

Within 'Reason': A Study of Normative Language

by

Eliot Watkins

B.A. (Hons), University of Cambridge, 2017

Submitted to the Department of Linguistics and Philosophy in partial fulfilment of the requirements of the degree of

Doctor of Philosophy

at the

Massachusetts Institute of Technology

September 2023

© 2023 Eliot Watkins. All rights reserved.

The author hereby grants to MIT a nonexclusive, worldwide, irrevocable, royalty-free license to exercise any and all rights under copyright, including to reproduce, preserve, distribute and publicly display copies of the thesis, or release the thesis under an open-access license.

Signature of Author: .....  
Department of Linguistics and Philosophy  
August 23, 2023

Certified by: .....  
Justin Khoo  
Professor of Philosophy  
Thesis Supervisor

Accepted by: .....  
Brad Skow  
Laurence S. Rockefeller Professor of Philosophy  
Chair of the Committee on Graduate Students

# Within ‘Reason’

## A Study of Normative Language

by

Eliot Watkins

Submitted to the Department of Linguistics and Philosophy on August 23, 2023, in partial fulfilment of the requirements for the degree of  
Doctor of Philosophy

### ABSTRACT

What do we mean when we say that someone ought to do something? What do we mean when we say that someone has a reason to do something? What do we mean when we say that someone has more reason to do one thing rather than another? The primary goal of this project is to shed light on these semantic questions.

The picture of normative talk that I develop across this thesis has a distinctive feature: the notion of *a* reason (roughly, a fact that counts in favour of something) isn’t given any fundamental role to play. Instead, the meanings of ‘ought’, ‘must’ and ‘is a reason for...’ are all understood in terms of something gradable – they’re understood in terms of facts about *how much* reason there is for something to be done.

Chapter One focuses on deontic modals like ‘ought’ and ‘must’. I argue that the standard semantics for these expressions is incompatible with the idea that facts about what you ought to do are connected with facts about what you have reason to do. I develop a new semantics for deontic modals which builds-in the connections between ought and reasons from the ground up.

Chapter Two centres on ‘reason’. We use ‘reason’ as both a count noun (as in “there is *a* reason for you to read my dissertation”) and a mass noun (as in “there is *some* reason for you to read my dissertation”). I argue that the best semantics for ‘reason’ will treat the mass form as fundamental. ‘Reason’ is a predicate of a particular kind of state – the state someone is in when they have reason to do something. I turn this result into an argument against the enduringly popular idea that count noun reasons are normatively fundamental.

Chapter Three stays with reasons. According to a standard picture, normative reasons do not extend beyond the boundaries of agency. If something isn’t an agent – if it can’t do rudimentary practical reasoning – then there can’t be normative reasons for it to do one thing rather than another. I argue that this standard picture gets things totally wrong: there are reasons for non-agents to be certain ways and do certain things. We must not analyse what it is to be a reason by appealing to distinctively agential capacities.

Thesis Supervisor: Justin Khoo  
Title: Professor of Philosophy

# Contents

Acknowledgements.....	5
Introduction.....	6
<b>Chapter 1: How to be Reasonable about the Meaning of ‘Ought’.....</b>	<b>10</b>
§1 Set Up.....	13
1.1 The Connection Principles .....	13
1.2 The Standard Semantics and monotonicity.....	16
§2 Normative Reason Claims are Non-Monotonic.....	18
§3 Monotonicity Strikes Back.....	25
§4 Towards a More Reasonable Semantics.....	31
4.1 The basic idea.....	31
4.2 A sketch of a new semantics for ‘ought’ and ‘have to’.....	33
4.3 Two tweaks to the semantics.....	38
4.4 The semantics resolves the tension.....	41
§5 Conclusion.....	44
Appendix.....	45
A.1. ‘Have to’ entails ‘ought to’.....	45
A.2. The reasonable semantics captures the Connection Principles.....	47
<b>Chapter 2: A Critique of Pure Reasons.....</b>	<b>50</b>
§1 What ‘Reason’ Can’t Mean.....	53
1.1 ‘Reason’ picks out a weight? .....	56
1.2 ‘Reason’ picks out reasons? .....	58
1.3 ‘Reason’ picks out reasons and weights? .....	62
1.4 Objection: agents always have infinitely many reasons.....	63
1.5 A different Count Fundamentalist account? .....	65
§2 What ‘Reason’ Might Mean.....	66
2.1 ‘Reason’ as a predicate of states.....	66
2.2 An outstanding puzzle.....	74
§3 Why This Matters.....	77
§4 Conclusion.....	84
<b>Chapter 3: Reasons for Non-Agents.....</b>	<b>86</b>
§1 Against Agentialism.....	91
1.1 Initial cases – reasons for non-agents to be certain ways.....	91
1.2 An early objection – normative reasons target doings.....	93
1.3 Reasons for non-agents to do certain things.....	97
§2 Objections and Replies.....	101
2.1 The Epistemic Diagnosis.....	101
2.2 The Ambiguity Diagnosis.....	103

2.3 The Anthropomorphism Diagnosis.....	108
§3 Implications.....	110
3.1 Implications for theories of normative reasons.....	110
3.2 Implications for the scope of normativity.....	115
§4 Conclusion.....	120
<b>Bibliography</b> .....	121

## Acknowledgements

I have most reason to thank my wonderful committee: Justin Khoo, Kieran Setiya, Brad Skow and Jack Spencer. Thank you for your guidance and encouragement. I could not have asked for better mentors.

I also have various good reasons to thank: Ido Benbaji, Sam Berstler, Gunnar Björnsson, Kenneth Black, Alex Byrne, Thomas Byrne, Kevin Dorst, Bennett Eckert, Patrick Elliott, Jonathan Fiat, Kai von Fintel, Enrico Flor, Daniel Fogal, Kelly Gaus, Felix-Antoine Gelineau, Sophie Gibert, Caspar Hare, Jessica Heine, Michele Odisseas Impagnatiello, Lilian Jin, Philipp Mayr, Gareth Norman, Tamar Schapiro, Haley Schilling, Roger Schwarzschild, Bob Stalnaker, Eliza Wells, Quinn White, Yoad Winter, and Steve Yablo.

I have two weighty reasons to thank Josh Pearson. Thank you for the comments you gave me on this dissertation, and thank you for helping make my time at MIT so enjoyable.

All three chapters benefited from feedback at MIT's *Work In Progress* series. Chapter 1 was improved in light of a useful set of comments from an anonymous reviewer. A version of Chapter 2 was presented at 2023's *St Louis Annual Conference of Reasons and Rationality*. I'd like to thank that audience for helpful questions and feedback, and particularly Chris Blake-Turner for a detailed and insightful set of written comments.

I'd also like to thank my parents and my brother. Thank you for the constant generous encouragement, and thank you for the love you send across the sea.

Finally, of course, Abi Thwaites. It would take a second dissertation to list all the reasons I have to thank you.

## Introduction

I ought to give more money to charity. I have to grade these papers by tomorrow. I should not be late to appointments. I must not steal money from my friends. There are good reasons for me to become a vegan. There's some reason for me to be distrustful of politicians. There is no reason for me to avoid stepping on cracks in the sidewalk. I have more reason to tear the bandaid off quickly than to tear it off slowly.

We use these kinds of sentences to describe the normative world – when we deliberate, advise, chastise, and praise. Within philosophy, we appeal to reasons and oughts to state our theses and build our theories. But what do sentences like these *mean*? The primary goal of this project is to shed light on that semantic question.

The three chapters are unified by their focus on the ways in which we talk about what there's reason to do. But the picture of normative talk that I develop across this thesis has a distinctive (and perhaps surprising) feature: the notion of *a* reason – roughly, a fact that counts in favour of something – isn't given any fundamental role to play. Instead, this project's central idea is that the meanings of 'ought', 'must' and 'is a reason for...' should all be cashed out in terms of facts about how *much* reason we have to do various things. It's this gradable notion – the notion of there being (some amount of) reason to act – that drives our talk about the normative world.

We start, in Chapter 1, with 'ought' and 'must'. Take a look at the following two popular and plausible principles. First principle: facts about what you ought to do are tightly bound-up with

facts about what you have reason to do. If you ought to watch *Oppenheimer* then there must be good reason for you to watch *Oppenheimer*. Similarly, if you ought to watch *Oppenheimer* then there must be more reason for you to watch *Oppenheimer* than not watch it. Second principle: ought is ‘upwards monotonic’ – if it’s true that you ought to  $\Phi$  in some specific way, then you ought to  $\Phi$ . If you ought to wear red socks today, then you ought to wear socks today. This principle is not only plausible on its face, but also follows from the standard quantificational semantics for ‘ought’.

I argue that these two independently plausible principles are, in fact, incompatible. Respecting the connections between reasons and ought requires giving up monotonicity, and so requires giving up the standard semantics for deontic modals. We need something new. I develop a non-monotonic semantics for ‘ought’ which builds connections with reason-talk directly into its semantics. The theory I sketch aims to have its cake and eat it: the semantics is developed with an eye both to capturing the reason/ought connections and to accommodating the data which made monotonicity look so appealing in the first place.

In Chapter Two, I turn my attention from oughts to reasons. Sometimes we use ‘reason’ as a count noun, to pick out facts in the world. For example: “the fact that you like Greta Gerwig is a reason for you to watch *Barbie*”. Other times, we use ‘reason’ as a mass noun, to pick out something we can have more or less of. For example: “there is not much reason for you to watch *Oppenheimer*” or “there is more reason for you to watch *Barbie* than *Oppenheimer*”.

It's natural to think that the count sense of 'reason' *comes first* – to think that the meaning of a comparative like 'there's more reason to watch Barbie than Oppenheimer' is comparing the weightiness of the particular reasons to watch one movie with the weightiness of the particular reasons to watch the other. I argue that this 'count fundamentalist' semantics is a non-starter – when we talk about what there's reason to do, we're not talking about the reasons there are.

I use this result to motivate a semantics for mass and count 'reason' which inverts the standard priority of reasons over reason. The central notion on this semantics isn't *there being a reason for S to  $\Phi$* , it's *there being some amount of reason for S to  $\Phi$* . 'Reason' is a predicate of a particular kind of state – the state which someone is in which they have reason to do something. I use this semantics to build an argument against the Reasons First programme (the enduringly popular idea that count noun reasons are normatively fundamental).

Chapter Three stays with reasons. According to a standard picture, normative reasons do not extend beyond the boundaries of agency. If something isn't an agent – if it can't do rudimentary practical reasoning, at the very least – then there can't be normative reasons for it to do one thing rather than another.

I argue that this standard picture gets things wrong. There are reasons for non-agents to *be* certain ways: there's good reason for the fence around my chicken coop to be tall, because if it were short my chickens would escape. There are also reasons for non-agents to *do* certain things. If my kitchen is full of smoke, then there's a good reason for my smoke detector to ring.



This has important consequences for theories of normative reasons. For one thing, we should not analyse reasons by appealing to distinctively agential properties or capacities (e.g. the standards of good practical reasoning). Instead, I suggest, all one needs to get reason-ascriptions going is some kind of teleology. There's a reason for my smoke alarm to go off in the presence of fire because that's what smoke alarms do when they're functioning well.

