Abilites de dicto and de re

Wolfgang Schwarz
Draft, 7 April 2018

Abstract. Many ability statements display a kind of de re/de dicto ambiguity. The classical (Kratzerian) semantics of modals only captures the de re reading. To derive the de dicto reading, which plays a central role in normative contexts, I suggest that that context can impose not only restrictions on the accessible worlds, but also on what counts as a performance of the relevant act. I give some independent arguments for this assumption and show that it provides an easy response to many objections that have been raised against the classical semantics.

Contents

1 Introduction

If you couldn't have prevented an outcome, you are not responsible for it. The President can't fire the Special Counsel. Tardigrades can survive temperatures of 150°C. Judgements about what someone can or cannot do play an important role in our conception of the world. What exactly is required for an agent to have a given ability? Many answers have been put forward, but they all seem to face serious problems. I will argue that a key to an adequate analysis of ability statements lies in a widespread, but often subtle and therefore neglected ambiguity.

Let me begin with a few examples. Suppose Eddie's favourite piano piece is the Moonlight Sonata. Maisy can play the Moonlight Sonata, but she doesn't know that it is Eddie favourite piece. Now consider (1).

(1) Maisy can play Eddie's favourite piece.

True or false? Intuitively, the statement is true in one sense and false in another. We have a kind of de re/de dicto ambiguity: Maisy can play what is in fact Eddie's favourite piece, but she can't play it "under that description".

The de dicto reading, on which (1) is false, roughly fits G.E. Moore's [1912] conditional analysis, according to which an agent can ϕ just in case she would ϕ if she tried to ϕ. If Maisy tried to play Eddie's favourite piece, she would most likely fail.

The conditional analysis has been influential in ethics and debates about free will. Meanwhile, in linguistics and philosophy of language, the orthodox account is the quantificational analysis originating in modal logic and influentially applied to natural language by Angelika Kratzer [1977, 1981, 1991, 2012]. On that account, modal expressions like 'must' and 'can' function as quantifiers over a contextually restricted set of possible worlds. Thus 'S can ϕ' is true just in case there are "accessible" worlds where S ϕs. In contrast to the conditional analysis, which seems to target the de dicto reading of sentences like (1), the quantificational analysis analysis plausibly targets their de re reading. To see why, note that Maisy can unambiguously play the Moonlight Sonata. So the accessible worlds should include worlds where she plays the Moonlight Sonata. Unless Eddie's musical preferences are mysteriously different in all these worlds, it follows that there are accessible worlds where Maisy plays Eddie's favourite piece. So (1) comes out true. (We will return to this line of thought.)

Another example. Assume Cyril does not know the first 10 digits of π. On its most natural reading, (2) is then false.

(2) Cyril can recite the first 10 digits of π.

But the sentence also has a true reading, assuming that Cyril can utter the words 'three', 'one', 'four', etc. If he can utter these words in the right order, he can do something that amounts to reciting the first 10 digits of π. So he can perform the act de re. But he can't perform it de dicto; he can't recite the first 10 digits of π "under that description".

In (2), the ambiguity is easy to miss because the de dicto reading is much more salient than the de re reading. Plausibly, that's because the de re is trivial. It is hardly worth pointing out that someone can utter the words 'three', 'one', 'four', etc. If a sentence has a trivial and a non-trivial reading, the non-trivial reading usually dominates. By comparison, (1) is non-trivial on either reading: most of us can utter 'three', 'one', 'four', etc., but few can play the Moonlight Sonata.

At the opposite end of the spectrum lie statements like (3).

(3) Usain Bolt can run 100 meters in 9.58 seconds.

This statement is most naturally understood de re. The claim is that Usain Bolt can do something that amounts to running 100 meters in 9.58 seconds – a highly non-trivial feat. The claim is not that Bolt is able to perform the relevant act "under that description",

which would presumably require that he can at will choose to run with an average speed of exactly 0.0958 m/s.

One more example, from [Carlson 1999]. Charley wants to open a safe, but doesn't know the combination. The combination is 448-961-5237, and Charley can dial that. Consequently, (4) is true in one sense (de re), and false in another (de dicto).

(4) Charley can open the safe.

Here again, the de dicto interpretation is more prominent, for the same reason as in (2). (4) illustrates that the two readings can't always be traced to the scope ambiguity of some definite description. One might have thought that the de re reading of (2) arises from 'the first 10 digits of π' taking wide scope over the modal 'can'; similarly for 'Eddie's favourite piece' in (1). But (4) doesn't contain a relevant description ('the safe' doesn't help); yet it intuitively displays the same ambiguity as (1) and (2). It's as if the entire verb phrase 'open the safe' can scope over the modal, so that (4) asserts *of opening the safe* that Charley can do *it*.

I will continue to use the mnemonic labels 'de dicto' and 'de re' for the two readings of statements like (1)–(4), without thereby assuming a particular analysis or syntactic derivation. What the two readings are, and how they come about, is the topic of the following sections.

Above I asked what it takes for an agent to have a given ability. In light of the examples just reviewed, we really need to ask two questions: What does it take to have an ability de re? And, what does it take to have an ability de dicto? We can hardly expect both questions to have the same answer. No wonder uniform analyses of ability statements all run into problems.

In the next section, I will give a preliminary answer to both questions, tailored to a certain use of ability statements in normative contexts. In the remainder of the paper, I will defend a variant of Kratzer's quantificational account that allows deriving both of these readings. In section 3, I will review the basic ideas behind the quantificational approach and defend it against certain misunderstandings. In section 4, I will explain how the approach could be tweaked to derive the de dicto reading of ability statements. In essence, the proposal is that when we talk about whether an agent has an ability to ϕ, context can impose a restriction not just on the accessible worlds, but also on what counts as a relevant way of ϕing. For a de dicto ability, the agent must have the ability to perform the relevant act "transparently" or "intentionally", rather than by sheer luck. As we will see, the mechanism that gives rise to the de dicto reading plausibly also explains some other puzzles about ability statements, such as their apparent gradability. Finally, in section 5, I will show that some prominent objections to the Kratzerian model involve de dicto abilities and are easily answered by the tweaked version of the model developed in section 4.

2 Oughts and cans

So, what does it take for an agent to have an ability de re? And what does it take to have an ability de dicto? Let's start with the de re reading, on which (1)–(4) are true. What is the sense in which Maisy *can* play Eddie's favourite piece?

As we saw, it's not that Maisy would succeed if only she tried. But a related idea seems to be on the right track. Maisy would play Eddie's favourite piece if she tried to play the Moonlight Sonata. That is, there is *something* such that if Maisy tried to do *it*, then she would play Eddie's favourite piece.

Let's say that an act is *under an agent's volitional control* if there is a possible variation of her volitional state – of what she wants or intends to do – that would cause her to perform the act. Arguably, (1) is true insofar as playing Eddie's favourite piece is under Maisy's volitional control. Similarly for (2) and (4). If Cyril wanted to utter the words 'three', 'one', 'four', etc., he would end up reciting the first 10 digits of π; if Charley wanted to dial 448-961-5237, he would end up opening the safe.

So here is my first stab at a general analysis of abilities de re:

> **Analysis 1a.**
>
> S can ϕ de re iff there is a possible variation of S's volitional state that would cause S to ϕ.

The restriction to "possible" variations is meant to deal with cases where an agent can't get into the relevant volitional state. We don't want to say that a person in a coma can get out of bed, even though she would presumably succeed in getting out of bed if she wanted to do so, because then she would not be in a coma (compare [van Inwagen 1983: 119]). For agents in a coma, wanting to get out of bed is therefore not a "possible" variation of their volitional state. (I will have more to say on this matter soon.)

Analysis 1a is not fully general. Among other things, it does not account for the ability of Tardigrades to survive temperatures of 150℃, which plausibly does not require any relevant volitions. It also does not account for so-called "general" ability statements like (3). In the next section, I will explain how the analysis should be generalised to handle these cases. For now, I want to focus on ability statements like (1), (2), and (4) that concern voluntary acts of (moderately) rational agents in a given choice situation.

Analysis 1a is motivated by a certain picture, or model, of rational agency. On that model, voluntary acts are consequences of corresponding intentions or volitions. When agents deliberate about what to do and finally reach a decision, the direct result is not an overt act, but a volitional state. Which volitional state comes about depends on the agent's desires, values, beliefs, and on the process of deliberation. Other values, other beliefs, or other deliberational processes may have led to different volitional states and consequently different acts.

