytic, for this connection does not derive exclusively from the contents of these concepts themselves, but necessarily also from our attuned affirmation of it.[74] So obviating the ought-ought gap requires us not only to deploy our understandings in comprehending sufficiently the content of the concept *good* (or *happy*) *person*, but also to deploy our imaginations in connecting this concept to ourselves. For only by elaborating imaginatively to ourselves what happiness can be like for us, and thereby generating confidence that we are capable of elaborating this concept more fully in response to novel situations, can we come to have happiness (goodness) in ourselves as an indefinitely extendable possibility with which we are capable of overcoming difficulties and realizing the good as it is for us.

---

[74] This is the point to which I promised to return.

I believe that the sort of imaginative activity I have in mind here is quite familiar, although none of the words with which we ordinarily refer to it—"wishing," "hoping," "daydreaming," and the like—is entirely free from distracting connotations or ambiguities. For it seems that insofar as an agent wishes, hopes, or daydreams, she is necessarily idle; and it seems that if she is idle, then she cannot be achieving anything so exalted as making a good person of herself. I will address this point in what follows, but first I need to settle on the best available word. The word "wish" is incorrigibly associated with idleness. The word "hope" perhaps carries with it less of this association than its competitors, but it has the disadvantage that it cannot really be right to say that someone spent a certain period of time hoping—hoping is not an *activity* in the same way that (say) daydreaming is. Also, to hope to become or accomplish something seems to involve having settled on that possibility in a way that excludes others, whereas the concept I am trying to isolate is necessarily supposed to perform the role of opening our minds to possible ways of being good or happy. So I will use the slightly stilted phrase "open-minded daydreaming" for the sort of imagination in question, and trust the reader to bear these many caveats in mind.

Now, I should clarify immediately that the claim advanced here is not of course that the thing to do on each occasion in which the ought-ought gap opens up is to engage in some open-minded daydreaming. The claim is rather that the sort of training in which a good person engages in order to prevent herself from becoming subject to the ought-ought gap necessarily includes a practice of imaginative exploration of the concept of happiness as it is a real possibility for her. And I think the worry I have registered about the idleness of an agent who merely wishes, hopes, or daydreams can be dispelled by reflecting on the phrase "as it is a real possibility for us."

What we must clarify here is the notion of reality this idea is supposed to involve. We may, for example, imagine a young child daydreaming about becoming a superhero, a way for things to turn out that is certainly impossible for her in some ways and possible in others. Perhaps (although Batman is a notable exception) there is no way to be a superhero without having supernatural powers, and perhaps it is physically impossible to have supernatural powers. What we need to understand is the sort of possibility one must recognize in or attribute to what one imagines in order for it to count as daydreaming of the open-minded sort that attunes one to the concept of a good person. For there is such a thing as exploring possibilities for oneself in a way that does not attune one to the concept of a good person—as when one daydreams about a life one does not conceive as possible for one at all, as a way of escaping, through fantasy, from the unpleasantness of one's life. (This sort of activity would be an "opiate," in Marx's usage.)

My response to this problem will develop from a loose Kantian argument in support of the place that I have assigned to the imagination in a good life. Kant certainly maintains that all knowledge of an object requires a faculty of intuition that attunes us to the form of this object, so as to make possible the synthesis that such knowledge consists in. In the case of theoretical knowledge, the object of knowledge is the world, the form of this knowledge is space and time, and the faculty of intuition that attunes us to the form of this kind of knowledge is the imagination. And we may well puzzle about how these different slots line up in the case of practical knowledge. Some of the answers are easy: the object of practical knowledge is the highest good. It is less easy to say what the form of this knowledge will be, but that does not really matter for our present purpose. For whatever form this knowledge takes, there will have to be a practical faculty of intuition that attunes us to it. So either the imagination has a practical role in addition to its theoretical role, or else there are two imaginative faculties—one for each of

130

the two branches of reason. I need not decide this issue in order to maintain that the faculty of open-minded daydreaming is the imaginative faculty that stands to *a priori* practical reason in the relation in which the faculty of imagination stands to *a priori* theoretical reason.

