

Trying without fail*

Ben Holguín & Harvey Lederman

February 2023

Abstract

An action is *agentially perfect* if and only if, if a person tries to perform it, they succeed, and, if a person performs it, they try to. We argue that trying itself is agentially perfect: if a person tries to try to do something, they try to do it; and, if a person tries to do something, they try to try to do it. We show how this claim sheds new light on the logical structure of intentional action, on the question of whether basic actions are tryings, and on the notion of “options” in decision theory. On the way to these central ideas, we argue that a person can try to do something even if they believe it is impossible that they will succeed, that a person can try to do something even if they do not want to succeed, and that a person can try to do something even if they do not intend to succeed.

1 Introduction

There is a vision of the divine that centers on a perfect correspondence between divine act and divine will. The divinity, it is said, does everything it wills, and wills everything it does. In a folksier manner of speaking, the divinity does everything it tries to do, and tries to do everything it does.

We mere mortals, of course, are not like this at all. We do not do everything we try to do. A tired parent tries to stay calm, but fails when they fly into a rage instead. An addict tries to get high, but fails because the dealer sold them poison, not their favored drug. A pacifist tries to avert a revolution, but fails because their speech instead incenses the crowd. And, as these examples show, we also do not try to do everything we do. The parent flies into a rage, the addict poisons themselves, and the pacifist starts a war—although none of them was trying to do what they did.

In general, then, the expression of the human will in action can be distorted in ways that make us quite unlike the divine. But is there a domain which is safe from such misfires—a domain in which the divine spark in our human will is guaranteed to shine through?

*Thanks to David Barnett, Bob Beddor, Kyle Blumberg, Mike Caie, Sam Carter, Imogen Dickie, Adam Elga, Davide Fassio, Anton Ford, Jane Friedman, Eric Funkhouser, Jessica Gelber, Simon Goldstein, Caspar Hare, John Hawthorne, Nico Kirk-Giannini, Hongwoo Kwong, Jack Mullaney, Jennifer Nagel, Carlotta Pavese, Antonia Peacocke, Bryan Reece, Gideon Rosen, Kieran Setiya, Michael Smith, Robert Stalnaker, Eric Swanson, Sergio Tenenbaum, and Timothy Williamson for conversations, correspondence and written comments. Friedman, Goldstein, Rosen and Williamson especially went above and beyond. Thanks also to audiences at the ACU Workshop on Action, Intention and Language, the Asian Epistemology Network, the University of Arkansas, the Ranch Metaphysics Conference, MIT, Princeton, and the University of Toronto.

A central goal of this paper is to argue that there is. We characterize a class of doings that we call *agentially perfect*, doings that (to put it roughly) by their very nature ensure a perfect match between attempt and success. More exactly, we will say that ϕ -ing is *agentially perfect* if and only if: necessarily, if a person tries to ϕ , they ϕ , and necessarily, if a person ϕ s, then they try to ϕ . We then argue that trying itself is agentially perfect: necessarily, if a person tries to ϕ , they try to try to ϕ , and necessarily, if they try to try to ϕ , they try to ϕ .

We go on to develop some consequences of this claim. If trying is agentially perfect, then plausibly trying is essentially intentional in Anscombe's sense, that is, necessarily, a person who tries to ϕ , intentionally tries to ϕ (Anscombe (1957, §47)). In fact, it becomes plausible that if a person tries to ϕ they intentionally intentionally...intentionally try to ϕ . We argue, in line with recent work by Shepherd & Carter (2021), that this leads to a striking divergence between the structure of intentional action on the one hand, and the structure of knowledge on the other. Moreover, we suggest that, if trying is agentially perfect, it opens a new argument for the claim that tryings may play the role of "options" in decision theory (cf. Pollock (2002), Hedden (2012), Hedden (2015)), and also makes plausible the idea that, on a natural way of understanding "basic action", all basic actions are tryings.

As a warmup to these central points, we begin by exploring a slightly different theme. Descartes writes of the will that it is "so free of its nature that it can never be constrained".¹ We defend a related doctrine for trying. A common view is that what a person can try to do is highly constrained by their attitudes toward the success of their attempt: that if a person tries to do something, they must think they can succeed, they must want to succeed, and perhaps they must even intend to succeed. As we will explain, we believe that if these constraints were correct, they would rule out the claim that trying is agentially perfect. So we begin the paper with extended arguments against them, and in favor of a "Cartesian" conception of trying, on which trying is not constrained by such attitudes to success.

Section 2 draws attention to some different uses of 'try' and isolates a use that will be our focus throughout. Sections 3-5 defend the Cartesian conception of trying. We consider and argue against particular connections between trying to ϕ and believing that ϕ -ing is possible (§3), between trying to ϕ and wanting to ϕ (§4), and between trying to ϕ and intending to ϕ (§5). Section 6 isolates the property of agential perfection. Section 7 argues that trying is agentially perfect: necessarily, a person tries to ϕ if and only if they try to try to ϕ . Section 8 presents an argument that trying is essentially intentional, and explores the consequences of this claim for the structure of intentional action. Section 9 connects the agential perfection of trying to the notion of options in decision theory and to the notion of basic action. Section 10 concludes.

¹ *Passions of the Soul* I. XLI: "la volonté est tellement libre de sa nature, qu'elle ne peut jamais être contrainte", Descartes (1996, XI.359), with translation in Descartes (1985-1992, vol 1, p. 343), cited by Albritton (1985).

2 On ‘trying’

Our arguments will often rely on intuitive judgments about trying. As we elicit these judgments, it will be important both to attend to the distinction between what is true and what is assertable, and to recognize the existence of possibly different contextual resolutions of the word ‘try’. So, in this preliminary section, we begin with some background on these ideas.

Before we come to them, however, we begin with a general methodological point. Throughout the paper, we will approach the topic of trying broadly from the perspective of lexical semantics.² Our interest is in the relation picked out by the ordinary expression ‘try’, as well as its connections to the relations picked out by the ordinary expressions ‘believe’, ‘want’, and ‘intend’, among others.³ We will not try to offer—or even assume that there exists—a reductive analysis of any of these relations. Nor will we make any prior assumptions about their theoretical roles. Our goal is simply to discover (or falsify) general principles about trying (again: what is denoted by ‘try’ in English), by relying on our general competence with the relevant natural language expressions. Our arguments can (and should!) be objected to by adducing counterexamples or by contesting the linguistic hypotheses on which they are based. But they cannot be objected to on the grounds that they conflict with principles stated in terms of theoretically loaded uses of the expressions ‘try’, ‘believe’, ‘intend’, and so on, for, by our lights, appeals to such principles simply change the subject.

With this in mind, we turn to our points about the range of uses of trying. We begin with a claim that seems widely accepted these days, but is still worth stating explicitly: that trying is compatible with foreknowledge of success.⁴ The mere fact that you knew you would succeed in raising your arm is no barrier to your having tried to raise it. We think this is intuitively clear in its own right. But we also think that the compatibility of trying and foreknowledge of success provides a natural way of explaining our knowledge of our abilities (cf. Mandelkern *et al.* (2017)). After all, the basis of our ordinary knowledge that we are able to raise our arms seems precisely to be our knowledge that, if we were to try to raise our arms, we would succeed.

Of course it can be misleading to describe a person who has succeeded in ϕ -ing as having tried to ϕ . We don’t typically say ‘They tried to ϕ ’ when we could instead have just said ‘They ϕ -ed’. But that’s at most weak evidence against the view that trying is compatible with foreknowledge of success. It is typically assumed that we try to do the things we succeed in doing, but not that we succeed in doing the things we try to do. So it is unsurprising that we might interpret ‘S tried to ϕ ’ as implying (without entailing) that they were unsuccessful in ϕ -ing.

² As a consequence, our approach is close in feel to work in semantics by Sharvit (2003), Grano (2011), although our target data will be quite different.

³ One important caveat: throughout the paper we consider only uses of ‘try’ where it takes an infinitival complement, meaning something like ‘attempt’ (e.g., ‘They’re trying to hit the bullseye’). We exclude those uses where ‘try’ takes an object complement, meaning something like ‘test’ (e.g., ‘They’re trying my soup’).

⁴ The *locus classicus* of the view that trying is incompatible with foreknowledge of success is Wittgenstein (1950, §622), but see also Heath (1971) and Jones (1983) for sympathetic discussion. The canonical argument in favor of the compatibility of trying and foreknowledge of success was given by Grice (1989, p. 7) (originally delivered, 1967), but cf. e.g., O’Shaughnessy (1973), Armstrong (1973), Hornsby (1980), McGinn (1982, ch. 8), Schroeder (2001).

Now for a subtler question: is trying compatible with effortlessness? That is: can a person who ϕ s without exerting any amount of physical or mental effort—a person who can ϕ at will, as it were—nonetheless count as having tried to ϕ when they succeed in ϕ -ing?

We think the answer is yes. We see nothing invariably problematic about describing a person who effortlessly succeeds in ϕ -ing as having tried to ϕ . Again, perhaps the most natural way to make sense of our knowledge of our abilities to do certain things effortlessly requires accepting the compatibility of trying and effortlessness. The basis of our ordinary knowledge that we are able to raise our arms effortlessly, for instance, seems precisely to be our knowledge that if we were to try to raise our arms, we would succeed whether or not we put any effort into it. So effortlessness must be compatible with trying, at least in one natural way of using ‘try’.

Why the qualification about ways of using ‘try’ here? Because sometimes we say things like ‘I can beat him in tennis without even trying’. Sometimes we order people to ‘*Actually* try!’. And sometimes we ask ‘Is she going to have to try on this question, or will it be as easy as the last one?’. On the assumption that ‘S is trying to ϕ ’ never entails ‘S is putting effort into ϕ -ing’, it seems hard to make sense of these ways of talking.

What should we say about these apparently different uses of ‘try’? One option is to give a pragmatic explanation of them, which doesn’t require changing our view of the meaning of ‘try’. Another would be to conclude that ‘try’ is ambiguous and/or context-sensitive, and that on certain interpretations (though not all!) it expresses something along the lines of what is expressed by ‘try hard’ or ‘try with sufficient effort’. The examples just given would be ones in which this sort of reading is salient, but that’s compatible with there being readings where (as above) a person can try without any effort.

These aren’t the only kind of examples that may suggest that ‘try’ has a complex array of uses. Suppose you promise your friend that you’ll try to climb Mt. Everest. You train intensely and buy plane tickets to Nepal in advance of your scheduled climb. But then disaster strikes in the form of an international pandemic, and you’re forced to cancel your trip. Obviously, you aren’t someone who has climbed Mt. Everest. But are you at least someone who has *tried* to climb it? On the one hand, if a stranger were to ask you ‘Have you ever tried to climb Mt. Everest?’, it would be pretty misleading to respond ‘Yes I have’. (Imagine how the conversation might evolve: ‘How far up the mountain did you get?’—‘Well, I never actually made it to the mountain, but I did buy plane tickets to Nepal. . .’) On the other hand, if your friend were to ask ‘Did you even *try* to climb Everest?’, it seems like you could speak truly in responding ‘Of course I did—I bought tickets and everything!’.

What to make of these facts? As before, it could be that there’s a pragmatic explanation. But also as before, it could be that for certain complicated, temporally extended actions—e.g., climbing Mt. Everest, writing a novel, starting a new career—whether one counts as having “tried” to perform the action in question is a context-sensitive matter. Perhaps in some contexts it suffices that one tried to perform certain preliminary steps (perhaps with the relevant intentions), while in other contexts one counts as having tried only if one manages to succeed in performing certain “core” actions—e.g., setting foot on

the mountain, typing sentences, or leaving one's current job.

For the purposes of this paper it won't be so important whether 'try' has a simple lexical entry and these alternative "readings" are given a pragmatic explanation, or whether 'try' is genuinely ambiguous or context-sensitive. Our official target is whatever relation (or relations) are expressed by very permissive readings of 'try'—that is, the readings of 'try' on which you count as having tried to climb Mt. Everest in virtue of having bought tickets to Nepal, and where *any* amount of effort could in principle be enough to count as having tried.⁵ It's clear that 'try' has such readings, whether or not it's context-sensitive or ambiguous.