The guiding thought behind this project is that natural language semantics and ethical theory have something to learn from one another. As I conceive of them, ethics and semantics share a common subject matter – both are in part concerned with giving theories of when it's true to say that someone ought (or has reason) to do something. Given this common subject matter, we should want our best ethical and semantic theories to fit together. There's a good reason to theorise about 'reason' with an eye to both disciplines.

## Chapter 1: How to be Reasonable About the Meaning of ‘Ought’

Last year, you solemnly promised your child, Billy, that you’d visit Disney World together over the summer vacation. But now the costs are starting to pile up, the weather doesn’t look good, and your partner wouldn’t be able to come with you. What should you do? Your partner tells you:

(1) (All things considered...) You ought to take Billy to Disney World this summer.

Here’s something you might begin to wonder (perhaps in an effort to avoid having to make a decision): what does (1) mean? Under what conditions is what your partner said true?

Luckily, you have two friends you can call to help. Your first friend is a natural language semanticist. Surely if anyone has anything helpful to say about the meaning of (1), she does. She tells you that ‘ought’ (like ‘should’, ‘must’ and ‘have to’) is a necessity modal – it is a kind of universal quantifier. The sentence “books must be returned to the library on time” says that in all the worlds which are ideal from the point of view of the library’s rules, books are returned on time. (1) works the same way: it says that in all worlds that are ideal according to some contextually salient standard (perhaps the standard of ‘being best from the moral point of view’, or ‘being best all things considered’), you take your child to Disney World.

Your second friend is an ethicist. He doesn’t think about the *meaning* of ‘ought’ too much, but he does take himself to have something helpful to say about the truth-conditions for (1). He tells you this: many ethicists think facts about what you ought to do are tied up with facts about your *reasons*

and the weightiness of those reasons. Suppose that (1) is true – all things considered, you ought to go to Disney World with Billy this summer. Then, at the very least, there must be some consideration that counts in favour of you doing so. Even stronger: if you ought to go to Disney World, then there must be *good* reason for you to do so. Strongest of all, he says, many ethicists think (1) is true if and only if there is *most* reason for you to take Billy to Disney World this summer: taking all relevant considerations into account – the promise, the cost, the weather, etc. – overall these considerations most favour taking the vacation.

After speaking to both of your friends, you end up more confused than you started. If the ethicist is right about the deep connections between the facts about what agents ought to do and the facts about what there's reason to do, then that would seem to impose constraints on the truth-conditions for (1). But these reason-connections aren't reflected anywhere in the semanticist's meaning for 'ought'. Why not?

I think you are right to be confused. But the situation is even worse than you realise. It's not just that the standard semantics is silent on the connections between reasons and 'ought'. The first thing I'm going to argue in this chapter is that the standard semantics is actually *incompatible* with those connections. Your friends cannot both be right.

Here's the short version of why that is. According to the standard semantics, 'ought' is *monotonic*: whenever  $p$  entails  $q$ , *ought*  $p$  entails *ought*  $q$ .<sup>1</sup> So, if your semanticist friend is right, (1) entails:

---

<sup>1</sup> Some authors, like Cariani (2011) and Blumberg & Hawthorne (2021a), call this principle 'Inheritance'. The 'monotonic'/'non-monotonic' terminology is more commonly used by natural language semanticists, as well as Von Steup (2012) and Muñoz & Spencer (2020).

- (2a) You ought to take Billy to Disney World sometime over the next fifty years.
- (2b) You ought to take someone's child to Disney World this summer.
- (2c) You ought to take Billy to Disney World or the local abattoir.

By contrast, I'll argue, reason-talk is *non-monotonic*. Generally, it needn't follow from the fact that your reasons support you doing the action in (1) that your reasons support you doing the actions in (2). It's a consequence of this mismatch in logical behaviour, I'll argue, that if reasons and oughts are connected in the way that the ethicist thinks they are, then the standard semantics for 'ought' must be wrong.

The rest of the chapter explores where we should go next in light of this. I start by making things even more puzzling. If we want to hang on to the connections between reasons and 'ought', then we must develop a semantics for 'ought' that's not monotonic, and so doesn't validate the inference from (1) to (2).<sup>2</sup> However, there's troubling data which suggests that any semantics for 'ought' worth its salt must validate that inference. In particular, we will see that the behaviour of strong modals like 'have to' or 'must' seems to strongly favour the monotonicity hypothesis.<sup>3</sup> Given all that, it looks difficult to see how we can respect both the ethicist and the semanticist at once: in order to accommodate the connections between reasons and 'ought' we need to abandon monotonicity, but doing *that* risks flying in the face of the linguistic evidence.

---

<sup>2</sup> For one important example of a non-monotonic semantics for 'ought', see Cariani (2011).

<sup>3</sup> See the data pointed to in von Fintel (2012), Muñoz and Spencer (2020), Blumberg & Hawthorne (2021a).

I aim to dissolve the tension and bring the semanticist and ethicist together by giving a meaning for ‘ought’ that does justice to the insights of each. I sketch a non-monotonic semantics for ‘ought’ which should be attractive to the ethicist insofar as it builds the connections with reasons right into the semantics itself, but also attractive to the semanticist insofar as it gives a plausible explanation of the data that originally favoured monotonicity. The trick will be to treat strong modals like ‘have to’ and ‘must’ differently from weak modals like ‘ought’ and ‘should’ – the former behave monotonically, but the latter do not.

§1 outlines some putative connections between reasons and ‘ought’, sketches the standard semantics, and defines monotonicity. In §2, I argue that environments like “S has good reason to...” are non-monotonic, and hence that the standard semantics is incompatible with the connections between reasons and ought. In §3, I explain why it’s so difficult to give a non-monotonic semantics for ‘ought’ that respects the linguistic data. In §4, I sketch a new non-monotonic semantics for ‘ought’ which aims to do just that.

## **§1 Set up**

### **1.1 The Connection Principles**

The idea that reasons and oughts are connected tends to be emphasised most by those in the ‘Reasons First’ programme, who want to understand all normativity in terms of reasons, or the

relation of ‘being a reason for...’.<sup>4</sup> But you don’t need to endorse that kind of ambitious reductive thesis to think that facts about what agents ought to do are bound up with facts about how much reason they have to do things – to think that if an ought-claim is true of an agent, a corresponding reason-claim must be true of them too.

One classic way of tightly binding together oughts and reasons is the idea that an agent ought to do the action which is, on balance, most favoured by all the relevant considerations:

*A ought to  $\Phi$ , all things considered, just in case the total reason for him to  $\Phi$  is stronger than the total reason to do anything else. (Setiya, 2014, p.229)*

When we have decisive reasons, or most reason, to act in some way, this act is what we should or ought to do in what we can call the decisive-reason-implying senses. Even if we never use the phrases ‘decisive reason’ or ‘most reason’, most of us often use ‘should’ and ‘ought’ in these reason-implying senses. (Parfit, 2011, p.33)

These authors are giving voice to our first Connection Principle:

(Most Reason)      Necessarily, S ought to  $\Phi$  iff S has most reason to  $\Phi$ .

---

<sup>4</sup> For classic statements of the reasons first programme, see Scanlon (1998) and Parfit (2011). For a more recent argument in favour of putting reasons first, see Schroeder (2021).

If it's true that I should bring Billy to Disney World this summer, then it can't be the case that I have even more reason to keep him at home. And vice versa: if bringing Billy to Disney World this summer is the thing I have most reason to do, then that's what I should do.

The basic idea behind (Most Reason) seems modest and plausible, but for much of this chapter we don't even need something as strong as this. Here are some weaker putative connections:

(Good Reason)      If S ought to  $\Phi$ , then, on balance, there is good reason for S to  $\Phi$ .

(More Reason)      If S ought to  $\Phi$ , then there is more reason for S to  $\Phi$  than not- $\Phi$ .

Together, (Good Reason), (More Reason) and (Most Reason) form what I call the *Connection Principles*.

The Connection Principles gain most of their appeal from the fact that something like them is endorsed or presupposed in much recent philosophy and moral theory. But there is also linguistic evidence that should make the natural language semanticist want to capture them. At least some of these kinds of connections are deeply embedded in our ordinary way of using oughts and shoulds. Just try conjoining (1) with the negation of the corresponding reason-claim:

(3a) # You should, all things considered, take Billy to Disney World, but on balance there isn't good reason for you to do so.

(3b) # You should, all things considered, take Billy to Disney World, but there's just as much reason for you not to.

On the face of it, a good semantics for 'ought' should at least be compatible with these kinds of connections. I'm going to argue that the standard semantics isn't.

## 1.2 The standard semantics and monotonicity

In contemporary natural language semantics, modals like 'ought', 'should', 'must', 'have to' and 'need to' are, in effect, analysed as universal quantifiers.<sup>5</sup> Imagine we have some function  $f$  which maps contexts to sets of worlds that are ideal according to the standards of that context. We can then state a simplified version of the standard semantics as follows:

**Standard Semantics:** *ought p* is true in context  $c$  iff all worlds in  $f(c)$  are worlds where  $p$ .

Various tweaks or additions to the standard semantics have been proposed, but the core idea that 'ought' is a universal quantifier over some set of ideal worlds remains incredibly popular. It's the core idea that's incompatible with the Connection Principles, as I'll now argue.

The argument goes via the claim that according to the standard semantics, 'ought' is *upwards monotonic*. The label 'upwards monotonic' sounds technical, but the basic idea can be put simply.

---

<sup>5</sup> For a classic statement of the standard semantics see Kratzer (1991). Kratzer's own semantics is more sophisticated, distinguishing between two contextually-determined variables (a modal base and an ordering source).



Let's say *linguistic environments* are sentences where phrases have been removed – things like “Some ... walked” or “Sally wanted...”. A linguistic environment is upwards monotonic just in case replacing the expression that fills the blank space “...” with something *weaker* preserves truth. For example, “Justin drank ...” is upwards monotonic: if “Justin drank beer” is true, then “Justin drank liquid” is true too.<sup>6</sup>

Upwards monotonicity is to be contrasted with downwards monotonicity and non-monotonicity. An environment is downwards monotonic just in case replacing the expression that fills the “...” with something *stronger* preserves truth. For example, “Justin didn't drink ...” is downwards monotonic: if “Justin didn't drink liquid” is true, then “Justin didn't drink beer” is true too.

An environment is non-monotonic iff it's neither upwards nor downwards monotonic. For example, “exactly two students drank ...” is non-monotonic. It doesn't follow from the fact that exactly two students drank beer that exactly two students drank liquid – perhaps the rest of the class drank water. So it's not upwards monotonic. It also doesn't follow from the fact that exactly two students drank beer that exactly two students drank lager – maybe one of them drank an ale. So the environment isn't downwards monotonic either.

---

<sup>6</sup> What is it, precisely, for one expression to be stronger/weaker than another? When we're dealing with propositions, the relevant notion of strength is entirely familiar:  $p$  is stronger than  $q$  iff  $p$  entails  $q$ . For expressions that don't denote propositions, we'll need to look to some more expansive notion of entailment – for example, Blumberg & Hawthorne suggest appealing to the definition of predicate entailment in Schlenker (2009, 14). This paper is going to end up focusing on action descriptions, and there the relevant notion of strength/weakness is intuitive:  $\Phi$ ing is stronger than  $\Psi$ ing iff it's impossible for an agent to  $\Phi$  without thereby  $\Psi$ ing. The action-description *eat a red apple* is stronger (more specific) than *eat an apple*, but weaker (less specific) than *eat a dark red apple*. I suspect that, for our purposes, nothing particularly deep turns on these issues – all of our discussion could be recast in terms of propositions and propositional entailment if we wanted.