Now, concepts without intuitions are empty, and intuitions without concepts are blind. So, in the practical realm, concepts without open-minded daydreaming are empty. The difference between open-minded daydreaming that does, and so-called daydreaming that does not, attune us to the concept of a good person is that the open-minded daydreaming that does attune us to the concept of a good person is an activity of this intuitive faculty in cooperation with some appropriate concept—the concept of a superhero, for example—whereas merely escapist daydreaming is an attempt to console oneself with the concept of a superhero without any participation of the intuition of one's own power to live well. Someone who merely entertains the *thought*, but without the participation of her faculty of open-minded daydreaming, that she could be a superhero, fails to think anything at all; her concept is empty. Such a person merely enjoys the *feeling* produced in her by the juxtaposition of this concept in her consciousness with the concept of herself, or of her future, without contributing the positive practical intuition of her power to live as she would like, a contribution upon which the contentfulness of any practical thinking depends in order to count as something valid for knowledge.

6.0   CONCLUSION

6.1   FIRST SECTION

If all the arguments of this dissertation are sound, then it has refuted three false views.

The first of these is bifurcationism. This is the view that every exercise of agency is a metaphysically heterogeneous composite of an inner element or aspect with an outer element or aspect sufficiently distinct from it that a priority relation can obtain between the value possessed by one of these elements or aspects and the value possessed by the other.

The second view refuted by this dissertation is pre-critical empiricism about practical options.   This is the view that every rational practical transaction with the world is a metaphysically heterogeneous composite of a given set of practical options in the having of which the agent's cognition is not active with an act of option-selection in which alone the agent's practical cognitive faculties do any work.

The third view refuted by this dissertation is what I call the "command picture of morality."   This is a picture on which an ethically excellent exercise of agency is a metaphysically heterogeneous composite of the raw applicability to the agent of a system of rules, commands or requirements with an exercise of agency from the agent that happens to conform to that system.

I have here characterized these three views in such a way as to highlight that they have in common an affirmation of *fundamental disunity* in various aspects of practical life, so that it must be some sort of accident in need of special explanation when these aspects fit together as they do when all goes well. The crucial question none of these positions can answer is then: *why* are these different elements supposed to fit together in the ways that they do when all goes well?

In the case of bifurcationism, the crucial question takes the form of asking why an inner element—a trying, say—that is supposed to be connected to a certain kind of outer element *is* supposed to be connected to an outer element of that kind rather than of another. That is a way of asking how it can be that an inner element can have a content for the agent definite enough that some outer elements do, while others do not, count as *what* it is a trying to do. Content marks a certain sort of unity of what has it with what it is about that must precede (because it explains) the possibility of error.

In the case of pre-critical empiricism, our crucial question takes the form of asking how it is that a practical option that an agent might select is one in which the agent may recognize the instrumental-justificatory relatedness among its various doables that makes it a practical option for her at all. For a practical option must have option-content in order to be one option rather than another, and this sort of instrumental-justificatory relatedness is a condition for the sort of unity constitutive of that content. Having a practical option therefore consists in the recognition of at least the potential for this sort of instrumental-justificatory relatedness, and therefore consists in a bit of work on the part of the agent's faculty for justifying what she does in this way (perhaps among others). And in general, whenever a bit of content appears in the life of a justifying being, that content marks at least the possibility of a certain sort of unity that is the work of that being's faculty for justifying.

In the case of the command picture of morality, our crucial question takes the form of asking why the system of rules, commands or requirements in which morality supposedly consists applies to agents in the first place. Now typically we say that someone should do something if and only if there is some good reason why it should be done, and a devotee of the command picture of morality can pretend to honor this fact by positing good reasons for action, the force of which is manifested in morality's rules, commands or requirements, and with which we therefore need not concern ourselves. But (as the example of Kant's shopkeeper illustrates) it is plain that a system of rules, commands or requirements is such that conformity of one's conduct to that system may express responsiveness to any of many distinct reasons, so there can be no fact of the matter in a bit of excellent conduct as the command picture conceives it sufficient to settle whether that conduct manifests responsiveness to the (right) reasons. There is therefore no way of getting around the fact that in order for our conduct to manifest responsiveness to the right reasons, we must somehow attend to those reasons. The difficulty into which bifurcationism fell here repeats itself in another context: just as the content of an inner element marks a kind of unity of what has it with what it is about, so too the fact that one should act a certain way for a certain reason marks a kind of unity of this way of acting with that reason—a unity that cannot be generated except somehow by the agent's recognition of it. To be a rational agent is (perhaps among other things) to be a self-recognizing nexus of this kind of unity.