But, while this is our official attitude, there is a stronger hypothesis which we find attractive, and which will help us to simplify the exposition below. The stronger hypothesis is that 'try' is indeed context-sensitive, and, furthermore, that there is a contextual resolution of 'try' that is minimal in the following sense: if for *any* context *c*, 'S is trying to ϕ ' expresses a truth in *c*, then 'S is trying to ϕ ' expresses a truth on the minimal resolution. Call the relation picked out by 'try' on this minimal resolution *pure trying*. According to our hypothesis, all trying entails pure trying. This hypothesis has several nice features. But one reason to attend to it in the present setting is that, if it is correct, then pure trying would have a certain distinguishedness in the semantics of 'try', which would make it apt for philosophical investigation. Of course, it is not in general true that, merely because there is a unique weak resolution of a philosophically interesting context-sensitive term, we should investigate that reading. But in the case of trying, such a weak notion would be a natural target. An important question about the metaphysics of agency is what the minimal conditions are for its exercise. If it exists, our weak reading of 'try', designates the very barest form of trying. Pure trying would be a place where—at least as far as trying is concerned—agency gets its start.

Whatever one thinks about this stronger hypothesis and its consequences, we now ask the reader to understand our uses of 'try' for the rest of the paper as having a permissive reading on which it expresses a form of trying that doesn't require effort or significant progress. In the next few sections, the restriction to this minimal reading won't matter so much—you can just use 'try' as it strikes you in the cases—but in later sections it will matter a lot. In those later sections, we'll sometimes step back to recall the ways in which the putative context-sensitivity of 'try' might be driving certain objections to our arguments. When we do that, we'll often assume for concreteness that the hypothesis we just sketched is correct, and refer to our target as 'pure trying'. But even there, for the most part we will simply talk about trying, and ask for the reader's charity in understanding us as we intend.

3 Trying and believing

We now turn to more substantive points, beginning with our arguments for the "Cartesian" conception of trying, on which what a person can try to do is comparatively unconstrained

⁵ In this sense our quarry is not the notion of *attempt* as used in the law, which requires the performance of certain canonical steps and not just mere mental preparation (think of what's required to be convicted of *attempted* murder). So we're aiming at a target quite different from that of (e.g.) Yaffe (2004, 2010).

by their attitudes toward success. A good deal of our discussion of this idea will build on arguments which have been known for some time now. But these older arguments will be cast in a new light by our linguistically-focused approach. And, as we show at the end of the section, they will set up an important pillar of our defense of our central thesis here, that trying is agentially perfect.

Suppose a person is trying to ϕ . What, if anything, follows about their beliefs with respect to the claim that they will ϕ ? Here are three natural answers, in order of decreasing strength:

BELIEVE WILL If S is trying to ϕ , then S believes that S will ϕ .

BELIEVE PROBABLE If S is trying to ϕ , then S believes it is probable that S will ϕ .

BELIEVE MIGHT If S is trying to ϕ , then it is compatible with what S believes that S will ϕ .

The first of these two principles are natural places to start, but they can't be correct. People routinely try to do things they know have an enormously low probability of success: you can try to hit a hole-in-one, try to guess the combination of a safe, try to win the lottery, or try to write a bestseller. In each case you know you're almost certain to fail. But that's no barrier to your trying.

The status of BELIEVE MIGHT is less obvious. It has many defenders.⁶ But we follow a number of philosophers in rejecting it. We're moved by cases like the following. Suppose you're in front of a brick wall. You are certain that no matter how hard you push it, it won't fall over. Does it follow that you cannot try to push the wall over?⁷

We think not. We think you can try to push the wall over as part of an exercise regimen. Or that you can try to push the wall over—perhaps as hard as you possibly can—to prove that you are not strong enough to push it over. More generally, when you know you can't ϕ and want to prove that you can't ϕ to someone else, often the best way of doing so is to try as hard as you can to ϕ . You'll fail having done all you can to succeed, thereby illustrating your inability to ϕ .

These sorts of judgments seem to us about as clear as they come. But there is a rich tradition of rejecting them; here we'll consider two kinds of objections.

Arguably the most prominent objection to such cases has been to claim that they are just not cases of genuine trying.⁸ Perhaps the agents are merely *pretending* to try, or *acting as if* they are trying. It's unclear whether the proponent of this response should be developing it as an error-theory or instead as a theory of the pragmatics of ordinary 'try' ascriptions. But either way we find this response unsatisfactory.

⁶ See, e.g., O'Shaughnessy (1973, 1980), Jones (1983), McCann (1986), and Adams (1995, 2007).

⁷ This case comes from Harman (1986, p. 370); it and cases like it are discussed in detail by Ludwig (1992), Adams (1995), Ludwig (1995), Hornsby (1995), who believes (n. 5) that James Thomson was the first to present them. The earliest explicit discussion we know of is Hampshire (1959, p. 134) ("I may attempt something that I know or believe to be impossible, e.g. in order to demonstrate its impossibility or to test my powers..."). More concrete examples include Thalberg (1962, p. 54)'s example of a lifeguard trying to revive a person who he knows to be unsaveable, McCormick & Thalberg (1967, p. 45)'s example of trying to break an indestructible watch to prove it can't be done, Ludwig (1992, p. 262)'s example of trying to start a stalled car to prove it's stalled (among others), and Hornsby (1995, p. 526-7)'s example of trying to lift a weight to show one can't. For an example caught in the wild, see Elon Musk's infamous request to his assistant Franz to "try to break the window" of Tesla's supposedly bulletproof Cybertruck: <https://www.youtube.com/watch?v=LMWwImDX3ks>.

⁸ See McCann (1986, n. 20), Adams (1986, p. 288), Adams & Mele (1992, p. 330), Adams (1997, p. 301).

First, suppose that you press on the wall as hard as you can and, to your shocked surprise, it actually falls over. To our ears it would be rather strange for you to explain what happened with a speech like ‘I wasn’t trying to push the wall over, I was merely acting as if I was trying!’. The speech is no better if we substitute ‘pretending to try’ for ‘acting as if I was trying’. Better would be to say something like ‘Yes, I was trying to push the wall over—but only because I was sure I couldn’t actually do it!’. This is especially clear if you were trying to push on the wall to demonstrate to shareholders of a wall-building company that this model of wall really could not be toppled. If in that context you merely pretended to try to push the wall over, the demonstration would be fraudulent: it is not at all impressive that someone who is merely pretending to try to push a wall over fails to do so. To prove that the wall is strong, you need to try to push it over.

Second, the alternative account of the intuitive judgments about these cases is *ad hoc*. Consider *deciding*. To our ears, a speech like ‘I decided to push the wall over’ sounds bad in the context of the wall case. A good explanation for this is that an analog of BELIEVE MIGHT holds for deciding: you can’t decide to ϕ if it is not compatible with what you believe that you will actually ϕ . But there is no barrier to *pretending* to decide to ϕ or *acting as if* you have decided to ϕ in cases where it is not compatible with what you believe that you will ϕ . A speech like ‘I knew I couldn’t push the wall over, but I acted as if I had decided to push it over’ is perfectly intelligible. So why don’t we charitably read (e.g.) ‘I decided to push the wall over’ as meaning ‘I pretended to decide to push the wall over’, given that the latter report could easily be true? This is allegedly what we do with trying-reports like ‘I tried to push this wall down’, so why not here too? Absent a plausible story that explains why we re- or misinterpret ‘S tried to ϕ ’ as meaning ‘S acted as if S was trying to ϕ ’ but *don’t* re- or misinterpret ‘S has decided to ϕ ’ as meaning ‘S is acting as if S has decided to ϕ ’, we should be inclined to take our judgments about the trying-reports at face value.⁹

Third and finally, anyone who is trying as hard as they possibly can to ϕ is trying to ϕ . But as far as we can tell there is just no intuitive pull to the idea that one can only try as hard as one possibly can to ϕ if one believes one might ϕ . Indeed, its being background knowledge that a person can’t actually ϕ tends to make it especially natural to tell them to try *as hard as they possibly can* to ϕ .

A second way to resist our judgments about these examples admits that they seem initially plausible, but argues that the relevant judgments weigh against other equally strong ones and so must, on reflection, be rejected. In this case, the opposing judgment is the following simple one: there are lots of things it seems like we cannot even *try* to do—let alone actually succeed in doing—and the best explanation of this fact is that something along the lines of BELIEVE MIGHT is true.¹⁰

⁹ McCann (1986, n. 20) suggests distinguishing between a “primary sense” and a “secondary sense” of ‘try’, writing that ‘try’ in our examples is used in its secondary sense, a use which is “deceptive: it arises from the fact that even though, in the primary sense, I am not trying to touch the ceiling, my action consists in precisely what I would do if I were to try, in that sense, to touch it.” But we find such a response unsatisfactory for two reasons. First, we’ve just seen that ‘decide’ seems not to have any “secondary sense” of this sort, making the suggestion that ‘try’ has it look *ad hoc*. And second, if there were different senses of ‘try’ at work here, then it should be possible to truly say that the wall-pusher did not try to push over the wall (since they didn’t in the primary sense). But instead it just seems unequivocally false.

¹⁰ In a series of important papers, Frederick Adams has developed a related objection. Adams presents an elegant, simple theory of the relationship between desire, belief, intention, attempt and intentional action, and suggests

For example: suppose there's a rock in front of you, and we ask you to try to make it to levitate. You might reasonably reply 'I don't know how to do that'—meaning not just that you don't know how to get the rock to actually levitate, but that you don't even know how to *try* to get it to levitate. This case might be thought to support BELIEVE MIGHT. For supposing this principle is valid, it would provide a good explanation of why it seems that people cannot try to get rocks to levitate: we can only try that which we take to be possible; we know (and thus believe) that nothing we can do could result in the rock levitating; so we can't try to levitate rocks.¹¹

But there is a better explanation of what's going on when people take themselves not to be able to try something in a case like this one, and this better explanation does not appeal to anything in the vicinity of BELIEVE MIGHT. The explanation is that the people in question simply haven't thought of a way to try to levitate the rock. If they did think of one—through cogitation or the right sort of prompting—they'd see that it's as easy to try to make the rock levitate as it is to try anything else. To see this, suppose that instead of merely asking you to try to get the rock to levitate, we asked you to try to get the rock to levitate *with your mind*, as though you had telepathic powers. You might still insist that you can't lift the rock as a way of resisting our request, but it's obvious now that you're not saying you can't *try* to lift the rock with your mind. Most people find it easy enough to do that. A typical response involves staring at the rock, scrunching up your face, and imagining the rock gently lifting off the ground. (All to no effect, of course.)¹²

These sorts of considerations make us think that it's not just that you can try to do things when you believe you won't succeed; but that you can even try to do things when you know that it's metaphysically impossible that you'll succeed. For example: someone who knows that it is mathematically (and hence metaphysically) impossible to trisect an angle of 60° using only a straightedge and compass can still try to trisect it using only these means. They might try it to "prove" that they can't do it, as for instance in front of a class, or try it to gain intuition as they search for a new proof of this known result, or try it to see

that given the virtues of this theory "[u]nless an example clearly cannot be fit to the theory, is overwhelming, or the rescue of the theory is clearly ad hoc, the virtues of the theory may well warrant the rejection of the purported counterexample." (Adams (1994b, p. 620), cf. Adams (1995, p. 554), Adams (1997, p. 299), cf. Adams (1994a)) Adams argues that, unlike his own theory, a theory which admits that our wall-pusher is trying to push over the wall, cannot offer an account of what makes some activities trying to do particular things, while others are not. We agree with Adams that a good theory should not be abandoned just because of some bad data. Still, there is a line between developing a strong theory and changing the topic of the discussion. In our view, rejecting the above examples amounts to discussing not trying (i.e. what is expressed by the English expression 'try') but something else. Here, we find ourselves largely in agreement with Ludwig, who writes:

The resulting picture is not a simple one, and does not promise to have a sharp informative set of necessary and sufficient conditions for what counts as a trying though it is thought of as having no chance of success, any more than there is a sharp informative set of necessary and sufficient conditions for something to be a wrench. But there is no shoehorn that will fit our concepts into a theory not made for them. Our first responsibility is to our data, our intuitions about cases. I agree with Adams that what we want is a simple view, but we do not want an overly simple view, one that does not accommodate the richness of our conception of human agency. (Ludwig (1995, p. 570))

¹¹ Our arguments below that people can try to do things they know to be metaphysically impossible undercuts other analyses of the cognitive condition on trying as well, e.g. Ludwig (1992, p. 268-9), Ludwig (1995), Hornsby (1995, p. 531-2).