It is easy to see that the standard semantics makes ‘ought’ upwards monotonic (or just ‘monotonic’, for short). Suppose “you ought to take your child to Disney World” is true in some context  $c$ . Then all the worlds in  $f(c)$  are worlds in which you take your child to Disney World. But the proposition *that you take your child to Disney World* entails *that you take someone’s child to Disney World*. So all the worlds in  $f(c)$  are worlds in which you take someone’s child to Disney World. So the sentence “you ought to take someone’s child to Disney World” will be true too.

The fact that the standard semantics makes ‘ought’ upwards monotonic has been controversial, especially in the philosophical literature.<sup>7</sup> But defenders of the standard semantics have historically found ingenious ways of blunting the force of putative counterexamples.<sup>8</sup> The next part of this chapter can be seen as presenting a new argument against monotonicity: only a semantics for ‘ought’ which rejects monotonicity will be able to capture the Connection Principles.

## **§2 Normative reason claims are non-monotonic**

The Connection Principles link ought-claims together with linguistic environments of this form:

On balance, S has good reason to ...

S has more reason to ... than they have not to ...

S has most reason to ...

---

<sup>7</sup> See the barrage of objections given in Cariani (2011), especially Ross’s paradox and his presentation of Jackson & Pargetter (1986)’s Professor Procrastinate.

<sup>8</sup> See von Fintel (2012) and the references therein.

In this section I am going to argue that these environments are non-monotonic (neither upwards nor downwards entailing).

It's easy to show that they are not downwards monotonic. For example, last Friday evening I had good reason to drink some wine, but I didn't have good reason to drink *twenty glasses* of wine. The less obvious claim I want to defend is that this kind of reason-talk is not upwards monotonic either: sometimes  $\Phi$ ing entails  $\Psi$ ing, S has good reason to  $\Phi$ , but S does not have good reason to  $\Psi$ . *Mutatis mutandis* for 'more reason' and 'most reason'.<sup>9</sup>

Take a look at the following:

*Suzy's game*: Suzy's bills are starting to pile up. Luckily for Suzy, she's about to be given the opportunity to earn some money. Suzy is led into a room with a table. On top of the table there's a game, which consists of a panel with buttons labelled 1-100. In order to play the game, she simply needs to press one of the buttons. If Suzy presses the winning button, £100 will be deposited into her bank account. If she presses any other button, £1000 will be extracted from her bank account. Also on the table is £90 in cash. Suzy can either play the game (push a button) or she can take the cash, but she can't do both. Suzy knows all this, but she doesn't know the location of the winning button. As it happens, the winning button is Button 3.

---

<sup>9</sup> Contra Portmore (2019), Kiesewetter (2015). Note that Kiesewetter, at least, is concerned with a slightly different question – the question of whether, when you have reasons to  $\Phi$ , you have reasons (of the same strength) to take the necessary *means* to  $\Phi$ ing.

There are, I take it, several strong reasons for Suzy to push Button 3. For example:

- (4) The fact that Suzy needs money is a reason for her to push Button 3.

Of course, Suzy doesn't know that there is this weighty reason for her to push Button 3. We may choose to say, then, that there's an ethically significant sense of 'have' according to which Suzy doesn't *have* such a reason. But (4) seems to have a true 'objective' reading nonetheless: it is bad for Suzy to be poor, and pushing Button 3 would go some (small) way to alleviating her poverty; so the fact she needs money is a reason for her to push Button 3.<sup>10</sup> What's more, there are no equally weighty reasons not to push Button 3 – it is not as though pushing that button comes with any downsides or risks. So, taking all relevant considerations into account, it seems that her (objective) reasons on balance support pushing Button 3. All of the following, then, have true readings:

- (5a) On balance, there is good reason for Suzy to push Button 3.
- (5b) There is more reason for Suzy to push Button 3 than there is for her not to.
- (5c) There is most reason for Suzy to push Button 3.

---

<sup>10</sup> There are some ways of thinking about normative reasons which may reject a true reading of (4). For example, theories which identify normative reasons with premises in good pieces of practical reasoning might reject it (Setiya 2014). I ask those who are allergic to objective reason claims to bear with me. We will see in footnote 18 that it is not necessary that the agent be ignorant of some normatively relevant fact, in order to get these kinds of cases to work.

*Pushing Button 3* entails *playing the game*.<sup>11</sup> So, if ‘good reason’, ‘more reason’ and ‘most reason’ were upwards monotonic, the following should have true readings too:

- (6a) On balance, there is good reason for Suzy to play the game.
- (6b) There is more reason for Suzy to play the game than there is reason for her not to.
- (6c) Suzy has most reason to play the game.

But these, I hope you agree, simply seem false.

Here are a few ways of drawing out this intuition. Firstly, note that if the sentences in (6) were true, then we’d expect to be able to point to some particular, weighty normative reasons for Suzy to play the game. But it doesn’t seem like there are any such weighty reasons. Granted, if Suzy plays the game then she might push Button 3, and that would be best; but the mere fact she *might* do something very good by playing the game is not a weighty reason for her to do so. Alternatively, take the fact that Suzy needs money. We just saw that this fact counted in favour of pushing Button 3. It does so, presumably, because if Suzy pushes Button 3, she will be richer. But this consideration does not seem to count in favour of her playing the game. After all, it simply isn’t true that if Suzy *plays the game* then she’ll be richer. On the contrary, if Suzy plays the game then she’s likely to push a losing button and be down £1000. (So, far from counting in favour of playing

---

<sup>11</sup> Suzy *plays the game* by pushing some button or other (either Button 1, or Button 2, ... or Button 100), and if Suzy pushes Button 3 then she’s pushed some button or other.

the game, it appears that the most salient consideration in the case – the fact that Suzy needs money – counts *against* her playing.)<sup>12</sup>

A second way of drawing out the reason non-monotonicity intuition is to focus on the amount of reason Suzy has to take the £90 cash. This comparative has a true objective reading:

(7a) There is more reason for Suzy to push Button 3 than take the cash.

After all, pushing Button 3 gets her an extra £10 and comes with no downsides. If ‘reason’ behaves monotonically, then the following should be true too:

(7b) There is more reason for Suzy to play the game than take the cash.

But I cannot get any true reading of (7b). Pushing some button or other between 1 and 100 is likely to lose Suzy £1000. There’s clearly more reason for Suzy to take the cash than do *that*.

---

<sup>12</sup> In the background here is the idea that particular count noun ‘reason’ ascriptions are themselves non-monotonic. Sometimes,  $\Phi$ ing entails  $\Psi$ ing,  $p$  is a reason for S to  $\Phi$ , but  $p$  isn’t a reason for S to  $\Psi$ . I take it that we see this kind of non-monotonicity exhibited even in quite run-of-the-mill cases. In the case of taking Billy to Disney World, the following is clearly true:

(\*) The fact you promised Billy you’d drive him to Disney World is a reason to drive Billy to Disney World.

But (\*) does not seem to entail (\*\*):

(\*\*) The fact you promised Billy you’d drive him to Disney World is a reason to drive to Disney World while there is another person in your car.

Suppose I start driving towards Disney World with Billy’s little sister in the passenger seat, leaving Billy at home to tidy his room. On the hypothesis that reasons transmit over entailment, the promise I made to Billy counts in favour something I am doing (namely, driving to Disney World while there is another person in my car). But that’s clearly wrong – the promise I made counted in favour of driving Billy in particular, not driving just anyone.

Here's a final case for thinking that the sentences in (5) are true but those in (6) are false. Consider some of the heuristics one would first reach for, in order to cash out 'the amount of reason an agent has to do something, overall'. These tend to be non-monotonic.<sup>13</sup> One heuristic, roughly following Michael Smith (1994), says that the amount of objective reason you have to do an act corresponds to the degree to which a fully informed, virtuous observer would want you to do it. How strongly would a fully informed, virtuous observer desire that Suzy push Button 3? Very strongly. If Suzy pushes Button 3, she'll win £100. Would they desire, to the same degree, that Suzy plays the game? It seems not, at least on the face of it. If Suzy plays the game, she's likely to lose £1000.<sup>14</sup>

A second heuristic says that you have most (objective) reason to  $\Phi$  iff a beneficent, omniscient observer would advise you to  $\Phi$ .<sup>15</sup> It seems to me that a beneficent, omniscient observer would give Suzy the advice *push Button 3*. But they wouldn't give Suzy the advice *play the game*, because if Suzy listened to that advice things would (most likely) go very badly.<sup>16</sup>

A final heuristic says that, at least sometimes, reason-to-act tracks goodness, or expectation of goodness. At least in cases where you're sufficiently isolated from other-regarding concerns, if

---

<sup>13</sup> 'Maximalism' about reasons and ought is an exception (Portmore 2019). According to maximalism, S ought to  $\Phi$  (or has most reason to  $\Phi$ ) iff  $\Phi$ ing is entailed by S's best maximally fine-grained option. Maximalism provides a pleasingly elegant solution to some of the puzzles in this paper, but at the cost of making surprising (and, in my view, implausible) predictions about the logic of the mass noun 'reason' in English. For example, maximalism would seem to predict that there is a true reading of the comparative in (7b) "There's more reason for Suzy to play the game than there is for her to take the cash", as the best way of doing the former is better than the best way of doing the latter. But, as I suggested above, this looks to be the wrong result – there are weighty reasons for Suzy to take the cash, but it's hard to identify weighty reasons for her to play the game.

<sup>14</sup> These are my judgements, but the precise monotonicity properties of desire-ascriptions are contested. See Phillips-Brown (2021), Blumberg & Hawthorne (2021b), and von Fintel (2021) for some interesting recent work.

<sup>15</sup> Kolodny and MacFarlane (2010, p. 117) suggest an analogous heuristic for the objective 'ought'.

<sup>16</sup> Contrast the argument in Muñoz and Spencer (2020, p.84).

$\Phi$ ing is expected to be better than  $\Psi$ ing, then you have more reason to  $\Phi$  than  $\Psi$ . But expected value behaves non-monotonically: sometimes,  $\Phi$  entails  $\Psi$  but the expected value of  $\Psi$ ing is lower than the expected value of  $\Phi$ ing. Again, the case of Suzy illustrates this: using the credence function of some suitably idealised observer, the expected value of Suzy's *pushing Button 3* is high, but the expected value of her *playing the game* is low.