6.2   SECOND SECTION

There are then three kinds of unity that the views against which this dissertation argues cannot countenance: the unity of an inner element of an exercise of agency with its outer element, the unity of doables into a practical option, and the unity of a reason to act with the way of acting that it is a reason for.   How, then, are these various sorts of unity to be understood?   If the appropriation of Kant that I attempted in my fourth chapter is on the right track, then at least the last of these unities is to be understood as the work of the agent's (practical) imagination.   In fact, I think the imagination is the key to understanding all three of these unities, and I would like to make some remarks here about what that does and does not mean.

Very roughly speaking, the imagination is the faculty for realizing what our concepts do not yet contain.[75] This characterization of the imagination may very well lead us to suspect that the Kantian strategy of attributing various sorts of unity to the work of the imagination is simply an *ad hoc* regress-stopper.   For if our puzzle is how a practical reason comes to be unified with an appropriate action across the self-recognizing nexus of an agent, then it seems that no great light is cast upon our puzzle if we are told that this transition is effected by a faculty for realizing what our concepts do not yet contain.   I will call this "Nietzsche's objection," since it seems to have been what he had in mind when he made these famous remarks:

> "How are synthetic judgments *a priori possible?"* Kant asked himself—and what really is his answer?   *"By virtue of a faculty"*—but unfortunately not in five words, but so circumstantially, venerably, and with such a display of German profundity and curlicues that people simply failed to note the comical *niaiserie allemande* involved in such an answer….

---

[75] The word "realizing" is here chosen for its ambiguity between a sense in which it means something purportedly passive such as to *appreciate* and another sense in which it means something purportedly active such as to *make real*. The point of choosing a word with this ambiguity is to combat the temptation to insist on this distinction.

> The honeymoon of German philosophy arrived. All the young theologians of the Tübingen seminary went into the bushes—all looking for "faculties." … "By virtue of a faculty"—he had said, or at least meant. But is that—an answer? An explanation? Or is it not rather merely a repetition of the question…?
>
> [I]t is high time to replace the Kantian question, "How are synthetic judgments *a priori* possible?" by another question, "Why is belief in such judgments *necessary?*"[76]

I think it is helpful here to focus on the historical fact that the unities Kant himself struggled to understand—quintessentially, the unity in which causes are bound up with their effects—first emerged as puzzles for him when he encountered Hume's skeptical arguments for the conclusion that the role of these unities in our thinking "are discoverable, not by reason, but by experience."[77] For in his way Hume attributed these unities to the work of the imagination—the faculty for realizing what our concepts do not yet contain—just as much as Kant did. To that extent, they agree. What they disagree about is whether this work that we do with the imagination is work that we have the right to do, for Hume denied that we have the right to do any cognitive work that reaches beyond what our concepts already contain together with whatever is given.

Kant's argument against Hume on this point can be summed up in the slogan that the work of the imagination is a condition for the possibility of *any* empirical knowledge whatsoever—including the sort of knowledge that Hume wants to take as given. The argument of my first chapter can be adapted to make this point: an *impression* of, say, a certain shade of red must have that shade of red as its *content*; and there cannot be such a thing as a content of a certain sort in the life of a being who is in principle never in a position to ascertain whether the world actually is as that sort of content portrays it. And if such acts of ascertaining are possible, then surely they are what we ought to take as our paradigm of informational commerce with the

---

[76] *Beyond Good and Evil*, §11

[77] Hume, *An Inquiry Concerning Human Understanding*, p. 42. Kant "confesses" that it was Hume's skeptical arguments that woke him from his dogmatic slumbers in the *Prolegomena* (260).

world. In these acts our cognition is unified through its content with what it is cognition of. And since this unity too is the work of the imagination, Hume's principle that we have no right to any cognitive work that reaches beyond what our concepts already contain will oblige us to disavow any right to do the work in which it consists to have the impressions with which Hume himself must hope to be able to equip us if we are really to be in his skeptical predicament. The position of the Humean skeptic is therefore implicitly inconsistent: it is a skeptical predicament that denies that we have a right to do something that we would have to have done in order to end up in it.[78]

Kant conceives a bit of knowledge, among other things, as a judgment that we somehow have a *right* to make. So at least one thing we can say on his behalf in response to Nietzsche is that part of what Kant is asking when he asks how synthetic a priori knowledge is possible is how we may have the appropriate kind of right. And his answer to this question is that we may have a right to make synthetic a priori judgments because the possibility of synthetic a priori knowledge is *necessary* for our (rational) form of life. So Kant is not indifferent to the question—why is synthetic a priori knowledge necessary?—that Nietzsche urges us to ask.