¹² See Albritton (1985, esp. p. 245) for a related idea.

how good of an approximation to such a trisection they can get.¹³

This concludes our defense of the claim that one can try that which one takes to be impossible. We take it to be interesting in its own right. But it is also of central importance for our main argument here. The reason is that, if a constraint like BELIEVE MIGHT were genuine, it would yield a straightforward argument against the thesis that trying is agentially perfect—that is, the claim that a person tries to ϕ if and only if they try to try to ϕ .¹⁴ For suppose that trying to ϕ really did require believing that one might ϕ . Now imagine that someone is (correctly) certain that squaring a circle is impossible, but falsely believes (remember our supposition) that it is possible to try that which one takes to be impossible. Given these stipulations, it would follow that our hero can't try to square the circle: they believe it's impossible to square the circle, and by hypothesis anything one believes to be impossible is something one cannot try to do. But for all we've said they can still try to try to square the circle, since they don't believe it's impossible to try the impossible. So, if these constraints were genuine, it would follow that there is something one can try to try to do but cannot try to do (namely squaring the circle). And so trying would not be agentially perfect. But, as we have shown, the constraints are not in fact genuine, so the objection is defused in advance.

This kind of objection doesn't arise just in the case of beliefs about one's success. Similar arguments could be made using other putative constraints on trying—most obviously, ones that tie facts about what an agent can try to do to facts about what they want or intend to do. So, in the next two sections, we'll continue to lay the foundations for the defense of our main claim by arguing against these latter constraints as well.

4 Trying and wanting

Suppose again that someone is trying to ϕ . Does it follow that they want to ϕ , in at least *some* sense of 'want'? In other words, what should we make of the following principle?¹⁵

¹³ One might worry that in allowing such cases of trying we have really gone too far, since they conflict with putatively plausible principles concerning the logic of trying. For instance: it seems plausible that someone who is trying to take a step with their left foot is thereby trying to take a step with at least one foot. Likewise, it seems plausible that someone who has a normal conceptual repertoire and is trying to paint a wall scarlet is thereby trying to paint it red. These observations might seem to suggest that trying is closed either under necessary implication, or under known implication, or something of the sort. And if it is, then a person who tries the metaphysically impossible would (absurdly) try everything whatsoever, including bringing about the ruin of the world.

But the fault here lies with the closure principles, not with our examples. Consider by analogy the attitude of wishing. One can wish to drink heavily, or wish that one had won yesterday's lottery, or wish that the continuum hypothesis were not independent of ZFC. But it seems clear to us that people who have these wishes do not automatically wish to become hungover, wish to have spent money on a lottery ticket, or wish that ZFC were inconsistent. This is not to say that the logic of wishing has *no* discipline. A person who wishes that they had taken a step with their left foot *does* seem to thereby wish that they had taken a step with at least one of their two feet, and someone who wishes to paint their wall scarlet *does* seem to thereby wish to paint their wall red. Trying seems to us to be the same: it allows some "closure-like" inferences, though not all of the ones licensed by the strong principles they objector assumes.

¹⁴ Pollock (2002, p. 12) uses BELIEVE MIGHT in a similar way to argue that trying is not "infallibly performable". See also Hermes (2006, p. 64-5).

¹⁵ For related discussion see Mele (1990, 1991, 1994), Adams (1991, 1994b,a), and Hornsby (1995). We remind the reader that our interest is in the relations picked out by the natural language expression 'want'. Our claim is that there is no interpretation of 'want' consistent with the conventions of English on which the inference from 'S is trying to ϕ ' to 'S wants to ϕ ' is valid.

TRYING \rightarrow WANTING If S is trying to ϕ , then S wants to ϕ .

Well, given that a person can try to ϕ even when they are certain that they won't ϕ no matter how hard they try, it isn't difficult to see how a person might try to ϕ even when they don't in any sense want to ϕ . The wall case demonstrates this vividly. The person who tries to push over the wall for exercise might not only have no desire to push the wall over, but in fact have a very strong desire that they not push it over.¹⁶ Perhaps if the wall fell it would destroy a priceless work of art, or kill the wall-pusher, or kill the wall-pusher together with those inside. It's only because the person is certain they won't be able to push it over that they are willing to try.

These cases show that you can try to ϕ without wanting to ϕ in cases where you are sure you won't ϕ . But what if you aren't sure? If a person is trying to ϕ and believes that it is possible that they will succeed in ϕ -ing, does it follow that there is some sense in which they want to ϕ ?

Once again we think the answer is no. Suppose you're playing your friend in chess. Your friend is a much stronger player—so much stronger, in fact, that you accurately take them to be a 99% favorite to win the match. However, you also know that there are two respects in which your friend is rather emotionally insecure. First, if they lose they'll be psychologically devastated. Second, if they detect that you aren't trying your hardest to beat them—and let's suppose they have a keen eye for a sham attempt—they'll again be psychologically devastated, whether or not they win. Knowing all this, and supposing that you enjoy playing chess but don't intrinsically care about whether you win or lose, can you (rationally) try to beat your friend in chess?

We think you can.¹⁷ Indeed, it seems the best thing to do in these circumstances is to try as hard as you can to beat your friend in chess. Yes, if you manage to succeed in doing what you're trying to do the consequences will be bad. But the consequences will also be bad if you don't try your hardest to win. And the chance that you win conditional on trying your hardest is very low—low enough to make the gamble worth it. But none of this means you want to win: plausibly there is no sense at all in which you do. So you can (rationally) try to ϕ without at all wanting to ϕ , even when you believe it's possible that you will succeed in ϕ -ing.

One might object to this argument on the following lines. The particular chess moves you make in the game are ones that you genuinely want to make. Moreover, it seems natural to think that part of what explains why you want to make those particular moves is that you recognize them as the ones that give you the greatest chance of winning. But the fact that you recognize a move as the one that gives you the greatest chance of winning only explains why you made that move if we further assume that you have the aim or goal of winning. And if you have this aim or goal, then you must want to win after all.

We have two responses to this objection. First, the fact that you want to make the moves you're making could just as well be explained by your recognizing them as the moves that

¹⁶ Here and throughout we stipulatively use 'desire' as the nominalization of 'want'. As such, we will move seamlessly between talk of agents wanting to ϕ and talk of agents having a desire to ϕ . Whether ordinary, non-stipulative uses of 'desire' license this inference is less clear: see [Davis \(1984\)](#) for discussion.

¹⁷ [Hornsby \(1995\)](#), p. 529 and n. 11) makes this point using a version of [Bratman \(1987\)](#)'s video game case.

give you the greatest chance of being (correctly) perceived as trying your hardest to win. You can want to make the relevant moves for this reason alone, and not because they are the winning moves.

Second, even supposing you really do want to make these moves because they are ones that will help you win, it would not follow that this can be explained only on the assumption that you want to win. Perhaps what generates your desire to make the winning moves is not the fact that you want to win, but the fact that you are trying to win. In other words, perhaps desires for means can be generated by the facts about what one is trying to do, and not merely the facts about what one wants as ends.

Indeed, this appears to be the lesson of the example. If I know you've figured out the best move, I can truly say that I know which move you want to make next. If an onlooker asks me how I know that, I could explain by saying 'They know that that move gives them the best chance of winning, and they're trying their hardest to win'. But intuitively I'd get your state of mind wrong if instead I said 'They know that that move gives them the best chance of winning, and they want to win'. So even conceding that you are trying to win and want to make the winning moves (as such), we deny that it follows that you want to win. If these claims are right, then the case carries a broader moral: ϕ -ing can be among one's aims or goals simply in virtue of the fact that one is trying to ϕ , even when one has no desire to actually ϕ (cf. Mele (1990, p. 252)). If we accept this principle, we can still hold that all desires for means are generated by aims or goals that one has; we'll just think that people can sometimes have aims or goals in virtue of what they are trying to do, rather than in virtue of what they want to do.

So, trying does not in general require wanting. Why, then, has the claim that it does struck so many as attractive? Our counterexamples to the principle have relied crucially on the fact that one can have motivating reasons to try to ϕ even when one has no reason whatsoever to ϕ (and perhaps quite strong reason *not* to ϕ).¹⁸ But it's natural to think that these cases are in a certain sense exceptional—that "paradigm" cases of trying are those in which an agent is trying to ϕ because they take it to be a necessary means toward an end they desire to bring about. So perhaps something like the following principle could be made to work:

RESTRICTED DESIRE If you are rational and all your reasons for trying to ϕ are also reasons for ϕ -ing, then if you try to ϕ , you want to ϕ .

To be clear, we are not suggesting that non-paradigmatic tryings are in any interesting sense rare. We suspect that the phenomenon brought out by the wall and chess cases is actually quite common, arising whenever we shoot for the stars ("Try to punch through him", "Try to touch the ceiling", "To see how close you can get, try to draw a perfect circle", "Try to count to a million, you'll fall sleep soon enough"...). However, we also suspect that the capacity for this kind of non-paradigmatic trying is a mark of cognitive sophistication, much in the same way that the capacity to resist the testimony of one's senses (in cases of suspected or known illusion) is a mark of cognitive sophistication. For much of the animal kingdom, it may well be that trying is a reliable indicator of desire. But adult humans can

¹⁸ See McCann (1986, p. 206-208), Mele (1990, p. 252) (cf. Mele (1992b, p. 61f.)), and Hornsby (1995).

think about the possible effects of tryings in ways that most animals cannot. For creatures like us, the connections between trying and desire are more tenuous.¹⁹

5 Trying and intending

Having argued that one can try to do things even when one believes it is impossible to succeed, and even when one in no sense wants to succeed, we now want to ask whether one can try to do things when one does not intend to succeed—that is, whether the following principle is true:

TRYING \rightarrow **INTENDING** If S is trying to ϕ , then S intends to ϕ .

Several authors have defended this idea.²⁰ But the arguments of the previous two sections already suggest strong reasons against it. This would follow immediately if one accepted the commonly (though not universally) endorsed idea that if one intends to ϕ , then one does not believe that one will not ϕ .²¹ Likewise, it would follow immediately from our earlier discussion if one endorsed the claim that if one intends to ϕ , then there is *some* sense in which one wants to ϕ . For, as we have seen, trying has neither of these entailments.

But independent of any general principles like these, the various cases discussed in the previous two sections also present intuitive counterexamples to **TRYING** \rightarrow **INTENDING**. A person who tries to push an unmoveable wall down for exercise, or to show that it can't be done, does not intend to push the wall down. And although you may be trying your hardest to beat your friend in chess, you do not intend to win. So trying does not entail intending.

In these cases, the person who tries to ϕ without intending to ϕ either does not believe that it's possible for them to ϕ , or does not want to ϕ . But what if a person tries to ϕ while wanting to ϕ and believing that it's possible that they will ϕ ? Must they then intend to ϕ ?

Once again, the answer turns out to be 'no' (and the case which gives us this answer also gives us an independent argument that trying doesn't require intending). Suppose you buy a ticket for a fair lottery. If I were to ask you why you did that, it would not be

¹⁹ In our view similar points apply to O'Shaughnessy's (1973, p. 369) claim that trying "is constituted by doing, intentionally and with just that purpose, whatever one takes to be needed if, the rest of the world suitably cooperating, one is to perform the action." As restricted to paradigmatic tryings, O'Shaughnessy's claim seems plausible enough. But counterexamples abound in cases where one's reasons for trying to ϕ are distinct from one's reasons for ϕ -ing. Take the wall case from earlier, but now imagine the wall has a strong section and a weak section. The strong section is strong enough that no matter how hard you push on it, the wall won't fall over, whereas the weak section is weak enough that any amount of pushing will cause the wall to collapse. Knowing this, you know that pushing the wall over will require pushing on the weak spot. But you still could try to push the wall over, by trying to push it over in the strong spot, whether to get some exercise, or to show that it can only be pushed over in the weak spot. Similarly, an excellent tennis player can try to beat a weak player by trying to beat the weak player with their non-dominant hand. These cases are all cases of trying even though they're not cases where the person is doing whatever they take to be needed to accomplish what they are trying to do. They are also counterexamples to Ludwig's (1995, p. 569) "least resistance principle", which says that "A is a trying to B only if it is conceived by the agent to be the path of least resistance to his end..."