Before wrapping up, there's one final thing to notice about this case. If I'm right that the sentences in (5) are true (when read objectively), but those in (6) are straightforwardly false, then we'd expect to be able to conjoin the sentences in (5) with the *negations* of the sentences in (6). And, indeed, I think we can do so. These sound acceptable to my ears:

- (8a) There is (on balance) good reason for Suzy to push Button 3, but there isn't (on balance) good reason for her to play the game.
- (8b) There is more reason for Suzy to push Button 3 than for her not to, but there isn't more reason for Suzy to play the game than for her not to.
- (8c) There is most reason for Suzy to push Button 3, but there isn't most reason for Suzy to play the game.<sup>17</sup>

Overall, reason-talk looks to be non-monotonic. The alternative view (that when  $\Phi$ ing entails  $\Psi$ ing, S's having good reason to  $\Phi$  entails that S has good reason to  $\Psi$ ) is in tension with our intuitions

---

<sup>17</sup> If I'm right that (8a-c) are acceptable, then this implies that monotonicity failures can occur *within a single context*. This is significant, as some recent attempts to accommodate our apparently non-monotonic intuitions in the case of 'ought' have appealed to shifts in contextually-salient sets of options (Blumberg & Hawthorne, 2021a). As it happens, I'm inclined to think that this sort of contrastivism is the wrong way to go with 'ought' too (see footnote 24).



in a case like *Suzy's Game* and implausibly detaches reason-to-act from heuristics like desirability, advisability and expected value.<sup>18</sup>

### §3 Monotonicity strikes back

Cast your mind back to your two friends: the semanticist and the ethicist. We now have all the materials we need to show that they cannot both be right about 'ought'. The sentence:

(9) Suzy ought to push Button 3.

has a true (objective) reading. If the standard semantics for 'ought' is correct then 'ought' is upwards monotonic, in which case (9) entails:

(10) Suzy ought to play the game.

If the ethicist is right, and the Connection Principles hold, then (10) entails:

---

<sup>18</sup> In *Suzy's Game*, Suzy does not have all normatively-significant information, as she is ignorant of the location of the winning button. But all we need in order to elicit the relevant non-monotonicity judgement is for it to be unlikely that Suzy will push Button 3 conditional upon her playing the game, and we can get this even when Suzy has full knowledge of all the relevant facts. Take a case like this:

*Suzy's Game Without Ignorance*: Once again Suzy is choosing between pushing a button and taking the cash. And once again, Button 3 remains her best option with the £90 cash in a close second place. This time, Suzy knows that Button 3 is the winning button. But Suzy is superstitious: she sees 3 as an unlucky number, so she would strongly prefer not to push it. Pushing Button 3 is still an option for Suzy (if she decided to push Button 3, she would do so) but if she plays the game she's very likely to push a different button.

Again, Suzy has most reason to push Button 3. But it still seems false to say that she has most reason to play the game. For again we can ask: what reasons are there for her to do so? As before, the fact she needs money isn't a weighty reason to play the game: if she plays the game, she'll almost definitely end up with *less* money than she started with.

(11) On balance, there is good reason for Suzy to play the game.

But I just argued that (11) is false. There is good reason for Suzy to do the specific action *push Button 3*, but there isn't good reason for her to do the non-specific action *play the game*.

Either the ethicist is right about the Connection Principles, in which case we need to reject the monotonicity of 'ought', or the semanticist is right about the standard semantics for 'ought', in which case we need to abandon the Connection Principles. What do we do now? In the next section, I develop a semantics for 'ought' that attempts to respect the insights of each. But before I do, I want to make the puzzle even more difficult for ourselves.

For all I've said thus far, you might think that the solution to all of this is easy: just abandon the standard semantics. The Connection Principles look deeply plausible, and I have not yet given you an argument for thinking that dropping the monotonicity of 'ought' will incur any great costs. But in fact things are not so simple. There is important data that's hard to square with the claim that 'ought' is non-monotonic. Some have concluded from this data that the monotonicity of 'ought' is non-negotiable – that any respectable semantics for 'ought' will validate the inference from *Suzy ought to push Button 3* to *Suzy ought to play the game*.<sup>19</sup>

---

<sup>19</sup> I originally found the following data in Muñoz & Spencer (2020), who themselves cite von Fintel (2012). See also Blumberg & Hawthorne (2021a).

The data centres on two facts about the modal ‘have to’ which are difficult to explain if it isn’t upwards monotonic. First, consider an abominable conjunction like the following:

- (12) # Charlie doesn’t have to wear a tie or a scarf, but he has to wear a tie. (Von Fintel, 2012, p.13)

The hypothesis that ‘have to’ is upwards monotonic gives an elegant explanation for why (12) is unacceptable. The proposition *that Charlie wears a tie* entails the proposition *that Charlie wears a tie or a scarf*. So if “Charlie has to wear a tie” is true – and ‘have to’ is monotonic – then “Charlie has to wear a tie or a scarf” must be true too. So, if ‘have to’ is monotonic, the two conjuncts contradict each other. If ‘have to’ isn’t monotonic, we need to look elsewhere for an explanation of why this sentence is degraded.

The second fact about ‘have to’ is more involved, but we can skip over most of the details here without losing too much.<sup>20</sup> One trick used by linguists to help discern whether an environment is upwards monotonic is to check whether its negation is *downwards* monotonic. And one way of figuring out whether an environment is downwards monotonic is to check whether it licenses Negative Polarity Items (NPIs) – expressions like *ever, at all, any, lift a finger, budge an inch*. And, indeed, NPIs can appear felicitously under ‘don’t have to’:

- (13a) You don’t have to pick up any groceries. (Muñoz and Spencer, 2020, p.83).

---

<sup>20</sup> Those who want more details should consult von Fintel (2012) and the references therein.

(13b) I would like it if you relaxed on the couch while I tidy up: you don't have to lift a finger.

(13c) The boss will give in to your demands if you are stubborn: you don't have to budge an inch.

If you want to claim that 'have to' is non-monotonic, you'll need to explain why it is that NPIs can felicitously appear in (13). It's not initially clear what form such an explanation could take.

Let's pause. This was supposed to be an argument supporting the monotonicity of 'ought'. Why are we taking a detour through a different modal, 'have to'? The detour is necessary. The tests for monotonicity – both the abominable conjunction test and the NPI test – rely on embedding the relevant linguistic environment under negation. And while 'have to' can easily be negated (as in, 'don't have to'), 'ought' and the other weak necessity modals (e.g. 'should') cannot.<sup>21</sup> Consider:

(14) You ought not dance.

There is no interpretation of (14) on which the negation takes wide scope over the modal (i.e. as *NOT ought dance*). The same thing even holds for expressions that lexicalise negation: the

---

<sup>21</sup> See Iatridou & Zeijlstra (2013) for interesting discussion on this issue.

sentence ‘I doubt you ought to dance’ is more naturally read as *I believe you ought not dance* than *I believe it’s not the case that you ought to dance*.<sup>22</sup>

Regardless, it is tempting to think that we can learn something about the logical properties of ‘ought’ by looking at the logical properties of ‘have to’. The two are obviously closely connected expressions in English, but cross-linguistically the story is even more interesting. Von Stechow & Iatridou (2008, 2022) point to so-called ‘transparent ought’ languages such as Spanish and Greek, in which ‘ought’ is pronounced by applying a layer of special morphology to ‘have to’: in these languages, sentences which would be translated into English as “Sam ought to dance” are expressed via something close to “Sam would have to dance”. Given this cross-linguistic picture, it’s difficult to resist the idea that the two modals have some kind of core meaning in common. Therefore, while the above data does not amount to a direct argument that ‘ought’ is monotonic, it does amount to an indirect one – the behaviour of ‘have to’ is unexpected and unexplained if ‘ought’ is non-monotonic.

---

<sup>22</sup> In English you can apparently force the negation to take wide scope with ‘it’s not the case that...’, ‘it’s not true that...’, ‘I doubt it’s true that’, etc. But it’s by no means obvious that these kinds of sentences will be probative in the present context. (We cannot just assume that “it’s not the case that you ought to  $\Phi$ ” or “it’s not true that you ought to  $\Phi$ ” have the logical form *NOT you ought to  $\Phi$* ). Regardless, even if the data were probative, the results are inconclusive. The following sound less bad to me than the abominable ‘have to’ conjunctions:

- ? It’s not the case that Suzy ought to play the game, but she nonetheless ought to push Button 3.
- ? It’s not true that Suzy ought to push some button or other between 1 and 100, but she ought to push Button 3.

It’s also unclear whether NPIs are felicitous under “it’s not the case that you ought to”:

- ? It’s not the case that you ought to ever worry about me.
- ? It’s not the case that you ought to budge an inch.
- ? It’s not the case that you ought to lift a finger.

I think it's fair to say that the above behaviour of 'have to' is a thorn in the side of all existing theories of 'ought' that drop monotonicity (as interesting and inventive as those theories may otherwise be).<sup>23</sup> One notable exception to this generalisation is Blumberg & Hawthorne (2021a), who have recently given an elegant contrastivist semantics for 'ought' that attempts to steer a course between the above pro-monotonicity data and our apparently non-monotonic intuitions in cases like Jackson & Pargetter's famous *Professor Procrastinate*. Blumberg & Hawthorne's analysis deserves a thorough treatment, but due to limitations of space I only include my worries about their view in a footnote.<sup>24</sup> Suffice it to say that while their analysis gives an elegant explanation of the pro-monotonicity data, I am sceptical that it will be a good fit for the Connection Principles.

In summary, a good semantics for 'ought' needs to do a kind of balancing act. It cannot give 'ought' a monotonic semantics, for otherwise the Connection Principles are lost. But it needs to say *enough* about the logic of 'ought' and 'have to' to explain the pro-monotonicity data. Steering

---

<sup>23</sup> This data is especially cutting against the non-monotonic contrastivist theories of Snedegar (2012) and Cariani (2011). See the discussion of Cariani's semantics in von Fintel (2012).

<sup>24</sup> Blumberg & Hawthorne's (B&H's) semantics is roughly this: *ought p* is true in *c* iff the top-ranked option in the contextually-determined set of alternatives  $ALT_c$  entails *p*. Recall the famous Procrastinate case – Procrastinate is the best person to write a review, but if he agrees to write the review he is unlikely to do so. B&H claim that the reason why we hear "Procrastinate ought to accept" as false is that speakers have a (defeasible) preference for evaluating ought-claims relative to sets of options that are immediately open to agents, e.g. {*accept*, *don't accept*}, rather than sets of temporally extended actions, e.g. {*accept & write*, *accept & don't write*, *don't accept*}. My worry with this strategy is that there seem to be cases in which 'ought' behaves non-monotonically, but all the actions happen at a single instant. In fact, *Suzy's Game* is a case of just this sort. B&H's semantics would seem to predict that the following has an easily accessible true reading:

(\*) Suzy ought to play the game.

After all, the top-ranked option in {*Push Button 1*, *Push Button 2... Take the cash*} is *Push Button 3*, which entails playing the game. And B&H cannot rule out using this set of actions to assess (\*), as all the actions in the set are immediately open to Suzy. But I struggle to access any true reading of (\*) – even an 'objective' reading. Regardless, if B&H bite the bullet and argue that (\*) is indeed true, then they will struggle to capture the Connection Principles: I have argued that there is not good reason, overall, for Suzy to play the game.

a course between those two constraints turns out to be rather difficult. In the next section, I suggest one way we might try and do so.

## §4 Towards a more reasonable semantics

### 4.1 The basic idea

We're caught in a tension between the linguistic data that favours monotonicity and the insight from ethical theory that ought-facts are hooked up with reasons to act. I think there's a way out of the tension. The trick is to take the pro-monotonicity data at face value – it really does show that 'have to' and the other strong modals (e.g. *must*, *need to*) behave monotonically. But it does not show us much more than that – in particular, it does not show us that 'ought' and the other weak modals (e.g. *should*) are monotonic. I think a plausible semantics can be developed on which 'ought' and 'have to' share similar meanings, but have different logics. As far as I know, this manoeuvre hasn't been attempted before, but it seems to be a plausible way of plotting a course between the two constraints.