Now suppose we ask what an agent *ought* to do in a given choice situation. Normative judgements are generally relative to norms and facts: in light of such-and-such norms and such-and-such facts, the agent ought to ϕ. To say that an agent ought to ϕ presupposes that ϕing is in some sense an available option for the agent, but it does not presuppose that the agent actually chooses to ϕ. After all, the agent may not know the relevant facts, and she may not adhere to the relevant norms. Relative to her own information or her own norms, some other act ϕ' may be preferable. Loosely speaking, when we ask what an agent ought to do (in light of such-and-such norms and such-and-such facts), we want to know what she would do if she were adequately responsive to the relevant reasons. Any act the agent could have performed if only she were differently motivated should therefore count as an available option. An act is genuinely unavailable only if no (reasonable) variation of the agent's volitional state – and accordingly, no (reasonable) variation of her beliefs, desires, values, and deliberational processes – would lead her to perform the act.

The parenthetical reasonability condition corresponds to the "possible variations" in Analysis 1a. If someone has severe arachnophobia, we may not want to say that she ought to pick up a spider, even if doing so would have desirable consequences. We certainly don't want to say that a person in a coma ought to get out of bed. In general, when we ask what an agent ought to do, we hold fixed constraints imposed by her physiology and her environment: in a trolley problem, we don't consider acts the agent could perform if she were 100 meters tall or if she had access to a button that stops the trolley; as the arachnophobia case illustrates, we also hold fixed certain neurological conditions that constrain the available volitional states.

All this suggests that when we consider what an agent ought to do, the relevant options are precisely the acts that are under the agent's volitional control. But things are more complicated. Return to Charley, standing in front of the safe. If Charley formed the intention to dial 448-961-5237, he would end up opening the safe. Analysis 1a (correctly) predicts that Charley has the de re ability to open the safe. But suppose Charley could prevent some tragedy by opening the safe. One might have thought that if an agent can perform an act that would prevent a tragedy (at no significant costs), and if no other available act would be comparably good, then the agent is obligated to perform that act. More cautiously, one might have thought that at least if an agent *knows* that they can perform an act that would prevent a tragedy etc., then they are obligated to perform the act. But on the de re reading, Charley *can* prevent the tragedy by opening the safe, and he *knows* that he can. Yet it seems wrong to claim that Charley is obligated to open the safe. We certainly wouldn't blame him if he doesn't open the safe. So we don't seem to count all the acts that are under an agent's volitional control as relevant options when we consider what she ought to do.[1]

1 One might bight the bullet and say that Charley is indeed obligated to open the safe, even though he

We need a more restrictive concept of abilities that excludes acts like Charley's opening the safe. This is, I think, where the concept of de dicto abilities comes into play.

What is required for an agent to have an ability de dicto? Return to (1). Why can't Maisy play Eddie's favourite piece, in the de dicto sense? Intuitively, it's because she doesn't know what that piece is. Similarly for (2) and (4). Cyril can't recite the first digits of π because he doesn't know what these digits are. Charley can't open the safe because he doesn't know the combination. What Maisy, Charley, and Cyril lack is neither skill nor opportunity, but plain old information. While some volitions they could form would in fact lead to the desired act, they don't know what these volitions are.

I therefore suggest the following (preliminary) analysis of de dicto abilities. (Like Analysis 1a, the analysis will be revised later.)

> **Analysis 1b.**
>
> S can ϕ de dicto just in case there is a possible variation V of S's volitional state such that (i) V would cause S to ϕ, and (ii) S knows that V would cause her to ϕ provided that ϕing is under her volitional control.

I will explain the 'provided that' clause in a moment. First, I want to argue that Analysis 1b not only gives the intuitively correct verdict in cases like (1), (2), and (4), but also fits the normative job description of delineating the options that are candidates for what an agent ought to do.

Consider what is required for an agent to be under an obligation to perform a certain act. Arguably, the following two conditions are necessary (though not sufficient). First, performing the act must be under the agent's volitional control: some possible variation of her motivational state would make her perform the act. That's what Analysis 1a gave us. But more is required. In addition, the agent must know what she would have to do in order to perform the act. Charley, for example, isn't obligated to open the safe because he has no idea what he would have to do to open it: he has no idea which of his available volitional states would make him open the safe. That excludes opening the safe from the range of available options.

Exactly why the second condition is required depends on one's views on normativity. But here is a plausible sketch of an explanation. A central role of normative judgements is often glossed as "motivational", "action-guiding", or "advisory". There is a close connection between the assumption that an agent ought to ϕ and the assumption that it would be appropriate to advise the agent to ϕ. But if an agent has no idea what she would have to do in order to ϕ, then it would not be appropriate to tell her to ϕ. If you

isn't blameworthy for flouting the obligation. As I will argue below, this is not only counter-intuitive, it also fails to respect the "action-guiding" or "advisory" role of oughts and obligations. If you'd nonetheless prefer to bite the bullet, Analysis 1b (below) could still be motivated by the connection between 'can' and 'blame' rather than the connection between 'can' and 'ought'.

(as an advisor) know the safe's combination, you might reasonably advise Charley to dial 448-961-5237. Perhaps that is what Charley ought to do in the "objective" sense of 'ought' – it is what he ought to do in light of all the facts. By contrast, advising Charley to open the safe would not be appropriate, since he would have no idea how to follow the advice.

The advisory function of normative judgements also sheds more light on the restriction to "possible variations" of the agent's volitional state. There is little point in advising a coma patient to get out of bed. It might be just as inappropriate to advise a person with severe arachnophobia to pick up a spider. In general, when giving advice, physiological conditions that are insensitive to reason must be accepted as constraints.[2]

Now let me explain the 'provided that' clause in Analysis 1b. Consider the following scenario. Dotty has been led into a room with a single door. There is nobody else in the room, the door is unlocked, and she can freely leave. However, she falsely believes (for good reasons, we may assume) that the door is locked. Can she open the door? Intuitively, yes. That is, (5) does not fit into the pattern of (1), (2), and (4).

(5) Dotty can open the door.

In contrast to the earlier examples, there is no relevant reading on which (5) is false. To be sure, Dotty *believes* that she can't open the door, but her belief is simply false.

Yet Dotty does *not* know that willing to open the door would cause her to open the door. On the contrary, she believes that nothing she could will to do would cause her to open the door. Without the proviso in (ii), Analysis 1b would therefore predict that (5) is false on its de dicto reading; it would predict that (5) patterns with (1), (2), and (4).

The proviso gets around the problem. Dotty does not know that willing to open the door would cause her to open the door, but she plausibly does know that *if opening the door is under her volitional control, then* willing to open the door would cause her to open the door. In other words, if she were informed that she can open the door, she would know how to go about achieving that result.

Notice how the judgement about (5) depends on the assumption that there is only one door. Suppose there are three doors; Dotty believes that all of them are locked, but in fact one is unlocked. For some reason, Dotty can't simply try all the doors to get out. Let's say that if she were to try a locked door, an alarm would ring and the unlocked door would lock as well. With these changes, Dotty no longer knows what she would have to do in order to open a door and leave the room, even on the supposition that it

2 There are difficult normative questions here. What exactly are the constraints? Just how strongly must you be afraid of spiders so that picking up the spider no longer counts as an available option? Can ordinary intentions sometimes function as constraints, as in Goldman's [1978] example of an agent who supposedly ought to accelerate given her ill-advised intention to change lanes? Some have argued that there are different flavours of 'ought', corresponding to different rules for what is held fixed (e.g. [Timmerman and Cohen 2016]). The account developed here is neutral on these debates.

is in her power to do so. Analysis 1b predicts (correctly, I think) that it is no longer determinately and unequivocally true that Dotty can open one of the doors, or that she can leave the room.

How does the proviso in Analysis 1b fit the normative job description of de dicto abilities? Again, the details depend on one's views about normativity, but superficially the condition goes well with the advisory picture. It would not be reasonable or appropriate to advise Charley to open the safe, but it would be perfectly reasonable to advise Dotty (in the original scenario) to open the door.[3]

To sum up, I have argued that our concept of de dicto abilities is tailored to a certain normative use of ability statements – namely, to delineate the options an agent might be obligated to perform. Analysis 1b spells out what is required for an act to be an option in this sense.

Analysis 1b is closely related to the conditional analysis, so a brief comparison may be in order. Suppose an agent has a de dicto ability in the sense of Analysis 1b. That is, some variant V of her volitional state would cause her to ϕ, and she knows that V would have this effect (provided that ϕing is under her volitional control). Under normal circumstances, trying or willing or intending to ϕ would then cause her to ϕ. So the agent also has the ability to ϕ in the sense of the conditional analysis. An exception might be cases where an agent knows that she can perform a given act only by trying to perform a different act. I am not sure whether, in such a case, the agent intuitively has or lacks the relevant de dicto ability, and whether the act should count as an option for normative purposes. As it stands, Analysis 1b says the agent has the ability, while the conditional analysis says she does not.