²⁰ See Hampshire (1959, p. 107) McCann (1975, 1986, 1989), Adams (1986, 2007), Adams & Mele (1992), Ludwig (1992, 2021), Yaffe (2004, 2010), Grano (2011, 2017), Shepherd (2016, p. 422).

²¹ But see Thalberg (1962), Hedman (1970), McCann (1986), McCann (1991), Ludwig (1992) and Buckwalter *et al.* (forthcoming) for dissent.

reasonable for you to respond ‘Because I intend to win the lottery’. An answer along those lines suggests that you are either joking or deluded. And if we assume that you’re neither of these things, it becomes natural to infer that you take yourself to have some means of fixing the result in your favor. There is something about fair lotteries that makes winning them unfit as an object of one’s intentions.²²

By contrast, a rational agent can absolutely *try* to win a fair lottery. If I were to ask you why you bought a ticket, it would be perfectly appropriate for you to respond ‘Because I’m trying to win the lottery’, and there would be no impression that you had to be joking, deluded, or planning on rigging the outcome. So one can try to win a fair lottery even if one does not intend to win it, and even though of course one wants to win it and thinks that it’s possible that one could. Trying is not a species of intention.²³

As we discussed at the end of §3, our arguments for the Cartesian conception of trying—on which trying to ϕ is compatible with not believing or wanting or intending to succeed—are necessary for responding to an important objection to our main claim: that trying is agentially perfect. We’ve now completed our preemptive strike against that objection, so it’s time to turn to the main claim itself.

6 Agential perfection

We start, in this section, by bringing the property of agential perfection into focus. In the next section we’ll argue that trying satisfies this property.

We opened the paper by suggesting that God’s attempts are immune from failure: if God tries, God succeeds. By contrast, humans routinely try and fail. But still, one might

²² McCann (1986), McCann (1989), Ludwig (1992, p. 269-270), Adams (2007), Ludwig (2021, p. 345) hold that a person can intend to A even if they think A is very low-probability (or, in Ludwig’s case, impossible), and argue against intuitions to the contrary on the basis of the pragmatics of “intend”-reports. They suggest that, since typically one does not intend to ϕ unless one believes one has a good chance of ϕ -ing, saying that a person intends to ϕ implicates that they have a sufficient degree of confidence that they will succeed. But it is some evidence against this claimed implicature that it does not appear to be cancellable. (Otherwise, we should expect to find acceptable speeches like “Mary intends to win. Of course she is certain she will not win, but she still intends it.”) The view would also seem to predict that ‘intend’ and ‘try’ would have the same patterns of implicature, which they do not.

²³ A connection between this observation and “Kraemer’s puzzle” may strengthen the argument here. In discussing a puzzle introduced by Butler (1978), Kraemer (1978) noticed that even when a person knows that they will do one thing if and only if they do another, they may do one of the two intentionally, and the other not. Here is an example taken from Blumberg & Hawthorne (2022) (cf. Ross (1978) and Nadelhoffer (2004)):

Uma has a button in front of her. If she pushes it, a machine will shoot an arrow randomly down one of nine paths, numbered one to nine. If (and only if) the arrow goes down path three, it will hit and kill Bill. Uma knows these facts. She wants to kill Bill, and knows that pushing the button is her only chance of doing so. So she decides to push the button. As it happens, the machine sends the arrow down path three, killing Bill.

It is standardly judged in this case that Uma intentionally killed Bill, although Uma did not intentionally send the arrow down path three. We suspect that facts about intention exhibit a similar pattern, though we acknowledge the judgments here are more subtle. We find it natural to think that, Uma intended to kill Bill by pressing the button, although it is not true to say that she intended to send the arrow down path three. So, plausibly, Kramer’s puzzle arises for intention in something like the way it arises for intentional action.

But crucially the puzzle does not arise for trying: Uma is both trying to kill Bill and trying to send the arrow down path three. Unlike intention, trying seems to be insensitive to the difference between the event of killing Bill and the event of sending the arrow down path three, providing further evidence that trying is not a species of intention.

ask, is there a class of doings such that even we humans cannot fail to do them if we try?²⁴ To have a name for this condition, we will say that

ϕ -ing is *essentially successful* iff: necessarily, if S tries to ϕ , then S ϕ s.

As we illustrated earlier, many doings lack this property. People try to stay calm and fail. They try to get high and fail. They try to assuage a crowd and fail. None of these doings is essentially successful.

But arguably some doings are. Maybe trying to think about Confucius requires having Confucius in mind enough that you're already thinking about him. If so, then thinking about Confucius would be essentially successful.

The class of essentially successful doings is interesting in its own right. But our aim here is to characterize a domain in which humans approximate the picture of divine agency with which we began, and so far we're still not there. For even supposing that thinking about Confucius is essentially successful, it can happen without being attempted—as when one wakes up in a sweat considering what the sage would say about one's relationship to one's parents, or finds oneself thinking about him in response to someone saying his name. This shows that even if a doing is essentially successful, it does not thereby rise to the level of divine agency which interests us here. God does not just do whatever God tries to do; God also doesn't do something unless God tries to do it.

These reflections motivate a second condition:

ϕ -ing is *essentially attempted* iff: necessarily, if S ϕ s, then S tried to ϕ .

In a way, it's easier to come up with at least initially plausible examples of doings that are essentially attempted (even if one may not ultimately accept them). Maybe *defrauding the state* is one: perhaps you can't defraud the state unless you try to. Likewise for *concentrating on one's breathing*: perhaps you can't concentrate on your breathing unless you're trying to. More generally, Anscombe (1957, §47) argues that some doings are *essentially intentional*: if a person performs them, they perform them intentionally. Putatively included in this class are *greeting*, *marrying*, and *promising*. Though we have doubts whether these sorts of things really are essentially *intentional*, perhaps they are essentially *attempted*.²⁵

To be clear: the examples we just considered are clearly not “essentially attempted” on all ways of interpreting ‘try’ as it appears in that condition. A person can clearly defraud the state without having tried *hard* to do so; some people are just naturals. If ‘try’ is context-sensitive or ambiguous, there are readings on which ‘try’ expresses what is expressed by ‘try hard’, and so there would be readings on which defrauding the state is not “essentially attempted”, for it can be done effortlessly. But as we emphasized earlier (and will from now on be especially important), we will be focused exclusively on very permissive readings of try—or, given our simplifying hypothesis about the context-sensitivity of try—on uses of

²⁴ Here we use the unlovely expression ‘doings’ in place of ‘actions’, to avoid contested questions about the characterization of action. We intend ‘doings’ rather expansively, covering roughly anything that can be expressed by a verb phrase in English, including for instance dying, digesting, and drooling.

²⁵ Relatedly, Beddor & Pavese (forthcoming) reject the claim that these actions are essentially intentional, but propose instead that they are essentially *intended*.

‘try’ in contexts where it expresses *pure trying*. The fact that someone can defraud the state without having put much effort into it is thus no counterexample to the intended interpretation of the claim that defrauding the state is essentially attempted.

The properties of being essentially successful and of being essentially attempted are both, in our view, of great interest on their own. But here we won’t say much about them separately. Instead, we’ll be interested in our promised notion of agential perfection, which we obtain by putting the two conditions together:

ϕ -ing is a *agentially perfect* iff: necessarily, if S tries to ϕ s, then S ϕ s, and if S ϕ s, then S tried to ϕ .

7 Trying is agentially perfect

We will now argue that trying satisfies this property—that necessarily, one tries to ϕ if and only if one tries to try to ϕ .²⁶ In doing so, we will also argue for a strictly stronger metaphysical thesis about the relationship between trying and trying to try, namely:

IDENTITY To try to ϕ just is to try to try to ϕ .

IDENTITY implies that trying is agentially perfect, but the converse need not hold: it is logically possible that trying and trying to try are merely necessarily equivalent rather than identical. But, while this is logically possible, it is hard for us to see a principled reason for believing in such a gap. So here we’ll take the agential perfection of trying to stand and fall with IDENTITY, and take arguments for either one of these theses to be arguments for both.

Our first argument for these claims begins from the fact that IDENTITY follows from a natural (albeit simplistic) picture of trying.²⁷ On this simplistic picture, to try to ϕ just is to begin or be in the process of ϕ -ing. If the picture were correct, then IDENTITY would, because to begin or be in the process of beginning or being in the process of ϕ -ing just is to begin or be in the process of ϕ -ing. To see this, note that we are typically *not* inclined to give a certain answer to the question ‘During what period of time had you begun or been in the process of quitting drinking?’, but a different answer to the question ‘During what period of time had you begun or been in the process of beginning or being in the process of quitting drinking?’. The latter question just seems like a bizarre way of asking the former question. And this is presumably because we do not distinguish the times for which one counts as beginning or being in the process of something from the times for which one counts as beginning or being in the process of beginning or being in the process of it. So, the simplistic picture implies that to try to ϕ just is to try to try to ϕ .

This simplistic picture is, however, too simple. Some cases of beginning or being in the process of ϕ -ing aren’t cases of trying to ϕ : if one begins to fall by accident, or is in the process of doing so, one is typically not trying to fall. Similarly, one may be in the process

²⁶ In fact, we suspect that trying is the *only* natural doing that is agentially perfect, but we won’t argue for this claim here.

²⁷ Hermes (2006, 63-4) considers a somewhat related argument (which he ultimately rejects) that McCann’s analysis of trying entails that trying is essentially successful. Cf. also Ludwig (2021, p. 346).

of digesting one's food even if that isn't something one can try to do. Still, the attractions of the simplistic picture suggest that it captures something important about trying. There is a structural analogy between trying to ϕ on the one hand and beginning or being in the process of ϕ -ing on the other: the former is something like the "agential" version of the latter. This structural analogy, together with the fact that the analog of IDENTITY holds for beginning or being in the process, provides support for IDENTITY itself.

Our second argument concerns linguistic judgments about iterated 'try' reports. Consider first the claim that trying is essentially successful, i.e. what we will call

COLLAPSE Necessarily, if S is trying to try to ϕ , then S is trying to ϕ .

In contexts where 'try' gets its minimal reading, it is hard to see what a counter-instance to this claim would look like. For example, if in such a context you were to ask me 'Did you try to lift that rock with your mind?', it does not seem I could felicitously reply with 'No, but I did *try to try* to lift it with my mind'. After all, if I tried to try to lift the rock but failed to actually try, then either my failing to actually try was intended or not intended. If it was *not* intended, then presumably my failure to first-order try was due to some kind of ineptitude on my part. But what could such ineptitude even look like here? Could there really be something I was supposed to have done differently? And if my failing to try was intended, then presumably I believed it was possible that I try to try to lift the rock without actually trying to lift it. To believe such a feat possible, I would have to have a sense for what it would look like to try to try to lift the rock with my mind without actually trying to lift it. But what could such a sense even come to? We thus find it hard to see how COLLAPSE could have false instances.

Similar points apply to the claim that trying is essentially attempted, i.e. what we will call

ITERATION Necessarily, if S is trying to ϕ , then S is trying to try to ϕ .

An argument in favor of ITERATION is that it explains why, if someone tells you to try to try to ϕ , you can comply by simply trying to ϕ —at least in contexts where 'try' gets its minimal reading. Suppose I tell you 'You should really try to try something you know to be impossible; it's liberating'. A good way of complying with this advice is to go and (say) try to push down the wall, or try to trisect a 60° angle. But this wouldn't obviously be a good way of satisfying the request if trying did not entail trying to try.