It's already well-established that there's some difference in meaning between the weak and the strong modals – between what you should do and what you must do.<sup>25</sup> We all understand the rules governing seatbelt usage on commercial flights: you have to keep your seatbelt on when the seatbelt sign is lit-up; when the seatbelt sign isn't lit up, you don't have to keep your seatbelt on

---

<sup>25</sup> See von Stechow & Iatridou (2008, 2022)

but you still ought to. Semantically, the distinction surfaces via asymmetric entailments. Notice the contrast between:

(15a) You ought to stay buckled in, but you don't have to.

(15b) # You have to stay buckled in, but it's not the case that you ought to.

It's at least a possibility, then, that this difference in meaning results in a difference in their logics. And in fact this kind of view seems a natural place to end up. Intuitively, an agent has to  $\Phi$  in some circumstances just in case any way of doing something minimally decent in those circumstances requires  $\Phi$ ing. But this has a distinctively monotonic feel: if  $\Phi$ ing requires  $\Psi$ ing, and any way of doing something minimally decent requires  $\Phi$ ing, then any way of doing something minimally decent requires  $\Psi$ ing. If you *have to* push Button 3, then you have to play the game. If you *must* drive Billy to Disney World, then you must drive him to Disney World or the local abattoir.

The idea of what an agent ought to do pulls in the other direction. Whether or not an agent ought to  $\Phi$  depends on whether or not they have most reason to  $\Phi$ . And that notion, we've seen, is a non-monotonic one. It needn't follow from the fact that your reasons most strongly favour *pushing Button 3* that your reasons most strongly favour *playing the game*.

Below, I sketch a new semantics for 'ought' and 'have to'. The semantics is unusual in two respects. Firstly, it gives 'have to' and 'ought' different logics in the way just suggested. Secondly, I will construct the meanings of 'have to' and 'ought' by appealing to facts about how much reason



there is for agents to pursue various options. This approach builds-in the Connection Principles from the ground up.

## 4.2 A sketch of a new semantics for ‘ought’ and ‘have to’

Let’s start by granting ourselves a simplifying assumption: for any agent  $S$  and option of theirs  $\Phi$ , there’s always some answer to the question *how much reason is there for  $S$  to  $\Phi$ ?* The most convenient way of representing this formally is via a function  $\mathbf{R}$  that maps pairs of agents and options  $\langle S, \Phi \rangle$  to real numbers that represent how much reason  $S$  has to take that option.<sup>26</sup> For short, I call  $\mathbf{R}(\langle S, \Phi \rangle)$  the *reasonableness* of  $\Phi$ ing for  $S$ .<sup>27</sup> (In what follows I normally suppress the argument place for the agent, and speak as though  $\mathbf{R}$  is a function from actions to numbers, rather than pairs of actions *and agents* to numbers – it should always be kept in mind that, strictly speaking, I mean the latter).<sup>28</sup>

There is a good case to be made that ‘reason’ is context sensitive, even once we’ve fixed on its normative sense. For example, it seems I can say “there is most reason for Suzy to push Button 3”

---

<sup>26</sup> We do not need to assume this rich cardinal structure, but doing so makes for more straightforward exposition. All that is ultimately required is that we can order an agent’s options by their  $\mathbf{R}$ -value. There are interesting and difficult questions, of course, about whether even assuming this much is too strong: perhaps sometimes the amount of reason an agent has to  $\Phi$  is incomparable to the amount of reason they have to  $\Psi$ . I myself am not persuaded, partly due to the sorts of concerns raised in Dorr et al. (2022). Nonetheless, I am optimistic that there are ways of building on the semantics that would make it compatible with some types of incomparability. (Perhaps, for example, we could have  $\mathbf{R}$  map options to  $n$ -tuples representing the reasonableness of the option along various dimensions – compare Hedden & Muñoz, forthcoming).

<sup>27</sup> In Chapter 2 I analyse the mass noun ‘reason’ in terms of states. Ultimately, then, we may want to unify the two semantics by having  $\mathbf{R}$  map states to real numbers.

<sup>28</sup> If one wanted, one could rewrite the semantics with *ought* taking a proposition as an argument, and with  $\mathbf{R}$  mapping propositions (*that  $S$   $\Phi$ s*) into real numbers. I prefer to stick with pairs of agents and actions, but only for convenience.

and communicate (truly) that given all the facts, pushing Button 3 would be best, or (falsely) that given the information she has available to her, pushing Button 3 would be best. For now, we can make room for this kind of context-sensitivity by positing many **R**-functions – at least as many as there are normative flavours of the English word ‘reason’. In general, we can say that  $\mathbf{R}_c$  corresponds to whichever sense of (mass noun) ‘reason’ is operative in the context  $c$ . Far more needs to be said, but the analysis I give here should be compatible with a wide range of approaches to the context sensitivity of reason talk.

For  $\mathbf{R}(\langle S, \Phi \rangle)$  to be well-defined,  $\Phi$  needs to be an option for  $S$ . It shouldn’t matter too much which theory of options you endorse, but what is important is that an agent’s options can be more or less fine-grained. One of Suzy’s options is *push Button 3*, and another option is *play the game*. Given the non-monotonicity of ‘reason’, sometimes a specific action will be more reasonable than a non-specific action it entails. For example:

$$\mathbf{R}(\langle \text{Suzy}, \text{push Button 3} \rangle) > \mathbf{R}(\langle \text{Suzy}, \text{play the game} \rangle)$$

Now, given an agent and their options, we can order the actions open to them by their **R**-value. Plausibly, an incomplete slice of the **R**-values in Suzy’s case is as follows:

$$\mathbf{R}(\text{push Button 3}) > \mathbf{R}(\text{take the cash}) > \mathbf{R}(\text{play the game}) > \mathbf{R}(\text{push Button 1}) \dots$$

One final piece to put in place. I follow Cariani (2011) and Snedegar (2012) in assuming that context also determines some baseline value  $b_c$ . I call actions whose **R**-value exceeds  $b_c$  ‘minimally

reasonable’. For now it will do no harm to assume the baseline in the above model goes just below *play the game*:

$$\mathbf{R}(\textit{push Button 3}) > \mathbf{R}(\textit{take the cash}) > \mathbf{R}(\textit{play the game}) > b_c > \mathbf{R}(\textit{push Button 1}) \dots$$

Enough of the set-up – onto the semantics. Let’s start with the strong modal, ‘have to’. The intuitive gloss I gave of ‘have to’ above said that *S has to  $\Phi$*  is true iff every minimally decent option of S’s requires  $\Phi$ ing. The semantics I give to ‘have to’ is just a formalisation of this idea in our new framework. Here’s the rough idea (again I suppress the agent variable ‘S’ for readability):

**Approximate meaning of ‘have to’:**

$$[[S \textit{ has to } \Phi]]^c = 1 \text{ iff } \quad \forall \Psi [(\mathbf{R}_c(\Psi) \geq b_c) \rightarrow \Psi \textit{ entails } \Phi]$$

i.e. *S has to  $\Phi$*  is true in context *c* iff all minimally reasonable actions (according to the standards of reasonableness in *c*) are ways of  $\Phi$ ing.

For example, “you have to take Billy to Disney World” is true iff every minimally reasonable option of yours entails taking Billy to Disney World.

What about ‘ought’? The asymmetrical entailments between the strong and weak modals suggest that while ‘have to’ tracks what needs to be done in order to be good enough, ‘ought’ tracks what’s *best*. That might suggest that ‘S ought to  $\Phi$ ’ is true just when  $\Phi$ ing has the highest **R**-value:

**Ought as best (incorrect semantics):**

$$[[S \text{ ought to } \Phi]]^c = 1 \text{ iff } \quad \forall \Psi [\mathbf{R}_c(\Phi) \geq \mathbf{R}_c(\Psi)]$$

But this cannot be quite right. We can imagine it's best for Suzy to push Button 3 using her right arm – perhaps her left arm is slightly sore. If so, when fleshed out a little more, our (still incomplete) model should look something like this:

$$\begin{aligned} & \mathbf{R}(\text{push Button 3 with right arm}) > \mathbf{R}(\text{push Button 3}) > \mathbf{R}(\text{take the cash}) > \mathbf{R}(\text{play the game}) \\ & > b_c > \mathbf{R}(\text{push Button 1}) \dots \end{aligned}$$

Now *push Button 3* isn't the highest ranked option. Nonetheless, the following still seems true:

(16) Suzy ought to push Button 3.

We can rescue the insight behind the 'ought as best' suggestion by weakening it slightly: 'ought  $\Phi$ ' doesn't say  $\Phi$ ing beats *all* other actions, only that it beats every action *except, perhaps, those which are themselves ways of  $\Phi$ ing*. This gloss is truth-conditionally equivalent to the claim that any action which is more reasonable than  $\Phi$  is itself a way of  $\Phi$ ing:

**Approximate meaning of 'ought':**

$$[[S \text{ ought to } \Phi]]^c = 1 \text{ iff } \quad \forall \Psi [(\mathbf{R}_c(\Psi) \geq \mathbf{R}_c(\Phi)) \rightarrow \Psi \text{ entails } \Phi]$$

i.e.  $S$  ought to  $\Phi$  is true in context  $c$  iff any action that's more reasonable than  $\Phi$  (according to the standards of reasonableness in  $c$ ) is itself just a way of  $\Phi$ ing.

(16) is true iff every action more reasonable than *push Button 3* is a way of pushing Button 3. According to the model above, the only action with a higher  $\mathbf{R}$ -value than *push Button 3* is *push Button 3 with right arm*. That's a way of pushing Button 3, so (16) is true on our model.

With this entry, we secure a rather nice result: the meanings of the strong modal 'have to' and the weak modal 'ought' are very similar indeed. *S has to  $\Phi$*  says that any action more reasonable than the *baseline* is a way of  $\Phi$ ing. *S ought to  $\Phi$*  says that any action more reasonable than  $\Phi$  is a way of  $\Phi$ ing. An equivalent way of thinking about the semantics for 'ought', then, is that 'ought  $\Phi$ ' means the same thing as 'have to  $\Phi$ ', except when we make an ought-claim the baseline is moved upwards from its default value to a stricter standard.

The similarity between the two semantic entries is especially important insofar as we want our theory to be a good fit for the cross-linguistic data. Recall that many languages' equivalent of 'ought' is pronounced by applying a dedicated kind of morphology to their equivalent of 'have to'. Von Stechow & Iatridou (2022) suggest that the function of this special kind of morphology is to signal that some parameter is taking a different value from the default that's set by context. The above semantics is a nice fit for that suggestion: the modal 'have to' is sensitive to  $b_c$ , whereas 'ought' shifts the baseline away from the default.

### 4.3 Two tweaks to the semantics

There are two tweaks to make to the semantics before I tie everything together. (Readers who are less interested in the details and just want the payoff can skip this subsection).