The two accounts more obviously come apart in the other direction. It can easily happen that an agent would ϕ if she tried (or willed, or intended) to ϕ without having the ability to ϕ in the sense of Analysis 1b. Notoriously, the conditional analysis implies that a person in a coma can get out of bed; Analysis 1b does not. More tellingly, perhaps, the conditional analysis faces problems with accidental success. Suppose if Maisy were to attempt to play Eddie's favourite piece, she would choose the Moonlight Sonata – not because she has any reason to think it is Eddie's favourite piece, but simply because she feels like playing it, and by her lights it is just as unlikely to be the right choice as

3 The advice would only be adequate from a perspective in which it is known that the door is unlocked. Thus in a more "subjective" sense, relative to Dotty's own information, it may still not be true that Dotty ought to open the door. I focus on "objective" oughts because these are most directly related to ability judgements. One might suggest that ϕ is a candidate for a subjective obligation iff the agent *knows* or *believes* that she has the de dicto ability to ϕ. But neither 'knows' nor 'believes' is entirely satisfactory: if we require knowledge, subjective oughts depend on external facts of which the agent may be unaware; if we merely require belief, there would seem to be cases where an agent is subjectively obligated to perform an act she cannot in fact perform. For more on the definition of "subjective options", see e.g. [Weirich 1983], [Pollock 2002], [Hedden 2012], [Schwarz 2017].

any other. Does that unambiguously render (1) true? Arguably not. Arguably, Maisy stills lack the de dicto ability; she is still unable to play Eddie's favourite piece "under that description". Relatedly, suppose Maisy in fact tries to play Eddie's favourite piece and happens to choose the Moonlight Sonata. Assuming that 'A and B' entails 'if A were the case then B would be the case' (a popular, although not entirely uncontested assumption in the logic of counterfactuals), the conditional analysis implies that (1) is unambiguously true. But arguably, it is not.[4]

Some have argued that just as one can try-and-succeed without having the relevant ability, one can also try-and-fail even though one has the ability (see [Austin 1961], [Fara 2008], [Vihvelin 2004: 187], [Mandelkern et al. 2017: 312]). John Austin's classical example is of a golfer who kicks himself for missing a putt, knowing that he had the ability to hole it (under the given circumstances). That kind of case seems possible to me. Analysis 1b gets it right, provided the golfer knows what he would have had to do in order to make the putt.[5]

Note that in these cases, the conditional analysis not only fails to match certain linguistic intuitions. It also fails to satisfy the normative job description: it does not identify the acts an agent might be obligated to perform. Charley, for example, is not obligated to open the safe even if by chance he were to succeed if he tried. By comparison, consider another problem for the conditional analysis: that it does not account for "general abilities" like (3). Most likely, Usain Bolt is not *right now* in a position to run 100 meters in under 10 seconds. Perhaps he has only just woken up, or has just eaten, or is in an airplane, is not wearing his running gear, etc. So he would fail if he now tried to run 100 meters in under 10 seconds. Yet there is a good sense in which he *is* able to run 100 meters in under 10 seconds. (At any rate, this was the case until recently.) Neither Analysis 1b nor the conditional analysis captures this sense. But that is not a serious problem if our target is the normatively relevant sense of ability statements. If it is not in Bolt's power to run 100 meters in under 10 seconds now, he can hardly be obligated to perform that act.[6] That the conditional analysis won't do as a fully general analysis

4 On the other hand, it sounds odd to say that an agent is unable to ϕ and yet ϕs; I will return to this observation in section 4. The judgement that successfully ϕing does not entail an ability to ϕ – at least in one normatively important sense of 'ability' – is widespread in the literature; see e.g. [Kenny 1976: 214], [Fara 2008], [Vihvelin 2013: 182], [Southwood and Wiens 2016], [Maier 2018b].

5 Must the golfer be able to identify the volitional state that would have made him succeed? See page 17.

6 Similarly, alleged counterexamples from "finks" and "reverse finks" (see e.g. [Vihvelin 2004]) arguably don't challenge the normative application of the conditional analysis. Imagine a sorcerer has temporarily paralysed Ella's right arm and now monitors Ella's intentions: if he should find that Ella wants to raise her left arm, he immediately paralyses that arm, so that the arm does not rise; if he finds that Ella wants to raise her right arm, he removes the paralysis from that arm so that the arm rises. In fact, Ella does not want to raise either arm, so the sorcerer does not intervene any further. (Ella's left arm constitutes a Frankfurt case (compare [Frankfurt 1969]), her right arm a Lehrer case

of ability statements will become abundantly clear in the next section (see examples 6, 7, 8, 9, 10, 11, 12).[7]

3 The canon

I now want to briefly set aside the de re/de dicto ambiguity, and review the classical quantificational semantics of ability statements as developed by Kratzer and others.[8]

On the quantificational approach, ability statements are regarded as a species in a large family of modal constructions all of which are analysed in terms of quantification over a restricted domain of ("accessible") possible worlds. An agent *can* ϕ iff she ϕs at some accessible worlds; an agent *must* ϕ iff she ϕs at all accessible worlds; and so on.

Before we dive into more details, it is important to clarify what this approach is trying to achieve. Take Maisy's ability to play the Moonlight Sonata. What makes it the case that she has this ability? Intuitively, the answer should appeal to certain physiological features of Maisy, relating to finger dexterity, hand coordination, muscle memory, etc., acquired through years of piano practice. The quantification account instead seems to suggest that Maisy's ability is grounded in facts about other possible worlds: Maisy has the ability to play the Moonlight Sonata, here in our world, because there is some other world where she – or worse: someone else, see [Lewis 1986: ch.4] – plays the Moonlight Sonata. That sounds wrong. (See e.g. [Vetter 2013] for this kind of complaint.)

Similar complaints have been raised against other applications of possible-worlds semantics, such as the Lewis-Stalnaker account of counterfactuals, or the analysis of physical necessity in terms of nomically possible worlds. (An especially well-known example is [Kripke 1980: 44f.].) But these complaints misunderstand the aim of possible-worlds semantics. The analysis in terms of possible worlds is not meant to identify the metaphysical grounds of the relevant phenomena, nor is it meant to offer a substantive reduction of the phenomena to psychologically more basic concepts. It is merely supposed

(compare [Lehrer 1968]).) The conditional analysis implies that Ella can raise her right arm while she cannot raise her left arm. The alleged problem is that it should be the other way round: Ella *cannot* raise her (paralysed) right arm while she *can* raise her (non-paralysed) left arm. But these judgements are at best controversial, and they don't fit the normative function of ability statements.

7 It is tempting to say that an agent can get out of bed only if she *can intend to get out of bed*; hence the "regress argument" against the conditional analysis. But people generally don't form intentions by intending to form these intentions. The model of agency outlined above (that motivated Analyses 1a and 1b) therefore doesn't apply to "acts" like intending to get out of bed. In terms of Analysis 1b, the coma patient can't intend to get out of bed simply because the intention isn't a possible variation of her volitional state: it is incompatible with the general physiological conditions that we hold fixed when we consider what an agent ought to do. How the word 'can' comes to have this different meaning when we talk about what an agent can intend (rather than what acts she can perform) will also become clear in the following sections.

8 I will be superficial on linguistic details. For more in-depth surveys, see e.g. [Portner 2009], [Hacquard 2011], [Kratzer 2012].

to offer a useful framework for modelling the phenomena themselves. As David Lewis explains, in response to the objection that counterfactuals should be made true by facts about our world, not by facts about other worlds:

> [It] is indeed the character of our world that makes the counterfactual true. But it is only by bringing the other worlds into the story that we can say in any concise way what character it takes to make what counterfactuals true. The other worlds provide a frame of reference whereby we can characterise our world. [Lewis 1986: 22]

In the case of 'can' and 'must', all the substance lies in the accessibility relation. In easy cases, accessibility is a matter of preserving certain features of the agent and her environment. (I'll get to more complicated cases in a moment.) When we talk about whether someone can or cannot play the Moonlight Sonata, we hold fixed whether she has the relevant finger dexterity, muscle memory, and whatever else is required to play the Sonata. Maisy can play the Moonlight Sonata because she has these features. It is her possession of these features, here in our world, that explains why she plays the Moonlight Sonata in some accessible world, and therefore why she *can* play the Moonlight Sonata in our world. As in the case of counterfactuals, the other worlds merely provide a convenient frame of reference for characterising the actual world.

What is convenient about the possible-worlds account only becomes apparent if we look at more examples, and a wider range of contexts. Consider (6), adapted from [Kratzer 1981].

(6) Mr. Percy can play the trombone.