These two arguments make us think that IDENTITY and the claim that trying is agentially perfect ought to have a kind of default status. But the two arguments on their own wouldn't have gotten us all the way to believing these claims, and we don't expect them to have done this for you either. The full case for the claims rests not just on the positive arguments in their favor, but also on the fact that, once one attends to the various readings of 'try' and takes care with the logic of trying, what might have seemed to be obvious arguments against them turn out to be quite weak. It is this fact, together with the positive arguments just presented, that moves us on balance to endorse IDENTITY and the agential perfection of trying.

In defense of the claim that the objections are weak, we'll present the four we find most compelling, and argue that they fail.

First, O'Shaughnessy (1973) argues that claims about whether one tried to try are unintelligible.²⁸ But we think this can't be right: such claims may be hard to assess, but they're not meaningless. Indeed, as we just argued, a request along the lines of 'You should really try to try something you know to be impossible; it's liberating' is just plainly intelligible.

A second and more serious worry about ITERATION (and thus about IDENTITY) arises from cases like the following.²⁹ A yoga instructor tells their students not to try to breathe from their diaphragm, but instead to breathe from their diaphragm naturally and automatically. A student tries to follow the teacher's directions, but fails: they breathe through their diaphragm in a way that is conscious and unnatural. It seems plausible that the student tried to breathe through their diaphragm and thus failed to follow their teacher's instructions. But it also seems plausible that the student did not try to do this—that is, that they did not try to try to breathe through their diaphragm. So plausibly one can try without trying to try.

We can see two ways of supporting the claim that the student does not try to try to breathe through their diaphragm, but we think neither of them quite works. A first way starts from the intuitive observation—which we are happy to concede—that it is natural to describe the student as trying to *not* try to breathe through their diaphragm. It then adds in the general principle that if you're trying to $\neg\phi$, then you're not trying to ϕ . From the observation and the principle together, it follows that the student is not trying to try to breathe through their diaphragm.

But we reject the general principle. People can knowingly both try to ϕ and try to $\neg\phi$, and even be rational in doing so. Suppose that, from long experience, you've learned that the best way for you to hit the bullseye in a game of darts is for you to try not to hit it, and, in particular, to try instead to hit a point about a foot above it. In a tense game, when you need a bullseye to win, you may try to hit the bullseye by trying not to hit it. Since, in such a case, you are both trying to hit it and trying not to, we deny that trying not to do something implies that one is not trying to do it. Accordingly the student may be trying not to breathe through their diaphragm, but it doesn't follow that they are not trying to breathe through it.³⁰

The second way of motivating the claim that the student does not try to try to breathe

²⁸ McCormick & Thalberg (1967), Hunter (1987), also voice worries along these lines. Cf. Mele (1992b, p. 64).

²⁹ We got this case from correspondence with Adam Elga. Thanks to Elga and Simon Goldstein for discussion here.

³⁰ More controversially, Bratman's (1984) video game case may provide another illustration of this. Suppose you're playing a game where you win if and only if you guide a missile to one of two targets. You know that if you happen to hit both targets at once, the game shuts down and you lose. You also know that the chances of hitting both targets simultaneously is very low, and thus that the best strategy is to try to hit both. The player in this game can naturally be described as trying to hit both targets (where 'both' scopes under 'try'). This description is especially natural if it is understood as a means toward their end of hitting exactly one. In addition, however, the player can also naturally be described as trying not to hit both targets: they are trying not to shut down the game, after all. Indeed, one can even describe the player as trying to hit both targets in an effort to try not to hit both targets (for detailed argument in favor of this position, see Hornsby (1995, p. 529); for argument against see Ludwig (1992, p. 273-4)). Similarly (but similarly more controversially) it seems plausible that Ulysses was trying not to heed the Sirens (that's why he tied himself to the mast), at the same time as he was trying to heed them (that's why he was struggling to free himself).

through their diaphragm is on the basis of a direct intuition: it's just supposed to be intuitively obvious that they aren't trying to try to breathe through their diaphragm. But we're skeptical there is any direct intuition to this effect. We suspect that those who claim to have this intuition are interpreting 'try' as meaning something like *consciously try*. On that reading, it does seem correct to say that the student is consciously trying to breathe through their diaphragm (against their instructions), but also that they aren't consciously trying to consciously try to breathe through their diaphragm.

But pure trying is not always conscious trying: tennis players can try to put certain kinds of spins on their shots without thinking about it, musicians can try to play in tune without thinking about it, and writers can try to spell their words correctly without thinking about it.³¹ So the fact that the student consciously tries without consciously trying to consciously try is no threat to ITERATION on its intended reading. If, on the other hand, 'try' *isn't* interpreted as meaning something like 'consciously try', then it's just not obvious that the student isn't trying to try to breathe through their diaphragm. In fact, we can give an argument that the student is trying to do just this. The student is trying to follow their instructor's advice. The instructor's advice is to try to breathe through one's diaphragm automatically rather than consciously. So the student is trying to try to breathe through their diaphragm automatically rather than consciously. But if the student is trying to try to breathe through their diaphragm automatically rather than consciously, then surely the student is trying to try to breathe through their diaphragm *tout court*. So we deny that there is a genuine counterexample to ITERATION here.

A third objection to ITERATION and IDENTITY claims that simple animals lack the conceptual wherewithal required to try to try to do anything. This objection seems to rest on two ideas: first, that trying requires a rich ability to conceive of the objects of one's attempts under a particular description; and second, that animals lack such an ability. But neither of these ideas is correct. With respect to the first idea, suppose a bear that has escaped from the zoo puts a girl's smartphone into its mouth. A distraught uncle calls to the zookeeper: "Sir, please help me, the bear is trying to eat my niece's smartphone!". Does this ascription show that in the relevant sense the bear has the concept of a smartphone, or of a niece? Presumably not. With respect to the second idea, it is far from clear that animals lack the concept of trying. We routinely say things like "The dog knows that its owner is trying to take out the food" or "The spider knows that the fly is trying to escape". To the extent one might be inclined to give an error-theory about these speeches, it's hard to see why one wouldn't want to do the same for the corresponding first-order 'try'-ascriptions. So animals do seem to have the conceptual wherewithal to try to try, provided they do try at all.

A fourth objection to IDENTITY is the one we raised at the close of §3 and that was a guiding thought behind our discussion in sections 3-5. To recall, the basic idea was that if

³¹ Pace Ginet (1990), who holds that trying has an "actish phenomenal feel", which we take to imply that trying is always conscious (for further arguments against Ginet's view see Adams & Mele (1992), Mele (1992a)). Shepherd (2016) (cf. Shepherd (2017)) endorses a more moderate position—consistent with ours—on which "the neural activities which subserve conscious trying are identical to those which subserve trying", a view he understands (p.c.) to imply only that all conscious tryings are identical with tryings, not that all tryings are conscious. For further discussion of the role of the experience of trying in the phenomenology of acting see Kriegel (2015, Ch. 2.5), and, on its connection to knowledge that we are acting, Peacocke (2004, Ch. 2.4), with Peacocke (2006, 2009).

trying requires believing that one will succeed, or wanting or intending to succeed, then we could imagine someone having these beliefs, desires or intentions toward trying to ϕ while lacking them toward ϕ -ing, thereby generating a counterexample to COLLAPSE.

In our view, this style of argument is extremely important, and that's why we took some time with it above. But since we've argued that each of the relevant principles fails, we take the objection to have been defused.³²

Perhaps more important than the specifics of these responses is that, once one takes care with relevant readings of 'try' and the facts about the logic of trying, the arguments above don't threaten the *motivations* for COLLAPSE, ITERATION, or even IDENTITY. In this sense, they're quite different than the cases we used to argue against various claims that trying requires certain attitudes toward success. Those earlier cases show that one can have a reason for trying to ϕ that has little or nothing to do with one's reasons for ϕ -ing. And once one sees this point, one loses the sense that there *should* be simple necessary connections between one's trying to ϕ and one's beliefs, desires, and intentions about one's success in ϕ -ing. The putative counterexamples to COLLAPSE, ITERATION, and IDENTITY we've just seen do not have this sort of feel, and thus do not threaten the default status of these principles. We thus take the case in favor of this strong, simple theory of trying to be secure.

8 Trying and intentional action

We now turn to drawing out some consequences of IDENTITY. We start in this section with consequences for the logical structure of intentional action. In the next section we turn to consequences related to the notion of options in decision theory and the notion of basic actions in action theory.

As a warm-up, we start by considering the idea that successful trying guarantees intentional action, that is:

SUCCESSFUL TRYING \rightarrow **INTENTIONAL** If S tries to ϕ and ϕ s as a result of this trying, then S ϕ s intentionally.

Although initially appealing, we think this principle should be rejected. One well-known problem is that a person can succeed in doing what they are trying to do because of

³² It might seem that we ourselves are vulnerable to a different instance of this form of objection, since we believe that if a person tries to ϕ , they have to have thought of a way to try to ϕ (recall the discussion of the rock example earlier). It might seem natural to describe a person who is trying to think of a way to ϕ as trying to try to ϕ , but according to us if they haven't yet thought of a way to try to ϕ , they can't yet be trying to ϕ , yielding an apparent counterexample to COLLAPSE.

Our response is that different cases require different diagnoses. In some cases we think that, even though it might seem to the trier that they haven't yet thought of a way to try, their very trying to think of a way to try is itself a way to try, so unbeknownst to them, they have thought of a way. In other cases, it seems to us that in trying to think of a way to try, the trier is not actually trying to try, but rather trying to put themselves in a position to try, which is a different thing altogether. If I deviously give someone a pill which makes them disposed to try to feed the llamas, then I haven't thereby tried to try to feed the llamas; I've just tried to change someone's mind so that they will try. Similarly, if I give *myself* the same pill, I also haven't tried to try to feed the llamas; instead I've just tried to put myself in a position to try. See also [Hermes \(2006, p. 65-8\)](#) for discussion of related cases (though drawing very different conclusions).

that very trying, but in a way that is too abnormal for the success to count as intentional.³³ For example: suppose Anna is trying to kill her uncle Bill. She fires a clear shot at what looks to be her uncle in the distance and hits her target, but the target is just a cardboard cutout of her uncle, not the man himself. Bizarrely (and unbeknownst to Anna), however, her uncle happened to be standing right behind the cardboard cutout at the time Anna took her shot, and he was incidentally killed as a result. So, Anna tried to kill her uncle and succeeded because of that very trying. But Anna did not kill her uncle intentionally.

And the problems with our principle don't stop there. Someone who tries and succeeds in winning the lottery by buying a ticket does not intentionally win the lottery. Likewise for cases of "successful" trying in the face of believed impossibility: a person who tries to push over a wall they thought they couldn't move may succeed, but they would not do so intentionally. Or again for cases of successful trying where one does not want to succeed: if the person who tries to hit the target by trying to hit the tree above it hits the tree above it, they have not done so intentionally.³⁴ There is more to intentional action than successful attempt.

What about a claim in the other direction? Does intentional action require an attempt?

INTENTIONAL → **TRYING** If S ϕ s intentionally, then S tried to ϕ .

At first sight it might seem plausible that anything you do intentionally is something you must have tried to do. Reflecting on paradigm cases of intentional action, it is hard to see how this connection between acting intentionally and trying could fail; indeed this connection has many prominent proponents.³⁵

But at the same time, considerations of the kind brought to bear by (among others) Harman (1976), Bratman (1984), and Knobe (2003, 2006) show that this principle does not hold in full generality. To take a case from Bratman, suppose I plan to run a marathon. I realize that by running it, I'll almost surely end up destroying my shoes, which happen to be a family heirloom. I don't want to destroy my shoes—they're a family heirloom!—but I love running and my family is long dead, so I treat the likely destruction of the shoes as a necessary evil. The time of the marathon comes. I run it and, lo and behold, my shoes are destroyed. I did not try to destroy my shoes, but I nonetheless destroyed them intentionally.

We will not attempt here to settle the question of how to distinguish this kind of intentional action from more paradigmatic cases. Instead, we will assume that the distinction can be understood in an intuitive way. We'll call the kind of intentional action involved in my destroying my shoes intentional in the *secondary* way, and the rest we'll call intentional in the *primary* way. Effectively by definition, then, my running the marathon was intentional in the primary way; my destroying my shoes was intentional in the secondary way. My arriving at the marathon's starting location was presumably intentional in the primary way. But my missing my lecture to run the marathon would have been intentional in the secondary way, assuming I merely saw it as a negative externality of my intended actions.