First tweak. As things stand, ‘S has to  $\Phi$ ’ is true only if *every* option more reasonable than the baseline entails  $\Phi$ ing, and ‘S ought to  $\Phi$ ’ is true only if *every* option more reasonable than  $\Phi$  entails  $\Phi$ ing. This turns out to be too strong. Recall that whatever Suzy does, it’s better if she uses her right arm to do it. Intuitively, both the claim “Suzy ought to push Button 3” and the claim “Suzy ought to use her right arm” are true.<sup>29</sup> But this isn’t captured by the semantics currently on the table. Suppose the **R**-values look like this:

$$\dots \mathbf{R}(\textit{push Button 3 with right arm}) > \mathbf{R}(\textit{push Button 3}) > \mathbf{R}(\textit{use her right arm}) > \dots$$

As it stands, our semantics is going to predict that “Suzy ought to use her right arm” is false. For it to be true, every action more reasonable than *use her right arm* needs to entail using her right arm. But the presence of the non-specific action *push Button 3* means that this condition isn’t met. Suzy can push Button 3 in good ways (by using her right arm) or bad ways (by using her left arm, or her nose).

---

<sup>29</sup> Importantly, the way I’m imagining the case is that Suzy’s using her right arm is independent of whether she pushes a button or takes the cash. e.g. It’s not more likely, conditional on her using her right arm, that she’ll push a losing button.

Intuitively, the fact that *pushing Button 3* can be performed badly (without using her right arm) is not enough to falsify the claim that she *ought* to use her right arm. Instead, it seems as though all the ought-claim turns on is whether the best *specific* things Suzy could do entail using her right arm. This thought inspires the following simple fix to the semantics. All we need to do is restrict our quantifiers to a set  $A$ , whose members are all and only the agent’s maximally-specific options:

$$\begin{aligned} [[S \text{ ought to } \Phi]]^c = 1 \text{ iff} & \quad \forall \Psi \in A [(\mathbf{R}_c(\Psi) \geq \mathbf{R}_c(\Phi)) \rightarrow \Psi \text{ entails } \Phi] \\ [[S \text{ has to } \Phi]]^c = 1 \text{ iff} & \quad \forall \Psi \in A [(\mathbf{R}_c(\Psi) \geq b_c) \rightarrow \Psi \text{ entails } \Phi] \end{aligned}$$

“Suzy ought to use her right arm” is true iff any *maximally specific* action more reasonable than using her right arm is, itself, just a way of using her right arm. The action *push Button 3* isn’t specific, so it’s not in  $A$ , and so it cannot falsify the ought-claim.

What is it for an action to be maximally specific? Officially, for the purposes of this project, I’m agnostic. Perhaps standards of specificity vary between contexts, and so while an action like *push Button 3* doesn’t count as maximally specific in the above context, it would qualify as maximally specific in other contexts. Alternatively, perhaps the semantics should ultimately be supplemented with some substantive theory of maximal specificity.<sup>30</sup> The only remaining thing we need in order to get the semantics to work is a constraint on  $\mathbf{R}$  to make sure it’s well-behaved with respect to  $A$ . I relegate discussion of this constraint to the first part of the Appendix.

---

<sup>30</sup> For example, perhaps an action is in  $A$  just in case it’s a member of the maximally fine-grained partition of actions that the agent has direct control over. Thanks to Jack Spencer for this suggestion.

That’s the first tweak to the semantics. We can deal with the second tweak more quickly. As things stand, ‘S has to  $\Phi$ ’ and ‘S ought to  $\Phi$ ’ can be true even when the reasonableness of  $\Phi$  is arbitrarily low. This is not only intuitively implausible, but it also ends up breaking the entailment from ‘have to’ to ‘ought’<sup>31</sup>. Here is my suggested fix: for ‘S has to  $\Phi$ ’ or ‘S ought to  $\Phi$ ’ to be true,  $\Phi$ ing must be minimally reasonable. The most straightforward way of cashing this out formally is to impose the following presupposition:

*S has to  $\Phi$*  (and *S ought to  $\Phi$* ) are well-defined in  $c$  only if either  $\mathbf{R}(\Phi) > b_c$  or  $\mathbf{R}(\sim\Phi) > b_c$ .<sup>32</sup>

If neither  $\Phi$  nor  $\sim\Phi$  appear among the options above the baseline, then neither ‘ought  $\Phi$ ’ nor ‘has to  $\Phi$ ’ can be true. This presupposition allows the semantics to capture the entailment from ‘have to’ to ‘ought’ as well as some connection principles between ‘have to’ and ‘reason’ claims. See the Appendix for more details.

This gives us our final entries for the two modals.

---

<sup>31</sup> Suppose our model was

$\mathbf{R}(\textit{push Button 3}) > b_c > \mathbf{R}(\textit{take the cash}) > \mathbf{R}(\textit{play the game}) > \mathbf{R}(\textit{push Button 1}) \dots$

Suppose too, for simplicity, that every action except *play the game* is maximally specific. Without tweaking the semantics, *Suzy has to play the game* would be true, as every specific above the baseline entails playing the game. But *Suzy ought to play the game* would be false, as there’s a specific action more reasonable than *play the game* which doesn’t entail playing the game – namely, taking the cash.

<sup>32</sup> The second disjunct is needed to accommodate cases where deontic modals are negated (we don’t want *you don’t have to commit mass murder* to presuppose that *commit mass murder* is minimally reasonable).



### A reasonable meaning for ‘have to’

$$[[S \text{ has to } \Phi]]^c = 1 \text{ iff } \quad \forall \Psi \in A \ [(\mathbf{R}_c(\Psi) \geq b_c) \rightarrow \Psi \text{ entails } \Phi]$$

presupposition: either  $\mathbf{R}_c(\Phi) \geq b_c$  or  $\mathbf{R}_c(\text{not-}\Phi) \geq b_c$

In other words: *S has to*  $\Phi$  is well-defined only if either  $\Phi$ ing or not- $\Phi$ ing is minimally reasonable, and (assuming it’s well-defined) it’s true iff all specific, minimally reasonable actions entail  $\Phi$ ing.

### A reasonable meaning for ‘ought to’

$$[[S \text{ ought to } \Phi]]^c = 1 \text{ iff } \quad \forall \Psi \in A \ [(\mathbf{R}_c(\Psi) \geq \mathbf{R}_c(\Phi)) \rightarrow \Psi \text{ entails } \Phi]$$

presupposition: either  $\mathbf{R}_c(\Phi) \geq b_c$  or  $\mathbf{R}_c(\text{not-}\Phi) \geq b_c$

In other words: *S ought to*  $\Phi$  is well-defined only if either  $\Phi$ ing or not- $\Phi$ ing is minimally reasonable, and (assuming it’s well-defined) it’s true iff any specific action that’s more reasonable than  $\Phi$  is just, itself, a way of  $\Phi$ ing.

## §4.4 The semantics resolves the tension

This semantics already gets us much of what we wanted. In particular, we can show that ‘ought’ is non-monotonic, with the Connection Principles built right into its meaning, but that ‘have to’ behaves monotonically.

Let's start with 'have to'. We get an elegant explanation for why a sentence like (12) will always be unassertable:

(12) #You don't have to wear a tie or a scarf but you have to wear a tie.

If the second conjunct of (12) is true, then all the specific actions above  $b_c$  entail wearing a tie. But that means all the specific actions above  $b_c$  entail wearing a tie or a scarf. So "you have to wear a tie or a scarf" must be true too, which contradicts the first conjunct. So (12) can never be true – hence its unassertability.

A similar story goes for the NPI data: our semantics can explain why expressions like *any*, *ever*, *lift a finger*, etc. can appear grammatically under 'don't have to'. According to our semantics, *have to* is 'Strawson monotonic': whenever  $\Phi$ ing entails  $\Psi$ ing, if  $c$  is a context in which both *S has to  $\Phi$*  and *S has to  $\Psi$*  are well-defined, then if *S has to  $\Phi$*  is true in  $c$ , so is *S has to  $\Psi$* . There's good evidence for the claim that Strawson monotonic environments license NPIs when negated.<sup>33</sup>

So, both pieces of pro-monotonicity data from §3 are accommodated by our semantics. However, despite the (Strawson) monotonicity of 'have to', 'ought' is non-monotonic. Take a look at this model of Suzy's situation again (and assume, for simplicity, that all the actions except *play the game* are maximally specific):

$\mathbf{R}(\textit{push Button 3}) > \mathbf{R}(\textit{take the cash}) > \mathbf{R}(\textit{play the game}) > b_c > \mathbf{R}(\textit{push Button 1}) \dots$

---

<sup>33</sup> For more on Strawson monotonicity and the licensing of NPIs, see von Stechow (1999).

Recall, *ought*  $\Phi$  is true iff (roughly speaking) any specific action that's more reasonable than  $\Phi$  is itself just a way of  $\Phi$ ing. Given our semantics, the above model makes (16) true but (17) false.

(16) Suzy ought to push Button 3.

(17) Suzy ought to play the game.

(16) is true because every specific action more reasonable than *push Button 3* is itself a way of pushing Button 3 (in fact, on the above simplified model, there is no action more reasonable than *push Button 3*). (17) is false because there's a specific action (namely, *take the cash*) which is more reasonable than *play the game* but which isn't a way of playing the game.

Generally, we will get counterexamples to the monotonicity of 'ought' whenever  $\Phi$ ing entails  $\Psi$ ing, the agent ought to  $\Phi$ , but there is a big gap between the amount of reason the agent has to  $\Phi$  and the amount of reason they have to  $\Psi$ . The gap between the reasonableness of  $\Phi$ ing and  $\Psi$ ing is 'big enough' to get monotonicity failure when there is a specific action which *isn't* a way of  $\Psi$ ing that appears between the two.

The fact that our semantics for 'ought' is non-monotonic means it is in a position to capture the Connection Principles. But in fact we end up with something more exciting than this: by constructing our semantics in terms of reason-to-act, the Connection Principles are simply consequences of the meanings we've assigned to 'have to' and 'ought'. (See the Appendix).

## §5 Conclusion

The idea that the ought-facts and the reason-facts are connected is both highly intuitive and philosophically fruitful. But the implications of this idea for natural language semantics have tended to be overlooked. This chapter argued for one such implication: if the Connection Principles hold, then ‘ought’ is not monotonic, and hence the standard semantics for deontic modals must be rejected. The second half of the chapter sketched the direction that I think we should go in next. I presented a semantics which took the Connection Principles seriously by building them into the meanings of deontic modals.

This approach to the meaning of ‘ought’ opens up exciting questions for both the ethicist and the semanticist. Of particular interest to the ethicist is my way of drawing the distinction between strong and weak modals (‘have to’ vs. ‘ought’). What determines the baseline value  $b_c$  in a context? Why do we care to track such a value? What might the connections and distinctions between ‘ought’, ‘have to’ and ‘reason’ teach us about the nature of moral obligation and supererogation?

One of the most important semantic issues that remains, I take it, is the question of what the above theory should say about *non-deontic* flavours of ‘ought’ and ‘have to’. As it happens, I am optimistic that something like the above semantics could be extended or adapted to apply in these other domains. For notice that there appear to be connections between ‘ought’ and ‘reason’ even outside the domain of ethics. Take epistemic flavours of ‘ought’:

(18) (According to my evidence...) Bob ought to be in his office.

An assertion of (18) would seem to commit a speaker to the following:

(19a) I have good reason to believe that Bob in his office.