Now here is how these points help us to respond to the heart of Nietzsche's famous objection. It is not that we first come to recognize various ways in which our form of life would be impossible if we did not have a faculty for realizing what is not yet contained in our concepts, and then we beat the psychic bushes and, lo and behold, we find just such a faculty in the imagination. Rather, we simply come to recognize various ways in which our form of life would be impossible if we did not have a faculty for realizing what is not yet contained in our concepts. That recognition *alone*—and not any independent act of finding such a faculty—is what puts us

---

[78] As I have said before, I derive this argument from McDowell, "Knowledge and the Internal."

in touch with that faculty and secures our right to it. So we do not search for and find exactly the faculty what we need to stop various regresses, and about which it is simply *ad hoc* to deny that we may ask how it enters into appropriate relations with the materials that we already had in view; rather, we recognize that our form of life is one in which these regresses do not start.

Perhaps a clearer way of putting this point would be to say that the unities with which this dissertation has been concerned are not such that we can get them sufficiently in view to seek a sort of explanation of them that refers to something *else*. We cannot need to explain how, say, a collection of doables enters into the unity of a task as if we could already understand what doables are without yet understanding what a task is, so that we must look at something *else*— the unity of doables—in order to understand the unity that puzzles us. For understanding the unity of a doable is already understanding the unity of a task: *A in order to do B* is just as much a doable as is A or B.

6.3   THIRD SECTION

Now I would like to make some remarks about what we may hope the positive results of this dissertation will help us to understand.

Part of what we might hope to accomplish by recognizing that ascertaining what our practical options are is a bit of work rather than something merely given is a sensible way of working out the ancient idea of virtues as quasi-perceptive faculties. For if ascertaining what the options are is already a part of the job of living an ethical life, then it follows that an investigation of the excellence proper to that bit of the job will be an investigation of an at least quasi-perceptive faculty. And since the branch of an agent's rational economy into which this

138

will be an investigation is the practical, rather than the theoretical, side of her cognition, we may hope to secure an understanding of how these faculties, although they are quasi-perceptive, can nonetheless be essentially practical.

This hope is also encouraged by our discovery of the role that the practical imagination must play in attuning us to various determinations of the concept of a good person. For what this process of attunement *is* is a process of opening oneself up to the possibility that various sorts of considerations can count as practical reasons for one to act in certain ways. The determination of the concept of a good person to which one thereby becomes attuned itself determines relations between at least a range of practical reasons and what they are reasons to do. So the process of identifying oneself with a way of being good is a process of opening oneself up to certain sorts of reasons as reasons to act in certain ways. One opens oneself up to these reasons—so the process involved here is at least quasi-perceptive—but the faculty whereby one does this is a practical faculty, a faculty that aims at action, for this is a context in which the imagination contributes toward the constitution of a bit of knowledge that is productive of what it knows.

6.4   FOURTH SECTION

There is also reason to hope that further investigation of the practical imagination may help us to understand what is involved in the recognition of one rational agent by another. I shall lead up to this point with some remarks about the perspective on rational agency that we have developed so far.

A rational agent is a self-recognizing nexus of the unity of practical reasons with the actions that they are practical reasons to perform. But there is no such thing as a *mere* rational

agent—an agent for whom there are no determinate facts about what can count as a reason for what. Such an agent, if she existed, could kill a hedgehog *entirely* because the Pythagorean Theorem is true, or run for President *entirely* because she was once thirsty. The point of the word "entirely" here is that there is *nothing* to say in response to an inquiry as to why the truth of the Pythagorean Theorem counts as a reason to kill that hedgehog, or as to why her thirst counts as a reason to run for President—those just are her reasons. And I doubt that there can be such a thing as a mere agent because what makes something a reality is the fact that it can be considered from multiple perspectives, whereas to say that there is nothing further to say about such a nexus of unity is to say that there is just one perspective from which it can be viewed.