³³ For classic discussion, see, e.g., Chisholm (1971), Davidson (1973), Armstrong (1973), Harman (1976). More recent discussions include Mayr (2011, Ch. 6) and Shepherd (2021, Ch. 3).

³⁴ For this sentence and the previous, see also Hornsby (1995, p. 530), arguing against Ludwig (1992).

³⁵ See, e.g., O'Shaughnessy (1973, 1980), McCann (1975), Gorr (1979), Hornsby (1980, 2010).

With this distinction in hand, we offer the following as a more promising account of the relationship between intentional action and trying:

PRIMARY INTENTIONAL \rightarrow **TRYING** If S ϕ s intentionally in the primary way, then S tried to ϕ .

In these paradigmatic cases—of acting intentionally in the primary way—if one ϕ s intentionally, one tries to ϕ . But it is not just that one’s trying to ϕ merely coincides with ϕ -ing; one ϕ s intentionally in such cases *by trying to ϕ* . Indeed, **PRIMARY INTENTIONAL** \rightarrow **TRYING** seems to be in part motivated by this stronger idea:

BY TRYING If S ϕ s intentionally in the primary way, then S ϕ s by trying to ϕ .

To solidify the case in favor of these principles, we will consider an objection to them. The objector claims that people can do some simple bodily movements intentionally without trying to do them: maybe, for instance, moving their fingers or tapping their toes. In response, note that, for reasons that will be familiar from our discussion in §2, this kind of example is misleading. The fact that you can do these things *effortlessly* does not mean that you can do them without trying to (cf. Ludwig (2021, §6)). Indeed, for the moving of your fingers to be something you do rather than something that merely happens to you, plausibly there must be something you are trying to do, by which you move your fingers. Certainly in paradigmatic cases, if you move your fingers intentionally, you do so by trying to do that very thing. More generally, whatever answer we get to the question ‘How many people here moved their fingers intentionally?’, we do not expect to get a smaller one to the question ‘How many people moved their fingers by trying to move them?’. Outside of the special secondary cases, what we do intentionally we do by trying.

So trying underwrites what is done intentionally in the primary way. But what about trying itself: how often is it intentional (in the primary way)? We’ll now argue, a bit more speculatively, that it *always* is:³⁶

TRYING IS INTENTIONAL If S tries to ϕ , then S intentionally tries to ϕ in the primary way.

We endorse this principle more tentatively than **PRIMARY INTENTIONAL** \rightarrow **TRYING**, because our intuitive judgments about the intentionality of tryings are less than perfectly clear. But like **IDENTITY**, we think it deserves something of a presumptive status.

For starters, to the extent that we have intuitive judgments about the intentionality of tryings, the idea that one might unintentionally try to do something seems as bizarre as the idea that one might unintentionally choose to think about Confucius. And however strange it sounds to say that you intentionally tried to do something, it sounds even stranger to actively deny that your trying was intentional. If I were to ask you ‘Why did you try to ϕ ?’, I would be quite surprised to hear that you didn’t mean to, or that your trying was a mere accident. Similarly, it’s hard to see how there could be tryings that are intentional but not in the primary way. Under what conditions might you try to ϕ not as a means toward some end, say, but instead as a merely foreseen consequence of something else

³⁶ See also McCann (1974, 1975, 1986), McGinn (1982, ch. 8), Ginet (1990), Adams (1995), Adams (1997).

you do intentionally? These reflections incline us toward the view that trying is essentially intentional, that is, that TRYING IS INTENTIONAL holds.

Against all this, one might object to TRYING IS INTENTIONAL for reasons that came up in our discussion of SUCCESSFUL TRYING \rightarrow INTENTIONAL. There we argued that not all successful attempts count as intentional successes. One might think that the same points could apply to trying itself—that whatever “glue” is missing between Anna’s trying to kill her uncle and her successfully killing him could equally well have been missing between (say) Anna’s trying to try to kill her uncle and her successfully trying to kill him. But an objection like this can’t be given by someone who endorses IDENTITY. If IDENTITY is correct, then if one tries to try and thereby tries, this success cannot be coincidental or the product of a deviant causal chain: one’s trying to try *just is* one’s trying. It is thus difficult to see why someone would think IDENTITY holds while denying that TRYING IS INTENTIONAL does.

With this in mind, we now turn to drawing out three consequences of TRYING IS INTENTIONAL.

First, if PRIMARY INTENTIONAL \rightarrow TRYING and TRYING IS INTENTIONAL are both valid, then we get a new argument for ITERATION. For suppose S is trying to ϕ . By TRYING IS INTENTIONAL, it follows that S intentionally tries to ϕ in the primary way. By PRIMARY INTENTIONAL \rightarrow TRYING, if S intentionally tries to ϕ in the primary way, then S tried to try to ϕ . So S tried to try to ϕ .

Second, TRYING IS INTENTIONAL opens a new route to IDENTITY itself. TRYING IS INTENTIONAL says that whenever you try, you try intentionally. BY TRYING says that whenever you do something intentionally, you do it by trying to do it. It follows that you try by trying to try, that you try to try by trying to try to try, and so on. If IDENTITY were false, this package of commitments might seem to involve a problematic regress: how could my intentionally doing one thing—say raising my arm—happen only by way of my intentionally doing infinitely many other things—trying to raise my arm, trying to try to raise it, and so on? It’s a point in favor of IDENTITY that if it holds, the higher-order tryings are identical to the first-order trying, so that this regress would only be apparent. We may use different words at each stage, but they all describe the same thing.

Third, TRYING IS INTENTIONAL supports a stronger claim:

IDENTITY+ To try to ϕ just is to intentionally try to ϕ .

Clearly, if someone intentionally tries to ϕ , then they try to ϕ ; you can’t do something intentionally without doing it. So, given TRYING IS INTENTIONAL, it would follow that, necessarily, a person tries to ϕ if and only if they intentionally try to ϕ . It is some evidence in favor of IDENTITY+ that it would explain this necessary equivalence. But there’s also the fact that, given IDENTITY, attempts to try are guaranteed to be realized, and realized tryings are guaranteed to have been attempted. This suggests there’s nothing more to intentionally trying than trying—that doing one is simply a way of doing the other. We thus find it hard to see a principled position that accepts IDENTITY and TRYING IS INTENTIONAL but denies IDENTITY+.

The weaker IDENTITY already implies that trying to ϕ is identical to trying to...try to ϕ , where the ellipsis can be filled in with an arbitrary number of repetitions of ‘trying to’.

IDENTITY+ seems to imply a much stronger claim: that trying is identical to intentionally...intentionally trying to ϕ , where the ellipsis can be filled in with an arbitrary number of repetitions of ‘intentionally’.³⁷ If this is right, then the logical structure of intentional action differs in a surprising way from the logical structure of knowledge, at least on one prominent view of the latter. According to [Williamson \(2000\)](#), we cannot have arbitrarily many iterations of knowledge of non-trivial truths. Since trying to ϕ is not trivial, if the above arguments are correct, then ‘intentionally’ can iterate indefinitely on non-trivial doings, yielding a sharp difference between knowledge and intentional action.³⁸

How should we understand this contrast? A key principle underlying Williamson’s view that we only ever have finitely many iterations of knowledge of non-trivial truths is that a person’s belief amounts to knowledge only if it is not true that they very nearly could have had a relevantly similar false belief. But whatever the status of this principle (or the various refinements of it which have since been offered), there seems to be no analogous constraint on intentional action. That is, it is not true that, if one’s relevant attempt succeeds, but one very easily could have made a relevantly similar failed attempt, then one does not act intentionally.

To see the difference, suppose an excellent tennis player is about to hit a second serve.³⁹ They are an extremely reliable second server, missing less than 1% of the time. If conditions are normal and they get the serve deep inside the service box, then it is plausible that they could know in advance that they will make their serve.⁴⁰ But if they happen to barely make their serve—say because they just barely catch the line—it is less plausible that they could know that they would succeed in advance. Whether the player knows they’ll make their second serve can depend on how close they come to missing it.

By contrast, whether they make their serve intentionally is not in the same way sensitive to where the ball lands in the box. Suppose for whatever reason the player doesn’t get to see where exactly their serve landed, but is informed that it landed in. If asked ‘Did you get your serve in intentionally?’, the player cannot reasonably respond ‘I’m not sure; it depends on where it landed’. Nor would seeing that the serve just barely caught the line cause the

³⁷ Depending on how one handles subtle questions related to Frege’s puzzle, this claim may not hold in full generality. It might be that people who do not realize that trying is identical to intentionally trying could intentionally try to ϕ , but fail to intentionally intentionally try to ϕ , in spite of the fact that trying to ϕ is identical to intentionally trying to ϕ , in much the same way that one might intentionally hug Superman, but not intentionally hug Clark, even if hugging Superman is hugging Clark. Even if that’s so, the result in the main text would still hold for ideally rational agents who do recognize the truth of IDENTITY+.

³⁸ In an important paper, [Beddor & Pavese \(forthcoming\)](#) argue that Anscombe’s category of essentially intentional actions is empty, a claim which would entail that both TRYING IS INTENTIONAL and IDENTITY+ are invalid (see also [Piñeros Glasscock \(2019\)](#)). Their argument relies on the premise that intentionally ϕ -ing entails knowing that one is ϕ -ing, and then invokes an argument like Williamson’s that no one can have arbitrarily many iterations of knowledge. But we are persuaded by a number of the counterexamples others have presented to Beddor and Pavese’s premise that ϕ -ing intentionally requires knowing that one is ϕ -ing (e.g., [Davidson \(1971\)](#), [Setiya \(2008\)](#), [Paul \(2009\)](#), and [Piñeros Glasscock \(2019\)](#), with [Paul \(2020, Ch. 6\)](#) for a survey). We believe that someone under anesthesia can intentionally clench their fist even if they’re highly unsure whether they have control over their hands, that doctors intentionally save their patient’s lives when they successfully perform operations that they know have a high chance of failure, and that professional tennis players intentionally ace their opponents despite not knowing whether their serves are in until the umpire makes a call.

³⁹ [Shepherd & Carter \(2021\)](#) develop an array of cases like this one, offer more sustained defense of the claim that the actions are intentional and provide a detailed analysis of the contrast in permissiveness between knowledge and intentional action. See also [Carter & Shepherd \(2022\)](#).

⁴⁰ Here we are assuming that epistemic skepticism about the near future is false—that we can in most situations know what we’ll be eating for dinner tonight, where we’ll be sleeping, how we’ll be getting home, and so on.

player to deny that they had hit their serve intentionally. In this sense, whether an action is intentional is different from whether it was known in advance that it would succeed.

Examples like this one are of course controversial. Some believe even in the face of such examples that intentional action requires knowledge. We have not tried to mount a sustained challenge against this distinguished tradition here.⁴¹ Our aim has been more modest: to illustrate how those with certain views about knowledge may have reason to reject an epistemic constraint on intentional action, since it conflicts with the idea that trying is intentional. If one rejects this constraint, then the contrast between ‘intentionally’ and ‘know’ is easy to explain. Whereas knowledge is fragile, intentional action is robust. Whereas knowledge does not allow near-failure, intentional action does. And as a result, whereas knowledge cannot iterate indefinitely, intentional action can.

9 Trying without fail

Finally, we turn to some connections between the theory of trying we’ve developed—on which both IDENTITY and TRYING IS INTENTIONAL hold—and issues in the theory of rational choice and the metaphysics of action.

We have argued that trying to ϕ is itself agentially perfect. This claim entails that if one tries to try to ϕ , one will succeed in trying to ϕ . But it does not guarantee that one will be in a position to try to try to ϕ , and in general one may not be.⁴² For example, one might be unable to try to ϕ because a malevolent neuroscientist has implanted a certain kind of chip in one’s brain that causes a fatal stroke the moment it detects whatever neural events would normally precede an event of one’s trying to ϕ (cf. e.g. [Storrs-Fox \(n.d.\)](#), [Koon \(2020\)](#), building on [Frankfurt \(1969\)](#)). Or one might be unable to try to ϕ because one isn’t in a position to grasp ϕ -ing under any relevant guise. For example, for some sufficiently large prime number n —which contains, say, more digits than there are atoms in the universe—the action of *proving that n is prime* is plausibly something one cannot even try to do; one might simply lack any relevant guise under which one can apprehend the question of whether it is prime. Agential perfection does not suffice for ability.