We might utter (6) to convey that Mr. Percy's physiology is compatible with playing the trombone, after sufficient training and practice. (This is the sense in which rhinoceroses can't play the trombone, while humans can.) Alternatively, we might want to convey that Mr. Percy has whatever features ordinary people acquire when they learn to play the trombone. In that sense, (6) can be true even if due to an acute asthma episode, Mr. Percy is presently incapable of playing the trombone. But (6) can also be used to rule out this kind of state, conveying that Mr. Percy is presently capable of playing the trombone. In other contexts, (6) might be used to convey that Mr. Percy has not only the capacity, but also the present opportunity to play the trombone: his trombone is right in front of him; he only needs to pick it up. Alternatively, one could use (6) to convey that although Mr. Percy does not at present have the opportunity to play the trombone, he can bring about such an opportunity with reasonable effort. (Returning from his year-long expedition to Antarctica, to which he did not bring his trombone, Mr. Percy says, in the airport, that he is glad that he can now play the trombone again.)

Each of these things we could express by (6) corresponds to a different ground or truth-maker for (6). Metaphysically, these grounds have almost nothing in common. If we wanted to give a semantics of (6) in terms of its grounds or truth-makers, the semantics would be a long and gerrymandered list. By refraining to give a substantive analysis, the classical possible-worlds account is able to bring out a common theme. What's common to all the cases just reviewed is that there are certain facts about Mr. Percy and his environment that are incompatible with playing the trombone, so that he is not playing the trombone in any world where those facts are held fixed.

The variety of truth-makers is even more evident if we look at other sentences. (7)–(9), for example, illustrate that abilities don't require intentional agency. In (10), there isn't even a recognizable subject.[9]

(7) I can hear you.

(8) New phones can instantly switch between available networks.

(9) Steel can withstand temperatures of 1000 ℃.

(10) It can take years to earn someone's trust.

Sometimes, the grounds for an ability or lack of ability lie in relevant norms or desires, as in (11) and (12):

(11) I can't come to the party; I need some rest.

(12) I can't come to the party; I have to look after my children.

Notice that it is not the *existence* of the need or obligation in (11) and (12) that is incompatible with coming to the party, but their *fulfilment*. In Kratzer's account, this is handled by assuming that the accessible worlds over which modals quantify are restricted by a combination of a "modal base", which holds fixed certain facts about the actual world, and an "ordering source" which ranks the worlds in the modal base.

Formally, and glossing over some details that are irrelevant for present purposes, Kratzer's truth-conditions for 'S can ϕ' may be rendered as follows.

$$[[S \text{ can } \phi]]^{w,R,\leq} = 1 \Leftrightarrow \exists w'(wRw' \wedge \forall w''(wRw'' \rightarrow w'' \leq w') \wedge [[S\ \phi\text{s}]]^{w'} = 1).$$

Here, w is the world of evaluation, R is a contextually supplied relation that determines what facts about w are to be held fixed, and $\leq$ is a contextually supplied partial order

9 (10) is also a rare example of a 'can' statement that is not equivalent to an 'able' statement. Arguably, this is not due to a deep semantic difference between 'can' and 'able', but simply to the syntactic fact that 'able' is a control predicate that can only occur right above the verb phrase, while 'can' can also appear above tense and aspect; see [Brennan 1993], [Hacquard 2010].

(which is often empty for ability modals), ranking w' above w'' just in case w'' better satisfies relevant norms, desires, or other kinds of standards. 'S can ϕ' is true at w iff S ϕs in some of the highest-ranked worlds R-related to w.

The lexical entry for 'can' that generates the above truth-conditions is the same for all "possibility" modals, of all flavours: deontic, bouletic, circumstantial, epistemic, etc. So we don't have to assume that 'can' is massively ambiguous, in English and a wide range of other languages. "Necessity" modals get a corresponding treatment with universal instead of existential quantification.

The Kratzerian approach has some other attractive features. For example, it explains why 'I must sneeze' is equivalent to 'It is not the case that I can not sneeze'. It explains the interaction of 'can' and 'if' in sentences like 'I can dance if nobody is watching', where, on the most natural interpretation, the if-clause has neither wide scope nor narrow scope, but rather seems to restrict the set of worlds over which the modal quantifies (see [Kratzer 1986], [Horgan 1979: 350]). It also explains why 'ought' implies (a specific kind of) 'can': if S ϕs in the normatively best of the contextually accessible worlds – as 'S ought to ϕ' claims – it follows that S ϕs in some contextually accessible worlds – as 'S can ϕ' claims.

This sense of 'can', implied by 'ought', is the 'can' we have studied in section 2. How do Analyses 1a and 1b fit into the Kratzerian picture? According to Analysis 1a, an agent has the de re ability to ϕ iff there is a possible variation of her volitional state that would make her ϕ. Technical details aside, this easily fits the Kratzerian picture: the corresponding accessibility relation is one on which a world counts as accessible iff it would come about through some possible variation of the agent's volitional state.[10]

Evidently, these accessibility standards won't give the right verdicts in cases like (8), (9), or (10), and they cover only one of the many readings of (6). A fully general analysis of (de re) abilities will have to allow for other accessibility relations, and it will have to deal with the way accessibility can be affected by norms or desires. I suggest that we can do no better than the classical quantificational semantics:

Analysis 2a.

S can ϕ de re iff there are accessible worlds where S ϕs.

As explained above, 'accessible' here serves as an umbrella term, covering a variety of aspects that are held fixed for different ability statements in different contexts. A more informative analysis would break down these differences and explain what kinds of aspects are held fixed for various kinds of ability statements. Philosophers sometimes distinguish two readings of ability statements, one "general" and one "specific" (see e.g. [Austin 1961: 230], [Honoré 1964], [Whittle 2010], [Mandelkern et al. 2017], [Maier

10 The main technical detail I'm setting aside is how to handle possible cases in which there is no single world that would determinately result from a certain volitional state.

2018a]). Roughly speaking, general ability is supposed to be skill, while specific ability is skill plus opportunity. But this picture is too simple: we have found no fewer than five readings for (6), and neither skill nor opportunity seems relevant to (9), (11), or (12). I will not attempt to give a more informative analysis. In any case, if we want to capture what is common to all kinds of abilities, I suspect we have to fall back on abstract principles like Analysis 2a that reveal little about the relevant metaphysical grounds.[11]

So I think the classical quantificational approach is right for the de re sense of ability statements. But it doesn't account for the de dicto sense. I already mentioned the basic problem in section 1. Here it is again.

On their de dicto reading, (13a) is true while (13b) is false.

(13) a. Charley can dial 448-961-5237.

b. Charley can open the safe.

The quantificational model easily accounts for the truth of (13a): since Charley has the skill and opportunity (although possibly not the motivation) to dial 448-961-5237, there are accessible worlds where he dials that combination. But then how could there fail to be accessible worlds where he opens the safe, so that (13b) comes out false? We would have to say that when we evaluate (13a), worlds where Charley dials 448-961-5237 are accessible, but when we evaluate (13b), these same worlds become inaccessible. But why? What other worlds become inaccessible when we evaluate (13b)? Does Charley not try any combination at all in the accessible worlds? Why not? Why do the worlds that render (13a) true suddenly become accessible again for (13b) if Charley gains information about the right combination?

Some attempts have been made to answer these questions. Let me briefly mention one. In [Mandelkern et al. 2017], Matthew Mandelkern et al. (in effect) suggest that the accessible worlds with respect to a given act ϕ are worlds where the agent performs some act ψ that is "practically available with respect to ϕ"; in normal contexts, ψ counts as practically available with respect to ϕ just in case the agent "could reasonably conclude in favor of doing ψ with respect to the goal of doing [...] ϕ" [Mandelkern et al. 2017: 319].

11 John Maier [2015, 2018a] suggests that what unifies abilities of all kinds is that they can be analysed in terms of an agent's "options". Unfortunately, I don't understand what he means by 'options', which he takes as primitive. Maier [2018a: 424] states that whenever an agent ϕs, even if by sheer luck, then ϕing was an option for the agent; his concept of options therefore does not fit the normative job description from section 2. (As Maier [2018a: 423f.:fn.10] acknowledges, it also does not fit the job description for "subjective options" that I mentioned in footnote 3 above.) If we really want to explain all ability statements in terms of options, we presumably have to assume that I have the "option" of hearing a loud noise outside my office, and that tardigrades have the "option" of surviving high temperatures. Much of the diversity in ability statements, it seems to me, is here hidden in the unanalysed notion of an option.