So every rational agent is an agent of some determinate kind. But if what we have argued is correct, then every agent must have deployed her practical imagination in order to attune herself to this determinate kind of agent that she already is. So the possession and application of the practical imagination is a condition for the possibility of being an agent. And the imagination is the faculty for realizing what our concepts do not yet contain. Every agent therefore both is a determinate kind of agent and also has a faculty whereby she may bring her first-personal concept of herself under the concept of a new kind of agent that she is not already, thereby realizing a unity not yet contained in her concepts. The fact that an agent must possess and apply a practical imagination entails that an agent must be self-determining by being self-recognizing.

So what it consists in for one rational agent to recognize another *as* a rational agent must be more than for her to recognize the determination of this concept that the agent she recognizes already exhibits. This determination must at least be recognized *as* a determination of the concept of a rational agent, which means it must at least be recognized *as* an identity forged by a

140

faculty for realizing unities not yet contained in our concepts. This faculty and the beings who have it always have some determinate shape, but they also always continue to develop from those shapes into new ones. And recognizing this openness to what is not yet contained in one's concepts is what recognizing another rational agent consists in.

I have brought out this point in order to issue a warning against a false conclusion that might seem to follow from it. For the fact that there is something that recognizing another rational agent consists in does not entail that there is such a thing as *mere* recognition of rational agency any more than the fact that there is something that rational agency consists in entails that there is such a thing as *mere* rational agency. Recognizing another rational agent as a rational agent is not recognizing that she has the faculty for realizing what is not yet contained in her concepts, *plus* (as if it were something in principle separable from that first element) some determination of this faculty. So recognizing another rational agent as a rational agent is a matter of *attuning* oneself to the particular determination of rational agency that the recognized being exhibits *as* a determination of rational agency. There is human recognition, Martian recognition, Athenian recognition, and so on, but there is not mere recognition as such.

Recognizing another rational agent as a rational agent requires recognizing that she has a practical imagination, which is a faculty to realize what is not yet contained in our concepts. So recognizing another rational agent as a rational agent requires a recognition of the possibility of what is not yet contained in our concepts. It follows from this that recognition of another rational agent as a rational agent itself requires us to deploy the practical imagination, since in this recognition we must recognize a faculty that reaches beyond what is contained in our concepts—something we can only do by using the faculty whereby we attune ourselves to what is not yet contained in our concepts.

I think that this point about how the practical imagination must be active in our recognition that another being also has the practical imagination may shed light on the function of narrative art forms in human life. To read a novel or to view a movie is, among other things, to explore various ways that life can go. To bring the appropriate amount of sympathy to the characters in a novel or a movie—to *recognize* them as rational agents—consists, perhaps among other things, in bringing one's imagination to the text or the images in an attempt to explore happiness as a possibility for oneself. But it is also to identify with other rational beings, sometimes with rational beings that are radically different from oneself—one can identify with a character of a different gender, nationality, or even biological species—and thereby to attune oneself not only to the particular determination of the concept *good person* that one bears, but also to the various determinables of which this is a determination. By identifying with a character from another *city*, one attunes oneself to the determinable concept *good compatriot*; by identifying with a character from another country, one attunes oneself to the more general determinable concept (say) *good democrat*; by identifying with a character of another gender, one attunes oneself to the more general determinable concept *good human being*; and by identifying with a character of another biological species, one attunes oneself to the mere determinable *good person*. All of this is achieved by the practical faculty of imagination, which explores happiness as it is a real possibility for us.