Similarly, to say that ϕ -ing is agentially perfect is also not to say that necessarily, ϕ -ing is always *under a person’s control*. Whether something is under a person’s control often depends on the nature of the motivational states that explain their doing it. When a person’s motivational states are the result of forces that are sufficiently alien to them—whether coercion, brainwashing, manipulative neurosurgery, psychosis, or what have you—it will often not be true that what they did was under their control. But the same is not true for

⁴¹ See, e.g., [Anscombe \(1957\)](#), [Thompson \(2011\)](#), [Pavese \(2020, forthcoming\)](#), and [Beddor & Pavese \(forthcoming\)](#). In addition to what we’ve discussed above, our approach requires denying the principle: if a person tries to ϕ intentionally, the person knows that they are trying to ϕ . The cases we’ve given don’t present *intuitive* counterexamples to this principle. But it is hard to see how anyone independently attracted to the anti-luminosity argument would accept this principle. Indeed, someone who accepted this and TRYING IS INTENTIONAL would be committed to the luminosity of trying itself.

⁴² [Lehrer \(1968\)](#) famously suggests that a person could be so revolted at the sight of a red candy, that they may be unable even to try to put it in their mouth, even though, if they were to try to try to put the candy in their mouth, they would thereby succeed in trying to do this (and doing it). But in our view, the person *can* still try to eat the candy, for instance, by trying to overcome their revulsion in order to eat it.

trying: someone who is brainwashed might try to do all sorts of things they would never have tried to do had they not been brainwashed. It does not follow that they tried without trying to try, or that they could have tried to try without trying. What follows is that it wasn't under their control what they tried to do. Agential perfection does not suffice for control.

Although agential perfection does not entail ability or control, agential perfection is important in its own right. We'll now tentatively offer two suggestions for domains where it could have some use, in the hope that others may pick up where we have left off.

First, normative decision theory standardly appeals to a set of *options* as the basis for its verdicts about what a person ought (subjectively) to do. But what are these options? A number of constraints on an answer have been proposed. Among them is the idea that options must be essentially successful (see, e.g., Pollock (2002, e.g. p. 13)).⁴³ The intuitive motivation for this thought comes, first, from the idea that decision-theoretic options should be things that the agent in question cannot fail to do, and, second, from the idea that to fail to do something one must try to do it. Given these assumptions, if ϕ -ing is essentially successful, no one can (in the target sense) fail to ϕ . We don't have space to explore the arguments for and against either of these ideas here. But those who find them attractive have reason to endorse the claim that trying has a special significance for decision theory, since trying is agentially perfect and thus essentially successful.⁴⁴

Indeed, the overall picture we've developed—on which trying itself is essentially successful—strengthens the idea that the property of being essentially successful is important for decision theory. Why? Well, suppose ϕ -ing was essentially successful but that trying to ϕ was not. Then even if one could not fail to ϕ in the sense described above, it might be that one could in the relevant sense still fail to *try to* ϕ . If this were possible, then in a case where one's only means of ϕ -ing was trying to ϕ , a person might try to try to ϕ , but still fail to try to ϕ , and thus not actually ϕ . Such cases won't be ones where the person *tried and failed to* ϕ , but they still would be cases where the person did not ϕ . So they would naturally raise the question of why we should care about the fact that ϕ -ing was something we could “do without fail” in the sense characterized above. But if trying is essentially successful, then this problem does not arise. For then essentially successful doings will be things we do without fail, not just because trying to do them suffices for doing them, but also because trying to try to do them suffices for trying to do them, and so on. If trying is essentially successful, then that which is essentially successful is that which we can do without fail, in the fullest sense of the phrase.

Second, a central notion in the theory of action is that of a 'basic action'. The correct characterization of this notion is contested. But on one natural way of understanding it, the idea is that a particular doing is basic if and only if the person does it intentionally (in

⁴³ Other proposed constraints include that options are things one is able to do (Pollock (2002); Hedden (2012, 2015); Schwarz (2021)), or things that one is able to do and certain that one is able to do (Koon (2020)), or even things that are perfectly under one's control (Storrs-Fox (n.d.)).

⁴⁴ Hedden (2015, p. 441) suggests briefly that options are tryings, but his official terminology quickly shifts to decisions, and the fact that he is thinking of decisions rather than trying plays a load-bearing role in his subsequent discussion, where he connects the relevant events to intention-formation. As discussed earlier, trying does not require intending. Also, deciding plausibly satisfies a version of BELIEVE MIGHT, so even those who hold that trying is agentially perfect should reject the claim that deciding is.

the primary way) without having to do anything else.⁴⁵ That is to say: a person's ϕ -ing on an occasion is basic iff: (i) the person ϕ s on that occasion intentionally in the primary way and (ii) for all ψ , if the person ϕ s on that occasion by ψ -ing, then their ϕ -ing on that occasion is identical to their ψ -ing on that occasion. We believe those attracted to such a conception of basic action have reason to accept IDENTITY. The reason is that basic doings are intentional in the primary way, and so—as BY TRYING tells us—one must perform them by trying to perform them. But a particular doing is basic only if one does it intentionally and not by doing something else. So either there are no basic actions in the above sense (because no doing is identical to its own trying) or basic actions are identical to their own tryings. It follows immediately that if there are basic actions, they are tryings. What's more, by accepting IDENTITY, we get a natural story about how there could be such basic actions: to be a basic action is in part to be an intentional doing that is identical to its own trying; and by IDENTITY, tryings are such doings. So anyone who is independently attracted to the above conception of basic action has reason to accept IDENTITY. And conversely, the harmonious way in which IDENTITY, TRYING IS INTENTIONAL, and BY TRYING fit together lends credence to a trying-theoretic conception of basic action.

10 Conclusion

Earlier we argued that the agential perfection of trying may have important implications for our understanding of the relationship between knowledge and intentional action. We want to close by developing a broader but more speculative analogy between epistemology and the philosophy of action.

There is a natural analogy between belief and knowledge, on the one hand, and trying and intentional action, on the other.⁴⁶ Trying is plausibly necessary for (primary) intentional action in the way that belief is plausibly necessary for knowledge.⁴⁷ Likewise, one's environment imposes fewer constraints on what one can try to do than on what one can do intentionally—in much the same way that (setting semantic externalism to one side) one's environment imposes fewer constraints on what one can believe than it does on what one can know. If one is unknowingly paralyzed, many of one's attempts will fail to manifest themselves in intentional action; if one is unknowingly envatted, many of one's beliefs will fail to constitute knowledge.

The analogy extends to questions about conceptual priority as well. Is knowledge to

⁴⁵ For helpful discussion see, e.g., see, Danto (1965), Goldman (1970), Davidson (1971), Hornsby (1980, Ch. 5-6), Thompson (2008), Lavin (2013).

⁴⁶ Williamson (2017) argues that *intention* (not trying) stands to acting as belief stands to knowledge. But intentions can fail to lead to action because the person changes their mind before it is time to act (or otherwise because they never try), while beliefs have no analogous failure mode. For related reasons, few hold that successful intentions are identical to intentional actions, but just as it is a popular view that appropriately successful beliefs are identical to states of knowledge, it is a popular view that at least some successful tryings are identical to intentional actions.

⁴⁷ It is controversial whether it is possible to know something without believing it is true. But it is not at all controversial that it is possible to perform a certain action without having tried to perform it. Accidental actions are a paradigm case of this. And if theorists like Knobe are right, it may even be possible to intentionally perform an action without having tried to, as when one intentionally scuffs up one's shoes by running a marathon. Still, whenever you act you are trying to do *something*, even if that something isn't the action you end up performing. It is in that sense that trying seems necessary for action.

be understood in terms of belief, or is belief to be understood in terms of knowledge (or is neither reducible to the other)? Is intentional action to be understood in terms of trying, or is trying to be understood in terms of intentional action (or neither)?

Post-Gettier epistemology sees few who take seriously the prospects of offering a non-trivial analysis of knowledge in terms of belief. But a good deal of the existing literature in the philosophy of action treats the question of whether trying or acting intentionally is prior as live. Among those who take trying to be primary, some claim that intentional actions are events that are caused by tryings in the right way (Armstrong 1973). And others claim that what makes an event an action is that it is a trying, or at least an event that can truly be described as one (Hornsby 1980, Pietroski 2000). Among those who take action to be the primary notion, some claim that trying is doing what one takes to be necessary to perform a certain action (O’Shaughnessy 1973). Others claim that to try to ϕ is to be such that if one were in the right circumstances, one would succeed in ϕ -ing (Ruben 2016, 2018). And yet others claim that to try to ϕ is to do something with the intention of bringing it about that one ϕ s (Ludwig 2021).

We have tried to sidestep this debate. We have argued that trying is necessary for (primary) intentional action, and that those sympathetic to certain concepts of basic action have reason to understand the notion in terms of trying. But we’ve remained essentially silent on the question of whether this means that intentional action itself can be analyzed in terms of trying, or vice-versa. Our own suspicion is that the notions of action and trying are likely both “prime”: that there is no non-trivial analysis of intentional action in terms of trying (or other notions);⁴⁸ and that there is no non-trivial analysis of trying in terms of intentional action (or other notions) either.

In epistemology, even those who think that belief is a kind of “defective knowledge” recognize that it may be subject to rich structural laws of its own—that many substantive questions about its nature and norms are both unobvious and of great theoretical interest. We hope to have shown that the same is true of trying, even for those who think of it as a kind of “defective intentional action”. Trying is subject to rich structural laws of its own, and questions about its nature and norms are worth studying in their own right.

References

- Adams, Frederick. 1986. Intention and Intentional Action: The Simple View. *Mind and Language*, 1(4), 281–301.
- Adams, Frederick. 1991. He Doesn’t Really Want to Try. *Analysis*, 51(2), 109–112.
- Adams, Frederick. 1994a. Of Epicycles and Elegance. *Canadian Journal of Philosophy*, 24(4), 637–641.
- Adams, Frederick. 1994b. Trying, Desire, and Desiring to Try. *Canadian Journal of Philosophy*, 24(4), 613–626.
- Adams, Frederick. 1995. Trying: You’ve Got to Believe. *Journal of Philosophical Research*, 20, 549–561.
- Adams, Frederick. 1997. Cognitive Trying. *Pages 287–314 of: Holmström-Hintikka, Ghita, & Tuomela, Raimo (eds), Contemporary Action Theory Volume 1: Individual Action*. Springer Netherlands.

⁴⁸ Here we see ourselves as aligned in different ways with Anscombe (1957), Thompson (2008), Levy (2013), and Ford (2017).