(19b) I have more reason to believe that Bob is in his office than that he is at home.

From the perspective of the ‘reasonable’ semantics, this is very intriguing data. It is too early to tell how far the semantics can be pushed, but the idea that modality is in general shot through with connections to reasons is exciting and worth taking seriously.

## Appendix

### A.1 ‘Have to’ entails ‘ought to’

Recall that in §4.3 I restricted the quantifiers in ‘have to’ and ‘ought’ to a set of maximally specific actions  $A$ :

$[[S \textit{ has to } \Phi]]^c = 1$  iff  $\forall \Psi \in A [(\mathbf{R}_c(\Psi) \geq b_c) \rightarrow \Psi \textit{ entails } \Phi]$   
 presupposition: either  $\mathbf{R}_c(\Phi) \geq b_c$  or  $\mathbf{R}_c(\textit{not-}\Phi) \geq b_c$

$[[S \textit{ ought to } \Phi]]^c = 1$  iff  $\forall \Psi \in A [(\mathbf{R}_c(\Psi) \geq \mathbf{R}_c(\Phi)) \rightarrow \Psi \textit{ entails } \Phi]$   
 presupposition: either  $\mathbf{R}_c(\Phi) \geq b_c$  or  $\mathbf{R}_c(\textit{not-}\Phi) \geq b_c$

In order to capture the entailment between ‘have to’ and ‘ought’, we need to ensure that **R** is well-behaved with respect to *A*. The following does the trick:

**Reason Requires a Reasonable Way (RRRW):**

$$\forall \Psi [\exists \Psi_w \in A (\Psi_w \text{ entails } \Psi \text{ and } \mathbf{R}_c(\Psi_w) \geq \mathbf{R}_c(\Psi))]$$

i.e. For all actions  $\Psi$ , there exists a specific way of  $\Psi$ ing,  $\Psi_w$ , such that  $\Psi_w$  is at least as reasonable as  $\Psi$ .

In effect, RRRW amounts to this plausible claim: it’s never the case that  $\Psi$ ing is strictly more reasonable than every specific way of  $\Psi$ ing. e.g. You can never have strictly more reason to push *some button or other* than you do to push any particular button.<sup>34</sup>

With this principle in hand, we can prove that if ‘S has to  $\Phi$ ’ is true in a context *c*, then so is ‘S ought to  $\Phi$ ’:

---

<sup>34</sup> Suppose we didn’t have this constraint. Then there’s nothing to rule out a model that looks like this (where  $\Phi_w$  is a specific way of  $\Phi$ ing, and  $\Psi_1, \Psi_2\dots$  are specific ways of not  $\Phi$ -ing):

$$\mathbf{R}(\Phi_w) > \mathbf{R}(\text{not-}\Phi) > \mathbf{R}(\Phi) > \mathbf{R}(\Psi_1) > \mathbf{R}(\Psi_2)\dots$$

On this model, it’ll end up being true that you ought to  $\Phi$ , as all the specific actions more reasonable than  $\Phi$  entail  $\Phi$ ing. But it’s also true that you have more reason to not- $\Phi$  than  $\Phi$ . That’s a violation of the Connection Principles.

- (1) Suppose  $S$  has to  $\Phi$  is true in  $c$ .
- (2) either  $\mathbf{R}_c(\Phi) \geq b_c$  or  $\mathbf{R}_c(\text{not-}\Phi) \geq b_c$  [1, presupp. of ‘have to’ satisfied]
- (3)  $\forall \Psi \in A [(\mathbf{R}_c(\Psi) \geq b_c) \rightarrow \Psi \text{ entails } \Phi]$  [1, truth conditions for ‘have to’]
- (4) *Assumption:*  $\mathbf{R}_c(\text{not-}\Phi) \geq b_c$
- (5)  $\exists \Psi \in A [\Psi \text{ entails not-}\Phi \ \& \ \mathbf{R}_c(\Psi) \geq b_c]$  [4, RRRW]
- Contradiction!* [3, 5]
- (6)  $\mathbf{R}_c(\Phi) \geq b_c$  [2, 4, disjunctive syllogism]
- (7)  $\forall \Psi \in A [(\mathbf{R}_c(\Psi) \geq \mathbf{R}_c(\Phi)) \rightarrow \Psi \text{ entails } \Phi]$  [3, 6]
- (8)  $S$  ought to  $\Phi$  is true in  $c$  [2, 7, truth conditions for ‘ought’]

## A.2 The reasonable semantics captures the Connection Principles

Here, I show that our reasonable semantics captures the Connection Principles:

- |               |   |
|---------------|---|
| (Good Reason) | If $S$ ought to $\Phi$ , then there is good reason, overall, for $S$ to $\Phi$ .        |
| (More Reason) | If $S$ ought to $\Phi$ , then there is more reason for $S$ to $\Phi$ than not- $\Phi$ . |
| (Most reason) | Necessarily, $S$ ought to $\Phi$ iff $S$ has most reason to $\Phi$ .                    |

I put aside (Good Reason) – mainly due to the tricky context sensitivity of ‘good’ – and assume that if the semantics captures (More Reason) and (Most Reason) it will capture (Good Reason) too.

It is very straightforward to show that the semantics captures (More Reason).

- (1) Suppose *S ought to  $\Phi$*  is true in  $c$ .
- (2)  $\forall \Psi \in A [(\mathbf{R}_c(\Psi) \geq \mathbf{R}_c(\Phi) \rightarrow \Psi \text{ entails } \Phi)]$  [1, truth conditions for ‘ought’]
- (3) *Assumption:  $\mathbf{R}_c(\text{not-}\Phi) \geq \mathbf{R}_c(\Phi)$*
- (4)  $\exists \Psi \in A [\Psi \text{ entails not-}\Phi \ \& \ \mathbf{R}_c(\Psi) \geq \mathbf{R}_c(\Phi)]$  [3, RRRW]
- Contradiction!* [2,4]
- (5)  $\sim[\mathbf{R}_c(\text{not-}\Phi) \geq \mathbf{R}_c(\Phi)]$  [*Reductio*]
- (6)  $\mathbf{R}_c(\Phi) > \mathbf{R}_c(\text{not-}\Phi)$  [5]

Now for (Most Reason). Here’s an assumption: you have most reason to  $\Phi$  iff you have more reason to  $\Phi$  than you have reason to do any of the *alternatives* to  $\Phi$ ing. What are the alternatives to  $\Phi$ ing? One plausible answer – those *specific actions which are not ways of  $\Phi$ ing*:

There is most reason for S to  $\Phi$  iff, for every  $\Psi \in \text{alt}(\Phi)$ ,  $\mathbf{R}(\Phi) > \mathbf{R}(\Psi)$

(Where  $\Psi \in \text{alt}(\Phi)$  iff  $\Psi \in A$  and  $\Psi$  does not entail  $\Phi$ )

With this comparison class in place, we can prove that S ought to  $\Phi$  iff there is most reason for S to  $\Phi$  (at least assuming the presupposition for ‘ought’ is met):

There is most reason for S to  $\Phi$  iff

$\forall \Psi \in \text{alt}(\Phi), \mathbf{R}_c(\Phi) > \mathbf{R}_c(\Psi)$  iff

$\forall \Psi [(\Psi \in A \text{ and } \Psi \text{ doesn't entail } \Phi) \rightarrow \mathbf{R}_c(\Phi) > \mathbf{R}_c(\Psi)]$  iff



$\forall \Psi \in A [\Psi \text{ doesn't entail } \Phi \rightarrow \mathbf{R}_c(\Phi) > \mathbf{R}_c(\Psi)]$  iff

$\forall \Psi \in A [\sim(\mathbf{R}_c(\Phi) > \mathbf{R}_c(\Psi)) \rightarrow \sim(\Psi \text{ doesn't entail } \Phi)]$  iff

$\forall \Psi \in A [\mathbf{R}_c(\Psi) \geq \mathbf{R}_c(\Phi) \rightarrow \Psi \text{ entails } \Phi]$  iff

S ought to  $\Phi$

A nice bonus of the fact that ‘have to’ is not fully monotonic (but merely Strawson-monotonic) is that we get some Connection Principles between ‘have to’ and ‘reason’. For example:

(More Reason\*) If *S has to  $\Phi$*  is true in *c*, then there is more reason for S to  $\Phi$  than there is reason for S to not- $\Phi$ .

(Most Reason\*) If *S has to  $\Phi$*  is true in *c*, then there is most reason for S to  $\Phi$ .

These follow from the fact that ‘S has to  $\Phi$ ’ entails ‘S ought to  $\Phi$ ’, and that ‘S ought to  $\Phi$ ’ entails both ‘there is more reason to  $\Phi$  than not- $\Phi$ ’ and ‘there is most reason to  $\Phi$ ’. What we won’t get is the other direction of ‘Most Reason\*’:

If S has most reason to  $\Phi$ , then S has to  $\Phi$ . (False)

But this seems like exactly the right result. Perhaps there’s most reason for me to keep my seatbelt on at all times on a flight – nonetheless I don’t have to.

## Chapter 2: A Critique of Pure Reasons

Suppose Sally has an excruciating toothache. Here are two things we can say. First: Sally has *a reason* to go to the dentist. Second: Sally has *reason* to go to the dentist.

There, I used the noun ‘reason’ twice. The first time, I used it as a count noun. Count nouns tend to pick out discrete, countable things – think nouns like ‘car’, ‘molecule’ and ‘idea’. A sign that you’re dealing with a count noun is the presence of an indefinite article (as in ‘*a reason*’) or number phrase and plural (as in ‘three reasons’). We’re using ‘reason’ as a count noun when we say familiar things like:

- (1a) The fact that Sally has a toothache is a reason for her to go to the dentist.
- (1b) Sally has at least one (good/strong/weighty) reason to go to the dentist.

The second time, I used ‘reason’ as a mass noun. Mass nouns tend to pick out things that come in amounts – think nouns like ‘lava’, ‘oxygen’ or ‘happiness’. Generally speaking, you can tell you’re dealing with a mass noun if you see a measure phrase (‘more’, ‘lots of’), or if you don’t see an indefinite article or plural marking. We’re using ‘reason’ as a mass noun when we say familiar things like:

- (2a) Sally has (some / lots of / plenty of) reason to go to the dentist.
- (2b) Sally has more reason to go to the dentist than the movie theatre.

You're probably more familiar with using 'reason' as a count noun than as a mass noun. But keep an ear out and you'll hear the mass form all the time, both inside and outside the philosophy classroom. Sometimes you probably slip between using 'reason' as a count noun and a mass noun without realising it.<sup>1</sup>

'Reason' isn't the only noun that has both a count and a mass form. 'Rope' is another dual life noun: there's *rope* (the flexible fibrous material), and there are *ropes* (lengths of rope). So is 'beer': there's *beer* (the tasty alcoholic beverage) and there are *beers* (bottles of beer, or varieties of beer).<sup>2</sup>

Reasons – the things picked out by the count noun – play a central role in ethical theory and epistemology. In fact, many philosophers think that reasons are *the* central normative notion, in terms of which all other normative phenomena can be explained. All of the facts about what's good, what one ought to do, what one is obligated to do, what attitudes are fitting – all of these can be explained by facts about normative reasons and their weights.<sup>3</sup>

---

<sup>1</sup> There's further still broadly normative senses of 'reason' which I'm not looking at in this project. For example, 'reason' can, as a mass noun, pick out a kind of faculty or rational capacity ("Immanuel thought that it was only by exercising reason that we could determine right from wrong"). It can also, as a verb, pick out a kind of rational activity or process ("Tomorrow, Immanuel will get up, sit at his desk, and reason his way through a puzzle until lunch time").