6.5   FIFTH SECTION

I would like to close with the grandiose suggestion that we may think of the activity of the practical imagination in open-minded daydreaming as a Hegelian successor of the concept of

prayer. The purpose of prayer for those who have not gone through Hegel's sublation of religion is to strengthen the connection between the individual and God. But anyone who has gone through Hegel's sublation of religion understands that this God is not something different from ourselves, but that it is rather the idea of our own fulfillment always implicit in all of our own activity (happiness). So the idea of a "connection" may be misplaced, since connections are between distinct things. Nonetheless, it is possible to become alienated from this idea in various ways, one of the worst of which is to be subject to the ought-ought gap. The way to prevent and cope with alienation from the concept of happiness (which, in pre-Hegelian language, is the feeling of being unbearably separate from God) is to attend imaginatively to the real prospect of happiness for oneself (that is, to daydream). So daydreaming is Hegelian prayer. For the more securely we grasp the concept of happiness as a real possibility for us (which we do by daydreaming), the less likely we are to misunderstand the point of some activity that plays an important role in it. And this is the way to set aside and obviate doubts about one's own worth, which is to say it is the way to set aside doubts about whether one has the good will. For since the good will has its uniquely unqualified kind of goodness in virtue of its own willing (its wanting of real possibilities), and not in virtue of its results, all one needs to do in order to secure one's confidence that one has a good will is to attend imaginatively to one's wanting of real possibilities.

Seen from this point of view, it may be useful to think of the command picture of morality as a form of "idolatry" that takes as its idol the *ethically excellent exercise of agency* in itself. It should be clear that this quasi-theological term "idolatry" resonates sympathetically not only with a longstanding tradition within analytic philosophy (for example, the concept of rule-fetishism as it figures in discussions of rule-consequentialism and Kantian ethics), but also with

similar usages in other philosophical traditions and contexts (for example, the concept of commodity-fetishism in Marx and the accusation of idolatry that Nietzsche levels against his opponents in *Twilight of the Idols*).[79] Intuitively, to take something as an idol is to try to locate a certain sort of value in a mere thing, in isolation from the vital activity in the context of which alone it has that value. It can perhaps be difficult to see how the ethically excellent exercise of agency can occupy the position of an idol, since it is of course a vital activity of a certain sort. But in the relevant sense of "mere thing," *anything* (including a vital activity) can be a mere thing; it simply depends on whether we see the vitality in it. And we fail to see the vitality of an ethically excellent exercise of agency if we regard its value as strictly independent of the reasons that motivate the agent to perform it, and of which it is therefore an expression. The deployment of our practical imaginations in open-minded daydreaming is the way to catch sight of that vitality again.

---

[79] The best philosophical treatment of idolatry as such that I have found is in the first chapter of Jean-Luc Marion's quasi-theological work *God Without Being* (Chicago: University of Chicago Press). Very roughly speaking, Marion defines an idol as whatever fills our gaze with what we already understand ourselves to be seeking, in contrast to an icon, which he defines as whatever overpowers our gaze with a recognition of the possibility of that which surpasses our understanding. Marion works out this distinction in a theological context in which I am not particularly interested, but I think the distinction itself is well worth philosophical attention, since the more comprehensive and true a philosophy is, the greater the danger that it will leave its adherents without a locus of wonder in which an icon, as opposed to an idol, can manifest itself. So the success of philosophy, which at its best springs from wonder rather than from anxiety, is a potentially great danger to what prompts and sustains its vitality—the openness to what our concepts do not yet contain.

BIBLIOGRAPHY

Anscombe, G. E. M. 1963: <u>Intention</u>; second edition. Cambridge: Cambridge University Press.

----------1958: Modern Moral Philosophy, in Cahn and Haber (eds.) <u>20th Century Ethical Theory</u>. Upper Saddle River, NJ: Prentice-Hall: 351-364.

Aristotle: <u>Nicomachean Ethics</u>, T.H. Irwin, transl. Indianapolis: Hackett Publishing Company.

Austin, J.L. 1956: A Plea for Excuses, in his 1961.

----------1961: <u>Philosophical Papers</u>. New York: Oxford University Press.

Bratman, M. Acting with an Intention, in his 1987: 128-138.

----------1999: <u>Faces of Intention</u>. New York: Cambridge University Press.

----------1987: <u>Intention, Plans and Practical Reason</u>. Cambridge: Harvard University Press.

----------1979: Castañeda's Theory of Thought and Action, in his 1999: 225-249.

Carroll, L. 1895: What the Tortoise Said to Achilles, <u>Mind</u>, n.s. (4): 278-280.

Castañeda, H. N., 1980: The Theory of Questions, Epistemic Powers, and the Indexical Theory of Knowledge, in P. French et al., eds., <u>Midwest Studies in Philosophy V</u>. Minneapolis: University of Minnesota Press: 193-237.