- Adams, Frederick. 2007. Trying with the Hope. In: Timmons, Mark, Greco, John, & Mele, Alfred R. (eds), *Rationality and the Good: Critical Essays on the Ethics and Epistemology of Robert Audi*. Oxford University Press.
- Adams, Frederick, & Mele, Alfred R. 1992. The Intention/Volition Debate. *Canadian Journal of Philosophy*, **22**(3), 323–337.
- Albritton, Rogers. 1985. Freedom of the Will and Freedom of Action. *Proceedings and Addresses of the American Philosophical Association*, **59**(2), 239–51.
- Anscombe, G. E. M. 1957. *Intention*. Harvard University Press.
- Armstrong, D. M. 1973. Acting and Trying. *Philosophical Papers*, **2**(1), 1–15.
- Beddor, Bob, & Pavese, Carlotta. forthcoming. Practical Knowledge Without Luminosity. *Mind*, 1–23.
- Blumberg, Kyle, & Hawthorne, John. 2022. *Kraemer's Puzzle and the Theory of Intentional Action*. Unpublished manuscript.
- Bratman, Michael. 1984. Two Faces of Intention. *Philosophical Review*, **93**(3), 375–405.
- Bratman, Michael. 1987. *Intention, Plans, and Practical Reason*. Cambridge, MA: Harvard University Press.
- Buckwalter, Wesley, Rose, David, & Turri, John. forthcoming. Impossible Intentions. *American Philosophical Quarterly*.
- Butler, Ronald J. 1978. Report on Analysis "Problem" No. 16. *Analysis*, **38**(3), 113–114.
- Carter, J. Adam, & Shepherd, Joshua. 2022. Intentional Action and Knowledge-Centred Theories of Control. *Philosophical Studies*, 1–21.
- Chisholm, Roderick. 1971. On the Logic of Intentional Action. *Pages 38–80 of: Binkley, Robert Williams, Bronaugh, Richard N., & Marras, Ausonio (eds), Agent, Action, and Reason*. University of Toronto Press.
- Danto, Arthur C. 1965. Basic Actions. *American Philosophical Quarterly*, **2**(2), 141–148.
- Davidson, Donald. 1971. Agency. *Pages 1–37 of: Marras, Ausonio, Bronaugh, R. N., & Binkley, Robert W. (eds), Agent, Action, and Reason*. University of Toronto Press.
- Davidson, Donald. 1973. Freedom to Act. *Essays on Actions and Events*, 63–81.
- Davis, Wayne A. 1984. The Two Senses of Desire. *Philosophical Studies*, **45**(2), 181–195.
- Descartes, Renée. 1985-1992. *The Philosophical Writings of Descartes*. Cambridge: Cambridge University Press. trans. John Cottingham, Robert Stoothoff, Dugald Murdoch, and Anthony Kenny.
- Descartes, Renée. 1996. *Oeuvres de Descartes*. Paris: J. Vrin. ed. Charles Adam and Paul Tannery.
- Ford, Anton. 2017. The Representation of Action. *Royal Institute of Philosophy Supplement*, **80**, 217–233.
- Frankfurt, Harry. 1969. Alternate Possibilities and Moral Responsibility. *Journal of Philosophy*, **66**(23), 829.
- Ginet, Carl. 1990. *On Action*. Cambridge, England: Cambridge University Press.
- Goldman, Alvin. 1970. *A Theory of Human Action*. Englewood Cliffs, NJ, USA: Princeton University Press.
- Gorr, Michael. 1979. Willing, Trying and Doing. *Australasian Journal of Philosophy*, **57**(3), 237–250.
- Grano, Thomas. 2011. Mental action and event structure in the semantics of 'try'. *Semantics and Linguistic Theory*, **21**, 426–443.
- Grano, Thomas. 2017. Control, temporal orientation, and the cross-linguistic grammar of trying. *Glossa*, **2**(1).
- Grice, Paul. 1989. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.
- Hampshire, Stuart. 1959. *Thought and Action*. London: Chatto Windus.
- Harman, Gilbert. 1976. Practical Reasoning. *The Review of Metaphysics*, **29**(3), 431–463.
- Harman, Gilbert. 1986. Willing and Intending. *Pages 363–380 of: Grandy, Richard, & Warner, Richard (eds), Philosophical grounds of rationality: Intentions, categories*. Clarendon Press Oxford.
- Heath, Peter. 1971. Trying and Attempting. *Aristotelian Society Supplementary Volume*, **45**(1), 193–227.
- Hedden, Brian. 2012. Options and the Subjective Ought. *Philosophical Studies*, **158**(2), 343–360.

- Hedden, Brian. 2015. Options and Diachronic Tragedy. *Philosophy and Phenomenological Research*, **90**(2), 423–451.
- Hedman, Carl G. 1970. Intending the Impossible. *Philosophy*, **45**(171), 33–38.
- Hermes, Charles. 2006. Does Attempting to Try to a Imply Trying to A? *Southwest Philosophy Review*, **22**(2), 63–70.
- Hornsby, Jennifer. 1980. *Actions*. Routledge and Kegan Paul.
- Hornsby, Jennifer. 1995. Reasons for Trying. *Journal of Philosophical Research*, **20**, 525–539.
- Hornsby, Jennifer. 2010. Trying to Act. *Pages 18–25 of: O'Connor, T., & Sandis, C. (eds), A Companion to the Philosophy of Action*. Blackwell Companions to Philosophy Series. Wiley.
- Hunter, J. F. M. 1987. Trying. *Philosophical Quarterly*, **37**(149), 392–401.
- Jones, O. R. 1983. Trying. *Mind*, **92**(367), 368–385.
- Knobe, Joshua. 2003. Intentional Action in Folk Psychology: An Experimental Investigation. *Philosophical Psychology*, **16**(2), 309–325.
- Knobe, Joshua. 2006. The Concept of Intentional Action: A Case Study in the Uses of Folk Psychology. *Philosophical Studies*, **130**(2), 203–231.
- Koon, Justis. 2020. Options Must Be External. *Philosophical Studies*, **177**(5), 1175–1189.
- Kraemer, Eric Russert. 1978. Intentional Action, Chance and Control: [Analysis "Problem" No. 16]. *Analysis*, **38**(3), 116–117.
- Kriegel, Uriah. 2015. *The varieties of consciousness*. Oxford: Oxford University Press.
- Lavin, Douglas. 2013. Must There Be Basic Action? *Noûs*, **47**(2), 273–301.
- Lehrer, Keith. 1968. Cans Without Ifs. *Analysis*, **29**(1), 29–32.
- Levy, Yair. 2013. Intentional Action First. *Australasian Journal of Philosophy*, **91**(4), 705–718.
- Ludwig, Kirk. 1992. Impossible Doings. *Philosophical Studies*, **65**(3), 257–281.
- Ludwig, Kirk. 1995. Trying the Impossible: Reply to Adams. *Journal of Philosophical Research*, **20**, 563–570.
- Ludwig, Kirk. 2021. Let Me Go and Try. *Philosophical Explorations*, **24**(3), 340–358.
- Mandelkern, Matthew, Schultheis, Ginger, & Boylan, David. 2017. Agentive Modals. *Philosophical Review*, **126**(3), 301–343.
- Mayr, Erasmus. 2011. *Understanding Human Agency*. Oxford University Press.
- McCann, Hugh. 1974. Volition and Basic Action. *Philosophical Review*, **83**(4), 451–473.
- McCann, Hugh. 1975. Trying, Paralysis, and Volition. *Review of Metaphysics*, **28**(3), 423–442.
- McCann, Hugh J. 1986. Rationality and the Range of Intention. *Midwest Studies in Philosophy*, **10**(1), 191–211.
- McCann, Hugh J. 1989. Intending and Planning: A Reply to Mele. *Philosophical Studies*, **55**(1), 107–110.
- McCann, Hugh J. 1991. Settled Objectives and Rational Constraints. *American Philosophical Quarterly*, **28**(1), 25–36.
- McCormick, Suzanne, & Thalberg, Irving. 1967. Trying. *Dialogue*, **6**(1), 29–46.
- McGinn, Colin. 1982. *The Character of Mind*. Oxford University Press.
- Mele, Alfred. 1992a. Review of *On Action*, by Carl Ginet. *Philosophy and Phenomenological Research*, **52**(2), 488–491.
- Mele, Alfred R. 1990. He Wants to Try. *Analysis*, **50**(4), 251–253.
- Mele, Alfred R. 1991. He Wants to Try Again: A Rejoinder. *Analysis*, **51**(4), 225–228.
- Mele, Alfred R. 1992b. *Springs of Action: Understanding Intentional Behavior*. Oxford: Oxford University Press.
- Mele, Alfred R. 1994. Desiring to Try: Reply to Adams. *Canadian Journal of Philosophy*, **24**(4), 627–636.
- Nadelhoffer, T. 2004. The Butler Problem Revisited. *Analysis*, **64**(3), 277–284.
- O'Shaughnessy, Brian. 1973. Trying (As the Mental "Pineal Gland"). *Journal of Philosophy*, **70**(13), 365–386.
- O'Shaughnessy, Brian. 1980. *The Will: A Dual Aspect Theory*. Cambridge University Press.
- Paul, Sarah. 2020. *Philosophy of Action: A Contemporary Introduction*. Routledge.

- Paul, Sarah K. 2009. Intention, Belief, and Wishful Thinking: Setiya on ?Practical Knowledge? *Ethics*, **119**(3), 546–557.
- Pavese, Carlotta. 2020. Probabilistic Knowledge in Action. *Analysis*, **80**(2), 342–356.
- Pavese, Carlotta. forthcoming. Practical Knowledge First. *Synthese*, 1–17.
- Peacocke, Christopher. 2004. *The realm of reason*. Oxford: Oxford University Press.
- Peacocke, Christopher. 2006. Mental Action and Self-Awareness. In: Cohen, Jonathan D., & McLaughlin, Brian P. (eds), *Contemporary Debates in the Philosophy of Mind*. Blackwell.
- Peacocke, Christopher. 2009. Mental Action and Self-Awareness : Epistemology. In: O'Brien, Lucy, & Soteriou, Matthew (eds), *Mental Actions*. Oxford University Press.
- Pietroski, Paul M. 2000. *Causing Actions*. Oxford University Press.
- Piñeros Glasscock, Juan S. 2019. Practical Knowledge and Luminosity. *Mind*, **129**(516), 1237–1267.
- Pollock, John L. 2002. Rational Choice and Action Omnipotence. *Philosophical Review*, **111**(1), 1–23.
- Ross, David. 1978. He Loads the Gun, Not the Dice: [Analysis "Problem" No. 16]. *Analysis*, **38**(3), 114–115.
- Ruben, David-Hillel. 2016. A Conditional Theory of Trying. *Philosophical Studies*, **173**(1), 271–287.
- Ruben, David-Hillel. 2018. *The Metaphysics of Action: Trying, Doing, Causing*. London: Palgrave Macmillan.
- Schroeder, Severin. 2001. The Concept of Trying. *Philosophical Investigations*, **24**(3), 213–227.
- Schwarz, Wolfgang. 2021. Objects of Choice. *Mind*, **130**(517), 165–197.
- Setiya, Kieran. 2008. Practical Knowledge. *Ethics*, **118**(3), 388–409.
- Sharvit, Yael. 2003. Trying to Be Progressive: The Extensionality of Try. *Journal of Semantics*, **20**(4), 403–445.
- Shepherd, Joshua. 2016. Conscious Action/Zombie Action. *Noûs*, **50**(2), 419–444.
- Shepherd, Joshua. 2017. The Experience of Acting and the Structure of Consciousness. *Journal of Philosophy*, **114**(8), 422–448.
- Shepherd, Joshua. 2021. *The Shape of Agency: Control, Action, Skill, Knowledge*. Oxford: Oxford University Press.
- Shepherd, Joshua, & Carter, J. Adam. 2021. Knowledge, Practical Knowledge, and Intentional Action. *Ergo: An Open Access Journal of Philosophy*.
- Storrs-Fox, David. *Non-Cartesian Agency and Normative Theory*. Unpublished manuscript.
- Thalberg, I. 1962. Intending the Impossible. *Australasian Journal of Philosophy*, **40**(1), 49–56.
- Thompson, Michael. 2008. Naive Action Theory. In: Thompson, Michael (ed), *Life and Action*. Harvard University Press.
- Thompson, Michael. 2011. Anscombe's Intention and Practical Knowledge. In: Ford, Anton, Hornsby, Jennifer, & Stoutland, Frederick (eds), *Essays on Anscombe's Intention*. Harvard University Press.
- Williamson, Timothy. 2000. *Knowledge and Its Limits*. Oxford University Press.
- Williamson, Timothy. 2017. Acting on Knowledge. *Pages 163–181 of: Carter, J. Adam, Gordon, Emma C., & Jarvis, Benjamin W. (eds), Knowledge First: Approaches in Epistemology and Mind*. Oxford: Oxford University Press.
- Wittgenstein, Ludwig. 1950. *Philosophical Investigations*. London: Macmillan.
- Yaffe, Gideon. 2004. Trying, Intending, and Attempted Crimes. *Philosophical Topics*, **32**(1/2), 505–531.
- Yaffe, Gideon. 2010. *Attempts: In the Philosophy of Action and the Criminal Law*. Oxford University Press.