The idea is that Charley could not reasonably conclude in favour of dialling 448-961-5237 with respect to the goal of opening the safe, since this would require knowledge that the former act is a way to bring about the latter (see [Mandelkern et al. 2017: 320]).[12] We also have to assume that Charley could not reasonably conclude in favour of opening the safe with respect to the goal of opening the safe (I'm not entirely sure why), while he *could* reasonably conclude in favour of dialling 448-961-5237 with respect to the goal of dialling 448-961-5237. We then get the right verdicts for (13a) and (13b). Unfortunately, we seem to get the wrong verdicts for other cases, like (5) and (11).[13] It is also unclear how considerations of what's reasonable could give rise to the many different readings of (6), or how the proposed account could deal with cases like (7)–(9) that don't involve intentional agency.

I don't claim to have established that one can't account for de dicto abilities by tinkering with the accessibility relation. Nonetheless, I want to put forward a different proposal. I want to suggest that the de dicto reading arises not through a restriction on the accessible worlds, but through a restriction on the interpretation of the embedded verb phrase ϕ (the "prejacent").

4 Restricting the prejacent

When we talk about whether someone can or cannot ϕ, we sometimes have in mind a particular way of ϕing. Consider a well-known puzzle from action theory. Tallulah's left arm is paralysed. Intuitively, that renders (14) true.

(14) Tallulah can't move her left arm.

But suppose Tallulah's right arm is not paralysed. As a consequence, she *can* move her left arm, in the way in which she can move a cup on the table, by grabbing it with her right arm. That's not the kind of "moving" we have in mind when we assent to (14). What we mean by (14) is that Tallulah can't move her left arm "actively", using the muscles in the arm.

(14) is true on one reading, and false on another. The two readings are orthogonal to the de dicto/de re ambiguity. But they motivate the idea that ability statements might be sensitive to context-dependent restrictions on what counts as a relevant way of ϕing.

12 Mandelkern et al. note that (13b) (or rather, their analogous (38)) also has a true reading, which they call "objective". Here, they suggest, we consider not what the agent knows, but what is in fact the case. Note the similarity to Analyses 1a and 1b.

13 Mandelkern et al. consider a case like (11), suggesting that whenever an agent has decided to do one thing, "it would not be reasonable for him to conclude in favor of doing something else", because "[p]eople should stick with their decisions." [Mandelkern et al. 2017: 322f.]. But we normally don't take the fact that an agent has decided to perform a certain act to imply that they are not able to do anything else.

If we wanted to account for the two readings of (14) in terms of accessibility, we would have to say that in contexts where we judge (14) to be true, worlds where Tallulah moves her left arm with her right arm are inaccessible. But that seems mysterious. After all, 'Tallulah can move her left arm with her right arm' is determinately true. Why would all the worlds where she moves her left arm with her right arm become inaccessible when we evaluate (14)?[14]

Here's another phenomenon that points towards contextual restrictions on the interpretation of the prejacent: context often determines *how well* an agent must be able to ϕ in order to count as having the ability to ϕ. When we say that someone can play the piano, for example, we usually don't just mean that they are able to produce sounds by pushing down piano keys. We mean that they can play the piano sufficiently well, but what that means depends on context.

Simplifying a little, let's imagine that every instance of playing the piano can be assigned a scalar *performance level* taking into account both the difficulty of the piece and the quality of the play.[15] We might then explain the context-dependence of sentences like (15) by assuming that context determines a minimal performance level that constrains what kinds of acts in accessible worlds count as falling under 'playing the piano':

(15) Maisy can play the piano.

(15) is true if there are accessible worlds where Maisy plays the piano at a level that exceeds the threshold.

The present suggestion might explain another puzzle about ability modals that has proved difficult to tackle in Kratzer's framework: the apparent gradability of abilities. We say that S can *barely* play the piano, or that S can play *better* than S'. If the interpretation of 'can ϕ' is relative to a scalar performance level, there is a straightforward explanation of these phenomena, familiar from gradable adjectives: S can barely play the piano if she can only play the piano at a low level; S can play better than S' if there is some level L such that S can play to a degree greater than L while S' cannot.

So there are reasons to think that context can impose restrictions not just on the accessible worlds, but also on what counts as a relevant way of ϕing. I suggest that this is also how the de re/de dicto ambiguity comes about: a de dicto ability to ϕ is an ability to ϕ in a particular way.

14 One might suggest that a "stereotypical ordering source" could do the job: worlds where Tallulah moves her left arm with her right arm are set aside because they are unusual. But that would be an unusual kind of ordering, ranking not the circumstances under which an act may or may not be performed, but different ways of performing the act. In addition, arguably (14) has a true reading even if a different ordering source is contextually salient.

15 The simplification is that these two aspects, and various sub-aspects in which they could be decomposed, can be measured on a single numerical scale.

What is that way? According to my proposal from section 2, Maisy lacks the de dicto ability to play Eddie's favourite piece because she doesn't know which of her volitional states, if any, would result in her playing Eddie's favourite piece. Consider what would change if Maisy acquired the missing information. When we look at the accessible worlds – the worlds that would result through some variation of Maisy's volitional state – we would then find not just worlds where she plays Eddie's favourite piece by chance, but worlds where she *knowingly* plays Eddie's favourite piece.

This suggests that we might analyse a de dicto ability to ϕ as an ability to *knowingly* ϕ, where an agent *knowingly* ϕs just in case she ϕs while knowing that she ϕs. But this isn't quite right. For suppose that, in the original Maisy scenario, if Maisy were to start playing the Moonlight Sonata, Eddie would immediately signal to her that she is playing his favourite piece. Then there would be accessible worlds where Maisy knowingly plays Eddie's favourite piece. Yet Maisy would still lack the de dicto ability to play Eddie's favourite piece.

It seems to matter *how* Maisy comes to know that she is playing Eddie's favourite piece in the accessible worlds. If Maisy had the relevant de dicto ability, she would know that she is playing Eddie's favourite piece without drawing on Eddie's reaction. In Anscombe's memorable words, should would know that she is playing Eddie's favourite piece "without observation" ([Anscombe 1957: sec.8]). That is, her decision alone would make it rational for her to believe that this is what she is doing.

I therefore suggest the following analysis.

> **Analysis 2b.**
>
> S can ϕ de dicto iff there are accessible worlds where S ϕs *transparently*.
>
> S ϕs *transparently* iff S ϕs as a consequence of a volitional state that warrants believing that she will ϕ given that ϕing is under her volitional control.

The 'given that' clause mirrors the 'provided that' clause in Analysis 1b. Remember (5): Dotty can open the door de dicto even though no decision she could make would, by itself, warrant believing that she will open the door. Nonetheless, a decision to (try to) open the door would warrant believing that *if she can open the door*, then she will open the door. By Analysis 2b, this is enough to render (5) true.

Analysis 2b is not fully equivalent to Analysis 1b, even if the standard of accessibility is given by possible variations of the agent's volitional state. For one thing, Analysis 1b seems to require that the agent can somehow pick out in advance the volitional state(s) that would make her ϕ; Analysis 2b does not. Moreover, while Analysis 1b requires knowledge, Analysis 2b only requires having a warrant for (true) belief. The two analyses therefore give different verdicts about Gettier cases (among others): if you're looking at a stopped clock that happens to show the correct time, and you have reason to believe the

clock works, Analysis 2b says that you have the de dicto ability to tell the time, while Analysis 1b says that you don't. In these respects, it seems to me that Analysis 2b is superior to Analysis 1b, but I won't argue the point.

In fact, I am not particularly wedded to the precise definition of transparency in Analysis 2b. What I want to suggest is that de dicto abilities are abilities to perform the relevant act *in a particular way* – a way that is important for normative contexts. Analysis 2b is an attempt to spell out that way, but it should be understood only as an illustration.

I have used the artificial adverb 'transparently' for the sake of neutrality, but it may well turn out that the relevant restriction matches some more familiar concept. In particular, it may be worth exploring whether de dicto abilities can be analysed as abilities to *intentionally* perform the relevant act. According to one influential tradition in action theory, intentional action is distinguished precisely by the agent having non-observational knowledge of what she is doing (see e.g. [Anscombe 1957], [Velleman 1989], [Setiya 2008]). My proposal lends some support to this idea, insofar as it explains why we might have this concept of intentionality. But it is not important to my proposal whether the idea is right.

The account I have outlined predicts that the de re/de dicto ambiguity only arises for ability attributions in which the subject is an agent capable of deliberation and belief, and in which the relevant act could result through a deliberate choice. In other cases, the de dicto reading is trivially false, so the de re reading should dominate. That seems correct: it is hard to sense the ambiguity in statements like (7), (8), or (9).

So far, I have assumed that somehow or other, context can effect a restriction on what counts as an instance of ϕing among the accessible worlds. How does context achieve this? How exactly should we modify the Kratzerian canon?