Davidson, D. 1963: Actions, reasons, and causes, in his 1980: 3-19.

----------1971: Agency, in his 1980: 43-62.

----------1980: <u>Essays on Actions and Events</u>. New York: Oxford University Press.

Descartes, R. 1640: <u>Meditations on First Philosophy</u>, in <u>The Philosophical Writings of Descartes, Vol. II</u>, J Cottingham, R. Stoothoff and D. Murdoch, transl. New York: Cambridge University Press.

145

Diamond, C. 1997: Consequentialism in Modern Moral Philosophy and 'Modern Moral Philosophy', in Oderberg and Laing (eds.), <u>Human Lives: Critical Essays on Consequentialist Bioethics</u>. New York: St. Martin's Press: 13-38.

----------1985: Missing the Adventure: Reply to Martha Nussbaum, in her 1991: 309-318.

----------1991: <u>The Realistic Spirit</u>. Cambridge: MIT Press.

Epictetus: <u>The Handbook</u>, N.P. White, transl. Indianapolis: Hackett Publishing Company.

Foot, P. 2001: <u>Natural Goodness</u>. New York: Oxford University Press.

Frankena, W. 1963: <u>Ethics</u>. Englewood Cliffs, NJ: Prentice-Hall.

Frege, G. 1884: <u>The Foundations of Arithmetic</u>, J.L. Austin, transl. 1950. Evanston, IL: Northwestern University Press.

Hume D. 1758: <u>An Inquiry Concerning Human Understanding</u>. New York: The Bobbs-Merrill Company, 1955.

Kant, I. 1781: <u>Critique of Pure Reason</u>, N. Kemp Smith, transl. New York: St. Martin's Press, 1929.

----------1785: <u>Grounding for the Metaphysics of Morals</u>, J. W. Ellington, transl., third edition. Indianapolis: Hackett Publishing Company, Inc., 1993.

----------1783: <u>Prolegomena to Any Future Metaphysics</u>, in <u>Philosophy of Material Nature</u>, J.W. Ellington, transl. Indianapolis: Hackett Publishing Company, 1985.

Kavka, G. 1983: The toxin puzzle, <u>Analysis</u> (43): 33-36.

Kripke, S. 1972: <u>Naming and Necessity</u>. Cambridge: Harvard University Press, 1980.

Lewis, D. 1979: Attitudes De Dicto and De Se, <u>The Philosophical Review</u> (88): 513–543.

Marion, J.L. 1982: <u>God Without Being</u>, T.A. Carlson, transl. Chicago: University of Chicago Press.

McDowell, J. 1995: Knowledge and the Internal, in his 1998: 395-413.

----------1998: <u>Meaning, Knowledge & Reality</u>. Cambridge: Harvard University Press.

Mill, J.S. 1863: <u>Utilitarianism</u>. Indianapolis, IN: Hackett Publishing Company.

Nietzsche, F. 1886: <u>Beyond Good and Evil</u>, W. Kaufmann, transl. 1966. New York: Vintage Books.

Perry, J. 1979: The Problem of the Essential Indexical, <u>Noûs</u>, Vol. 13, No. 1: 3-21.

Scheffler, S. 1985: Agent-Centered Restrictions, Rationality and the Virtues, <u>Mind</u>, Vol. 94: 409-419.

Sellars, W. 1957: <u>Empiricism and the Philosophy of Mind</u>, in his 1963: 127-196.

----------1963: <u>Science, Perception and Reality</u>. Atascadero, CA: Ridgeview Publishing Company.

Thomson, J.J. 1985: The Trolley Problem, in <u>Yale Law Journal</u>, Vol. 94, Issue 6: 1395-1415.

Thompson, M. "Knowing What's Up: The Unity of its Empirical and Practical Forms." Anscombe's *Intention*. University of Chicago. Chicago, April 24, 2009.

Wittgenstein, L. 1951: <u>On Certainty</u>, D. Paul and G. E. M. Anscombe, transl. Oxford: Blackwell Publishers Ltd.,1969.

----------1953: <u>Philosophical Investigations</u>, G. E. M. Anscombe, transl. Malden, MA: Blackwell Publishers.