Technically, we could include the restriction in the accessibility relation. For note that there are R-accessible worlds at which an agent ϕs ψly (e.g., *transparently*, *actively*, or *sufficiently well*) just in case there are R^*-accessible worlds at which the agent ϕs, where R^* is defined (say) as follows:

$$wR^*w' \Leftrightarrow wRw' \text{and the agent } \phi\text{s } \psi\text{ly at } w'.$$

So we get the de dicto reading of ability statements by using R^* as the accessibility relation. The downside is that we would have to give up the idea that accessibility is a matter of holding fixed salient facts about the agent and the environment (plus the fulfilment of relevant needs or norms). Indeed, the present account arguably doesn't answer, but only relocates our question: by what mechanism does context determine R^* as the relevant accessibility relation?[16]

16 With R^* as the accessibility relation, 'S can ϕ' is true at w iff there is some accessible world from w –

Another way to account for the de dicto reading in a broadly Kratzerian framework is to assume that the relevant verb phrases are simply ambiguous: 'open the safe' can mean either *open the safe transparently* or *open the safe either transparently or intransparantly*. On the Kratzerian model, we then get the de dicto reading of (4) ('Charley can open the safe') if, in the relevant context, 'open the safe' means *open the safe transparently*. To handle (14) in a parallel fashion, we'd have to assume that 'move an arm' can mean either *move an arm actively* or *move an arm either actively or passively*; for (15), we'd have to assume that 'play the piano' means *play the piano sufficiently well*, where the performance threshold is determined by context.

An attractive feature of this idea is that it explains the oddity of statements like (16).

(16) Charley wasn't able to open the safe, but he did open the safe.

According to Analysis 2b, lack of a de dicto ability is compatible with accidentally performing the relevant act. There should therefore be no contradiction in assuming that Charley wasn't able to open the safe (de dicto) even though he did open the safe, by luck. But many people hear a strong tension in (16). On the present proposal, this could be explained by assuming that the ambiguity in 'open the safe' is resolved the same way in both parts of (16), so that (16) is ambiguous between (17a) and (17b), both of which are contradictory (in the absence of an ordering source).[17]

(17) a. Charley wasn't able to open the safe transparently, but he did open the safe transparently.

b. Charley wasn't able to open the safe either transparently or intransparantly, but he did open the safe either transparently or intransparantly.

On the other hand, one would like to see independent evidence for the supposed lexical ambiguities. Indeed, the present proposal implies that (18) has a false reading if Charley opened the safe by luck, not knowing the combination.

(18) Charley opened the safe.

As far as I can tell, such a reading does not exist. Even if Charley only opened the safe by luck, (18) is unambiguously true.[18]

full stop. The definition of R^* ensures that the agent ϕs at every accessible world. Could we define R^* by merely requiring that at R^*-accessible worlds the agent ψly does *something or other*, without fixing that the agent ϕs? No: we want (13b) to be false even though Charley transparently dials 448-961-5237 at accessible worlds; similarly, we want (14) to be false even though Tallulah actively moves her right arm at accessible worlds.

17 Thanks to Brian Rabern for drawing my attention to this point.

18 What about the oddity of (16)? [Hacquard 2009], drawing on [Bhatt 2006], argues that on an independently motivated account of how modals interact with aspect, a perfective aspect effectively restricts the domain of accessible worlds to the singleton of the actual world, rendering 'can ϕ' equivalent to 'ϕ', and 'can't ϕ' equivalent to 'not ϕ'.

Here is another idea. On Kratzer's account, ability modals have two contextual parameters, one selecting a modal base, the other an ordering source; together, these determine the sphere of accessible worlds. Perhaps we need an additional parameter that functions somewhat like the adverbial modifiers in (19a)–(19c).

(19) a. I can *easily* get a new certificate.

b. I can fire him *at will*.

c. I can fall asleep *on command*.

Grammatically, these arguably modify 'can', not the embedded verb phrase. And they can be left tacit. I have exploited the possibility of adverbial modification all along, speaking of what an agent can do 'de dicto', 'de re', 'transparently', or 'intentionally'. Now suppose English already has a 'de dicto' modifier, which may or may not be lexicalised. The de dicto reading of ability statements might then arise from a tacit presence of that modifier. To handle (14) and (15), we could similarly postulate an 'actively' modifier and a 'sufficiently well' modifier.

Semantically, these modifiers are simply passed on to the embedded verb phrase: you can^{actively} raise your arm iff you can perform the act of actively raising your arm; you can$^{sufficiently\text{-}well}$ play the piano iff you can perform the act of playing the piano sufficiently well; you can^{transparently} open the safe iff you can perform the act of transparently opening the safe. The full semantics of 'can' might therefore look roughly as follows, with a new parameter ψ of adverbial type:[19]

$$[[\text{can}]]^{w,R,\leq,\psi} = \lambda P_{<s,<e,t>>}\lambda x_e.\exists w'(wRw' \wedge \forall w''(wRw'' \rightarrow w'' \leq w') \wedge \psi(P)(w)(x) = 1).$$

I have to leave a more thorough investigation of this proposal to another occasion. My aims in the present section have been relatively modest. I wanted to make it plausible that (a) the de re/de dicto ambiguity of ability statements can be explained by a contextual restriction on the interpretation of the embedded verb phrase, (b) postulating such restrictions is not entirely ad hoc, as the same kind of mechanism arguably explains the context-sensitivity of (14) and (15), and (c) the existence of such a mechanism does not call for a complete revision to the orthodox Kratzerian account of modals.

In section 1, I mentioned that one might try to explain the ambiguity in (1) and (2) as a scope ambiguity. I argued that such an explanation would not be general enough. Nonetheless, we can now explain why scope seems to play a role – for example, why (1) is intuitively true if 'Eddie's favourite piece' is given wide scope, as forced in (20).

(20) Eddie's favourite piece is something Maisy can play.

19 For the sake of perspicuity, I here assume that 'can' is of type $<< s, et >, et >$, taking a property as argument. Linguistic orthodoxy instead assumes that 'can' applies to propositions; see [von Fintel and Heim 2002: ch.5].

According to Analysis 2b, (20) is true on its de dicto reading iff (informally speaking) there is some x such that (a) x is Eddie's favourite piece, and (b) Maisy has the ability to play x transparently. Is there such an x? Yes: the Moonlight Sonata. Giving 'Eddie's favourite piece' wide scope here makes the de dicto reading effectively equivalent to the de re reading, on which (1) is true.

5 Defending the canon

I have suggested that the orthodox quantificational approach correctly captures the de re sense of ability statements: an agent can ϕ de re iff there are accessible worlds where she ϕs. To get the de dicto reading, I have argued that we should appeal to a contextual restriction on the prejacent ϕ: an agent can ϕ de dicto iff there are accessible worlds where she ϕs transparently. To conclude, I want to revisit some objections to the quantificational account. As we will see, all of them involve de dicto abilities and are easily answered by the account I have outlined.

The first objection I want to revisit might be called the *no-fluke argument* (it is related to an objection to the conditional analysis reviewed in section 2). On the quantificational account, an agent has the ability to ϕ as long as she ϕs in some accessible world. Often, however, an agent intuitively lacks the ability to ϕ even if her skills and the relevant environmental circumstances are compatible with her ϕing by chance. To use the standard example from [Kenny 1976], suppose Betsy is a novice dart player who can barely hit the board. If we look at accessible worlds where Betsy attempts to hit the bullseye, we find the dart landing all over the place. In a few of these worlds (loosely speaking), the dart happens to hit the bullseye. The quantificational account therefore seems to imply that (21) is true. But intuitively, it is not.

(21) Betsy can hit the bullseye (on the next throw).

Relatedly, as I mentioned in section 2, it has often been suggested that even actual performance of an act does not establish an ability to perform the act. The general point is that abilities like the one attributed in (21) seem to requires a kind of robustness: it shouldn't be a mere fluke that the agent succeeds if she tries. Several alternatives to the quantificational account have been developed around this intuition; see e.g. [Nowell Smith 1960], [Brown 1988], [Greco 2007], [Vihvelin 2013], [Jaster 2016], [Maier 2018b].

In response, we need to first distinguish the de re sense of (21) from the de dicto sense. Interpreted de re, (21) is arguably true: it is in Betsy's power to do something that amounts to hitting the bullseye. This sense is usually trivial and therefore not very salient, but it can be made salient by supplying more context. For example, imagine Betsy has been in an accident and is slowly regaining the ability to move her arm. We discuss how far she could throw a dart in her present state. More precisely, we discuss

which points on the wall she could reach if she were to throw a dart from where she is standing. Suppose any point above a height of 3 meters on the wall is ruled out by the condition of Betsy's arm, but points below 2 meters within a certain radius are not. If the bullseye falls in that region, we could truly utter (21).

In normal contexts, however, the de dicto reading of (21) dominates. Since the orthodox quantificational account only captures the de re reading of ability statements, we should not be surprised that it has trouble with (21). The tweak I suggested in the previous section gets around the problem. According to Analysis 2b, Betsy has the de dicto ability to hit the bullseye iff there are accessible worlds where she hits the bullseye transparently. These would be worlds where Betsy has reason to believe that she will hit the bullseye, merely based on a relevant decision. By normal accessibility standards, there won't be any such worlds. So (21) comes out false.

In effect, the transparency condition in Analysis 2b ensures that the success events aren't mere flukes. But note that the no-fluke requirement is epistemic: by Analysis 2b, if an agent has a de dicto ability, then *from her own epistemic perspective* it is not an accident that she succeeds. We do not require that the agent succeeds across a whole range of worlds, with varying environmental circumstances. The difference shows up in unusual cases like the following. Suppose Betsy has been informed by a time-traveller that if she will throw the next dart, then she will hit the bullseye. Arguably, that renders (21) true, even on its de dicto reading. A de dicto ability to hit the bullseye therefore does not always require general skill or success across a range of circumstances.

A second, though related, objection to the quantificational account is what I'll call the *argument from general abilities*. (21) attributes a "specific" ability, talking about what Betsy can do in her present circumstances. Now we focus on "general" ability statements in which the present circumstances are not held fixed. The objection is that whatever we say about specific abilities, general abilities surely require more than successful performance at a single accessible world.

For example, consider (2) ('Cyril can recite the first 10 digits of π'). When we say that someone can recite the first 10 digits of π, we usually attribute a "general ability". We don't mean that they can do it right now, in whatever circumstances they currently find themselves. On the other hand, we seem to require more than the existence of *some* possible circumstances under which the agent would recite the digits, perhaps by chance. Intuitively, if an agent can recite the first 10 digits of π, then she can do so systematically whenever the need arises – in good weather and in bad weather, in the morning and in the evening.

In response, we again need to distinguish the de re reading and the de dicto reading of the relevant statements. The objection is most plausible for the de dicto reading, which is also more salient for sentences like (2). On the account I have outlined, an agent has the de dicto ability to recite the first 10 digits of π just in case there is at least one accessible

world where she transparently recites the digits. Suppose there is such a world. In that world, the agent knows that she is reciting the first 10 digits of π merely on the basis of her decision to utter a certain sequence of sounds. That is, she knows that uttering those sounds would amount to reciting the first 10 digits of π. How could she know this? Most likely because she knows what the first 10 digits of π are. So if (2) is true on its de dicto reading, there should be accessible worlds at which Cyril knows the first 10 digits of π.[20] But whether or not an agent knows the first 10 digits of π is plausibly one of the aspects of the actual world that we normally hold fixed when we evaluate statements like (2). That is, if Cyril doesn't actually know the first 10 digits of π, then he also doesn't know them in any accessible world.[21] The hypothesis that an agent transparently recites the first digits of π in some accessible world therefore implies that she knows what these digits are in *all* accessible worlds. This, in turn, implies that across the accessible worlds, the agent almost always succeed to recite the first 10 digits of π if she tries, irrespective of the weather or the time of the day.

The upshot is that the account I have outlined can vindicate the intuition that "general abilities" like the one attributed in (2) imply successful performance across a wide range of trying circumstances, without explicitly putting this into the semantics.

By contrast, consider a variant of (3):

(22) Usain Bolt can run 100 meters in under 10 seconds.

Intuitively, (22) is true (or was true until recently). As mentioned earlier, this means that it doesn't attribute a "specific" ability: the claim is not that in his present circumstances, Bolt can run 100 meters in under 10 seconds. So (3) attributes a "general ability". But unlike (2), it does not imply successful performance across a wide range of circumstances. On the contrary, it is enough if there are a few, fairly special conditions under which Bolt can run 100 meters in under 10 seconds. The popular idea that "general abilities" require success across a variety of circumstances thus gets (3) wrong. The account I have outlined gets it right: holding fixed Bolt's strength, but possibly varying his environment, his state of alertness, etc., we find *some* worlds at which he runs 100 meters in under 10 seconds – including some in which he knows from the outset that he is going to achieve this goal. That is enough to render (3) true. Unlike in (2), here successful performance in some accessible worlds does not imply successful performance in a wide range of accessible worlds.

I have given one explanation for the apparent robustness requirement in "general" ability statements like (2). In other cases, different mechanisms might play a role. For

20 This ignores the proviso in Analysis 2b. In fact, according to Analysis 2b, it would suffice for the truth of (2) if (say) Cyril gives credence 0.4 to the digits being $3, 1, 4, 1, 5, 9, 2, 6, 5, 3$ and much lower credence to every other possibility. That seems correct to me.

21 As usual, the standards depend on context: 'Can you recite the first 10 digits of π for this video I'm making?' – 'Sure, if you tell me what they are.'

example, a stereotypical ordering source might sometimes restrict the accessible worlds to relatively normal ones (compare [Kratzer 1981: 60f.]). The claim that steel can withstand temperatures of 1000 ℃, for instance, might be understood to mean that steel can withstand such temperatures *under otherwise normal conditions*, even if the actual utterance context is non-normal.

Another mechanism is at work in statements like 'I can draw animals', which seems to imply that I can draw a whole range of animals. This "free choice effect" is plausibly a scalar implicature (see [Klinedinst 2007], [Schwarz 2018]).[22]

Sometimes, more elementary pragmatic mechanisms might be invoked. For example, suppose you can only play the piano in a dark and quiet room at temperatures below freezing, and you know that I am looking for a pianist to entertain a group of children. It would be misleading to say, without further qualification, that you can play the piano, even if what you say is literally true.

In sum, the tweaked version of the quantificational semantics that I have defended plausibly has enough resources to answer the problem of general abilities.

Let me turn to a third objection, due to [Kenny 1976]: the *argument from the failure of K*. Imagine a well-shuffled deck of cards, lying face-down on the table. Can you draw a card that's either red or black? Sure: any card will do. But can you specifically draw a red card, in a single try? Arguably not. Nor can you specifically draw a black card. So, (23a) seems to be true while (23b) and (23c) are false.

(23) a. You can draw a red or black card.
b. You can draw a red card.
c. You can draw a black card.

The problem is that if 'can' is an existential quantifier over accessible worlds, then 'can(ϕ or ψ)' should entail 'can(ϕ) or can(ψ)' – a (consequence of a) principle known as 'K' in modal logic: if there are no accessible worlds where you draw a red card, and none where you draw a black card, how could there be accessible worlds where you draw a card that's either red or black?

You can probably anticipate my diagnosis. (23a)–(23c) are ambiguous between a de dicto reading and a de re reading. Since the de re reading of (23b) and (23c) is

22 Paradigmatic "free choice effects" involve disjunctive prejacents. For example, 'I can come tomorrow or on Friday' seems to entail that I can come tomorrow and that I can come on Friday. This would be a problem for the quantificational model, since the existence of a world where I come tomorrow or on Friday clearly doesn't entail the existence of a world where I come tomorrow. However, there are strong reasons to think the apparent entailment is a special kind of scalar implicature. (For example, it disappears under negation.) How exactly the implicature arises is a controversial topic. Notably, many promising accounts – including [Klinedinst 2007] and [Schwarz 2018], but also e.g. [Fox 2007] or [Franke 2011] – presuppose the quantificational account of modals. Thus, rather than presenting a problem for the quantificational canon, "free choice effects" arguably support that account.

trivial, the de dicto reading is more salient. On the account I have put forward, the de dicto reading of ability statements does not validate the K principle: there may well be accessible worlds where you ϕ or ψ transparently even if there are no accessible worlds where you ϕ transparently nor worlds where you ψ transparently. Thus (23b) and (23c) can be false while (23a) is true.

Notice that this response not only blocks the problematic inference from (23a) to [(23b)$\vee$(23c)] in some way or other; it also gives an intuitively plausible account of *why* the inference fails: the problem is that you can't knowingly/intentionally/transparently draw a red card.[23]

Finally, the *argument from hyperintensionality*. Possibile-worlds accounts are coarse-grained. If the set of possible worlds where an agent ϕs is identical to the set of worlds where she ψs, then the quantificational account implies that the agent has the ability to ϕ just in case she has the ability to ψ. But intuitively, (24a) is true while (24b) is false.

(24) a. Cyril can recite the digits $3, 1, 4, 1, 5, 9, 2, 6, 5, 3$.

b. Cyril can recite the first 10 digits of π.

The prejacents in (24a) and (24b) are (necessarily, and a priori) equivalent: there is no possible world at which anyone recites the digits $3, 1, 4, 1, 5, 9, 2, 6, 5, 3$ without reciting the first ten digits of π.

Again, you can probably anticipate my response. On their (non-salient) de re reading, (24a) and (24b) really do stand and fall together. But on their more salient de dicto reading, they can come apart. The hyperintensionality comes from the hyperintensionality of warrant and belief: the fact that there are worlds where Cyril transparently recites $3, 1, 4, 1, 5, 9, 2, 6, 5, 3$ does not imply that there are worlds where he transparently recites the first 10 digits of π.

6 Conclusion

Let me summarize what I've tried to show. I have argued that many ability statements display a kind of de re/de dicto ambiguity. The ambiguity is easy to miss because one of the two readings is often more salient than the other. The classical quantificational approach, as developed by Kratzer and others, only accounts for the de re reading. Many

23 We can also explain why some authors have resisted Kenny's argument and claimed that (23) actually does imply [(23b)$\vee$(23c)]: the argument is valid if all the statements are read de re.

My own experience is that the longer I think about the relevant instances of (23b) and (23c), or indeed (1) or (2), the more I become inclined to say that they are simply true. For some reason, the de re reading seems to becomes more salient. If the de dicto reading arises through a contextual restriction on the interpretation of the prejacent, this might be an instance of "accommodation" (see [Lewis 1979]): it is generally easier to lift a contextual restriction than to re-impose it.

objections and counterexamples to this approach accordingly involve ability statements that are most naturally interpreted de dicto.

To account for de dicto abilities, I have suggested that ability modals are sensitive to contextual restrictions on the interpretation of their prejacent. A de dicto ability is an ability to perform the relevant act in a particular way. The same mechanism is arguably in play when we say that a person with a paralysed arm can't move her arm, or when we say that a mediocre piano player can't play the piano; it might also explain the gradability of ability statements.

More specifically, I have suggested that having an ability de dicto requires having the ability to perform the relevant act "transparently", meaning that the agent can perform the act in such a way that she has warrant to belief that she will perform the act merely on the basis of the relevant decision, provided the act is under her volitional control.

Why should we be interested in this peculiar kind of ability? Because it plays a central role in normative contexts, when we consider what an agent ought to do. Informally, an act is a candidate for being obligatory only if the agent knows what she would have to do in order to perform the act. The de dicto sense of ability, I have argued, is tailored to this normative job description.

References

G.E.M. Anscombe [1957]: *Intention.* Ithaca, NY: Harvard University Press

John L. Austin [1961]: "Ifs and Cans". In *Philosophical Papers,* London: Oxford University Press, 151–180

Rajesh Bhatt [2006]: *Covert modality in non-finite contexts.* Walter de Gruyter

Virginia Mary Brennan [1993]: *Root and epistemic modal auxiliary verbs.* PhD Thesis, University of Massachusetts at Amherst

Mark A Brown [1988]: "On the logic of ability". *Journal of Philosophical Logic*, 17(1): 1–26

Erik Carlson [1999]: "The Oughts and Cans of Objective Consequentialism". *Utilitas*, 11(01): 91–96

Michael Fara [2008]: "Masked abilities and compatibilism". *Mind*, 117(468): 843–865

Danny Fox [2007]: "Free choice and the Theory of Scalar Implicatures". In U. Sauerland and P. Stateva (Eds.) *Presupposition and Implicature in Compositional Semantics,* Basingstoke: Palgrave Macmillan, 71–120

M. Franke [2011]: "Quantity implicatures, exhaustive interpretation, and rational conversation". *Semantics and Pragmatics*, 4(1): 1–82

Harry Frankfurt [1969]: "Alternate possibilities and moral responsibility". *The Journal of Philosophy*, 66(23): 829–839

Holly S. Goldman [1978]: "Doing the Best One Can". In A. Goldman and J. Kim (Eds.) *Values and Morals,* Dordrecht: Reidel, 185–214

John Greco [2007]: "The nature of ability and the purpose of knowledge". *Philosophical issues*, 17(1): 57–69

Valentine Hacquard [2009]: "On the interaction of aspect and modal auxiliaries". *Linguistics and Philosophy*, 32(3): 279

— [2010]: "On the event relativity of modal auxiliaries". *Natural language semantics*, 18(1): 79–114

— [2011]: "Modality". In C. Maienborn, K. von Heusinger and P. Portner (Eds.) *Semantics: An International Handbook of Natural Language Meaning,* vol 2. Berlin: de Gruyter, 1484–1515

Brian Hedden [2012]: "Options and the subjective ought". *Philosophical Studies*, 158(2): 343–360

A.M. Honoré [1964]: "Can and Can't". *Mind*, 73(292): 463–479

Terence Horgan [1979]: "'Could', possible worlds, and moral responsibility". *The Southern Journal of Philosophy*, 17(3): 345–358

Romy Jaster [2016]: *Agents' abilities*. PhD Thesis, HU Berlin

Anthony Kenny [1976]: "Human abilities and dynamic modalities". In J. Manninen and R. Tuomela (Eds.) *Essays on explanation and understanding,* Springer, 209–232

Nathan W. Klinedinst [2007]: "Plurality and Possibility". Phd Dissertation, UCLA

Angelika Kratzer [1977]: "What 'must' and 'can' must and can mean". *Linguistics and Philosophy*, 1(3): 337–355

— [1981]: "The notional category of modality". *Words, Worlds, and Contexts*: 38–74

— [1986]: "Conditionals". *Chicago Linguistics Society*, 22(2): 1–15

— [1991]: "Modality". In A. von Stechow and D. Wunderlich (Eds.) *Semantik Handbuch,* Berlin: de Gruyter, 639–650

— [2012]: *Modals and Conditionals*. Oxford: Oxford University Press

Saul Kripke [1980]: *Naming and Necessity*. Oxford: Blackwell

Keith Lehrer [1968]: "Cans without ifs". *Analysis*, 29(1): 29–32

David Lewis [1979]: "Scorekeeping in a Language Game". *Journal of Philosophical Logic*, 8: 339–359. Reprinted in [Lewis 1983]

— [1983]: *Philosophical Papers I*. New York, Oxford: Oxford University Press

— [1986]: *On the Plurality of Worlds*. Malden (Mass.): Blackwell

John Maier [2015]: "The agentive modalities". *Philosophy and Phenomenological Research*, 90(1): 113–134

— [2018a]: "Abilities". In Edward N. Zalta (Ed.) *The Stanford Encyclopedia of Philosophy,* Metaphysics Research Lab, Stanford University, spring 2018 edition

— [2018b]: "Ability, modality, and genericity". *Philosophical Studies*, 175(2): 411–428

Matthew Mandelkern, Ginger Schultheis and David Boylan [2017]: "Agentive modals". *Philosophical Review*, 126(3): 301–343

George E. Moore [1912]: *Ethics*. London: Williams and Norgate

P.H. Nowell Smith [1960]: "Ifs and cans". *Theoria*, 26(2): 85–101

John L. Pollock [2002]: "Rational Choice and Action Omnipotence". *The Philosophical Review*, 111: 1–23

Paul Portner [2009]: *Modality*. Oxford University Press

Wolfgang Schwarz [2017]: "Options and Actions". URL https://umsu.de/papers/options.pdf. Manuscript

— [2018]: "Discourse, Diversity, and Free Choice". URL https://umsu.de/papers/freechoice2.pdf. Manuscript

Kieran Setiya [2008]: "Practical knowledge". *Ethics*, 118(3): 388–409

Nicholas Southwood and David Wiens [2016]: ""Actual" does not imply "feasible"". *Philosophical Studies*, 173(11): 3037–3060

Travis Timmerman and Yishai Cohen [2016]: "Moral Obligations: Actualist, Possibilist, or Hybridist?" *Australasian Journal of Philosophy*, 94(4): 672–686

Peter van Inwagen [1983]: *An Essay on Free Will*. Oxford: Clarendon Press

David Velleman [1989]: *Practical Reflection*. Princeton: Princeton University Press

Barbara Vetter [2013]: "'Can' without Possible Worlds: Semantics for Anti-Humeans". *Philosopher's Imprint*, 13(16)

Kadri Vihvelin [2004]: "Free will demystified: A dispositional account". *Philosophical Topics*, 32(1/2): 427–450

— [2013]: *Causes, laws, and free will: Why determinism doesn't matter*. Oxford: Oxford University Press

Kai von Fintel and Irene Heim [2002]: "Lecture notes on intensional semantics". MIT

Paul Weirich [1983]: "A decision maker's options". *Philosophical Studies*, 44(2): 175–186

Ann Whittle [2010]: "Dispositional abilities". *Philosophers' Imprint*, 10(12): 1–23