<sup>2</sup> Kiss et al. point to dual life nouns as diverse as *fruitcake*, *forest*, *anticoagulant*, *war*, *eccentricity*, *ribbon*, *mockery*, *silence*, *imperfection*, *curvature*, and *analysis*.

<sup>3</sup> See the 'Reasons First' programme, and in particular Scanlon (1998), Parfit (2011), and Schroeder (2021).

It's surprising, then, that *reason* – the thing picked out by the mass noun – has for the most part received little attention.<sup>4</sup> Why is this so? I suspect that at least part of the explanation has to do with a particular kind of semantic assumption that philosophers have been tacitly making. The assumption is this: all mass noun reason-talk can be cashed out in terms of claims about count noun reasons. It's just been assumed, I think, that when we say "Sally has lots of reason to go to the dentist" what we mean is that she has reasons of a certain strength; or that when we say "She has more reason to go to the dentist than the movie theatre", what we mean is that her reasons favour the former course of action more than the latter. If that were right, then of course there'd be no interesting, distinctive phenomenon here for a philosophical theory of the mass noun to study or explain. By developing a theory of reasons we'd get a theory of *reason* for free.

But this semantic assumption looks to me to be false. In §1, I argue that there are fundamental obstacles to giving an analysis of mass noun reason sentences in terms of reasons. All such 'count fundamentalist' semantics I've been able to think of generate bad predictions when 'reason' is

---

<sup>4</sup> An important exception is Daniel Fogal (2016), to whose rich and careful work on this topic I owe a serious debt. Like me, Fogal thinks that what's picked out by the mass-sense of 'reason' is more fundamental than what's picked out by the count-sense. But there are at least three important differences. Firstly, Fogal does not commit himself to any particular picture of what mass-noun 'reason' picks out, whereas below I sketch a positive proposal in terms of states. Secondly, the relationship between reasons and reason that I sketch in §3, while close in spirit to some of what Fogal says, is strictly speaking incompatible the kind of theory he gives in his 2016 paper and later in Fogal & Risberg (2023a). For Fogal (and Fogal & Risberg), reasons are the sources of quantities of *pro tanto* reason ('normative support'), with distinct reasons being the sources of different amounts of reason. On my picture, we'll see, the mass noun 'reason for S to  $\Phi$ ' is true of a single thing (a state S is in), and S's reasons to  $\Phi$  are the sources of one and the same thing (that state). Thirdly, Fogal's argument for taking the mass-sense of 'reason' to be fundamental is very different from my own. The key piece of data, for Fogal, is that we speak as though reasons (the things picked out by the count noun) generate/provide the stuff picked out by the mass noun (as in "Sally's toothache gave her reason to go to the dentist"). Fogal then deploys a kind of family resemblance argument: each dual life noun that can appear in these 'generational' environments is such that what's denoted by the mass noun is more fundamental than what's denoted by the count noun. I'm wary of grounding the case for mass-fundamentalism in this data. We say things like "his description of the mugger gave his testimony some detail" or "the drum solos give the song some variation", but it's certainly not obvious that *detail* picks out anything over and above details, or *variation* picks out anything over and above variations – it seems to me that something has *detail* in virtue of having many details, or *variation* in virtue of having many variations.

embedded within comparatives. The lesson: when we talk about what there's reason for an agent to do, we're not talking about the reasons there are.

This result opens up a host of new questions. Chief among them: how should we understand talk about what there's reason to do, if not in terms of reasons? In §2, following the treatment that Wellwood (2019) gives to abstract mass nouns (and, in particular, the treatment that she, Fabrizio Cariani and Paolo Santorio give to 'confidence'), I suggest that the mass noun 'reason' is a predicate of *states*. In short, the sentences in (2) are to be understood as quantifying over and measuring the sizes of reason-states – the kind of state you're in when you have reason to do something.

§3 returns to count noun reasons and their metaphysical fundamentality. I argue that a states-based semantics for 'reason' is at best uncomfortable for the Reasons First programme (the idea that count noun reasons are normatively fundamental) and most likely just incompatible with it. If the mass noun 'reason' ought to be understood in terms of states, then the count noun should be too.

## **§1. What 'Reason' Can't Mean**

Take a comparative like the following, which features 'reason' as a mass noun:

(\*) There is more reason for Sally to go to the dentist than there is reason for Tim to go.

We're going to spend this section thinking about what (\*) might mean. Let's introduce a character, the Count Fundamentalist.

**Count Fundamentalism:** The best semantics for (\*) will appeal to reasons for Sally to go to the dentist.

It's not too hard to think of the kind of gloss of (\*) that the Count Fundamentalist will want to give. Standard theories of normative reasons say that every reason has a certain *strength* or *weight*, which represents the degree to which it counts in favour of a particular course of action.<sup>5</sup> The Count Fundamentalist is likely to want to paraphrase (\*) as something like:

(\*)<sub>C</sub>: The weight of Sally's reasons for and against going to the dentist, taken together, exceed the weight of Tim's reasons for and against going to the dentist, taken together.

I'm going to try and convince you that Count Fundamentalism simply cannot be right. The problem isn't that (\*) and the paraphrase (\*)<sub>C</sub> are true under different conditions – I'm not going to argue against Count Fundamentalism by pointing to cases in which (\*) is true but the paraphrase is false, or vice versa. Rather, the trouble for Count Fundamentalism is the task of building up their proposed truth-conditions compositionally.

---

<sup>5</sup> See §§4-5 of Fogal & Risberg (2023b) for a critical survey of various ways philosophers have tried to understand the idea that reasons have weights.

A good semantic theory won't just state truth-conditions for all the sentences in a language one by one. Instead, it needs to assign meanings to the parts of the language which then combine to form the meanings of whole sentences. In order for Count Fundamentalism to be a plausible semantic theory, it needs to tell us what meaning is being assigned to the mass noun that appears in (\*):

[[*reason for Sally to go to the dentist*]] = ???

The meaning they assign to this noun, when combined with the rest of the material in the sentence (the meanings of the expressions 'more', 'there is', 'reason for Tim to go to the dentist', etc.), needs to yield something like the truth-conditions in (\*)<sub>C</sub>. It's that job – the job of getting their truth-conditions to be generated compositionally – which turns out to be very difficult.<sup>6</sup>

In order to show this, I'm going to argue against a series of semantics for 'reason' that the Count Fundamentalist might propose. I hope that what the survey below shows is that the challenges facing a Count Fundamentalist theory are both very serious and quite general – that there are obstacles to the very idea of an analysis of 'reason' in terms of reasons.

---

<sup>6</sup> Finlay (2014, p.91-2) gives a Count Fundamentalist analysis. He concedes that his analysis of mass noun reason sentences is non-compositional, but insists "[these] expressions are evidently idiomatic" (p.92, fn 22), as they "appear to be idiosyncratic to English" (ibid.). Finlay cites Swedish as a language in which 'more/most reason' translate as 'stronger/strongest reasons'. Finlay seems to be right that some languages don't have 'reason' appearing as a mass noun (e.g. French and Italian, according to my informants). But it's certainly not true that mass noun 'reason' is idiosyncratic to English. I've checked with Icelandic, Danish and Dutch speakers, and all those languages can use their translation of 'reason' as a mass noun. It also looks to me like Hebrew does too (though this is harder to test). An informant tells me that even the case of Swedish is complex. 'Reason' can be translated as either 'skäl' or 'anledning', with the former being the more commonly used expression in philosophical contexts. While 'skäl' is more commonly graded in the way Finlay indicates, both nouns can appear as mass nouns. The expressions "mer/mest skäl" (more/most reason) do occur. 'Mer' picks out a non-counting flavour of 'more', used for mass nouns, which is contrasted with 'fler', which measures by counting. (Thanks to Gunnar Björnsson, Jonathan Fiat, Ido Benbaji, Omri Doron, Hedde Zeijlstra, and Oktavía Jóns).

## 1.1 ‘Reason’ picks out a weight?

Let’s get an initial thought out of the way. The Count Fundamentalist thinks that a comparative like (\*) compares the weight of Sally’s reasons to go to the dentist to the weight of Tim’s reasons to go to the dentist. So here’s a natural idea – the mass noun ‘reason for Sally to go to the dentist’ simply picks out the weightiness of her reasons, like so:

[[*reason for Sally to go to the dentist*]] = *d*, where *d* is the weight, on balance, of all of Sally’s reasons to go to the dentist.

But this semantics is no good: it confuses *a measurement* with *the thing being measured*.<sup>7</sup> Take a look at:

(3)        There’s more coffee in Sally’s cup than there is coffee in Tim’s cup.

(3) compares two measurements – the *amount* of coffee in Sally’s cup, with the *amount* of coffee in Tim’s cup. But it would clearly be a mistake to give the noun the following meaning:

[[*coffee in Sally’s cup*]] = *d*, where *d* is the volume of coffee in Sally’s cup.

The coffee in Sally’s cup is a hot tasty liquid, not an amount:

---

<sup>7</sup> See Wellwood (2019). Nouns like ‘coffee’ don’t come with degrees built-in to their extensions. On Wellwood’s semantics, degrees only enter the semantic derivation when the extension of the noun combines with (a morpheme corresponding to) ‘much’. So the more sophisticated way of putting the argument I make above is this: if a noun like ‘reason’ or ‘coffee’ denotes a degree, then that leaves ‘much’ without a job.



(4a) Sally smelled the coffee in her cup.

(4b) # Sally smelled 16oz.

Weights and measures don't have smells. When Sally smells her coffee, it might be true that she smells 16oz *of coffee*, but it's certainly not true to say she smelled 16oz.

The same goes for 'reason'. In (\*):

(\*) There is more reason for Sally to go to the dentist than there is reason for Tim to go.

we're comparing two amounts – the amount of reason Sally has to go to the dentist, with the amount of reason Tim has to go to the dentist. But Sally's reason to go to the dentist isn't *itself* an amount or degree (e.g. 47). When I say "There is reason for Sally to go to the dentist" or "Sally has reason to go to the dentist", I am clearly not saying "There is *d*" or "Sally has *d*" (where *d* is some degree).

'Reason for Sally to go to the dentist' picks out something that can be measured to yield a number, or a degree. But 'reason for Sally to go to the dentist' doesn't itself pick out a number or a degree.

## 1.2 ‘Reason’ picks out reasons?

Next try. The Count Fundamentalist should want ‘reason’ to pick out something that’s measured to yield the weightiness of Sally’s reasons to go to the dentist. So why not just take ‘reason for Sally to go to the dentist’ to pick out all of Sally’s reasons to go?

According to this suggestion, unlike the mass nouns ‘lava’, ‘water’, ‘iron’, etc., the mass noun ‘reason’ doesn’t pick out *stuff*. Rather, it picks out discrete, countable *things* – reasons.

Fortunately for the Count Fundamentalist, there are many examples of mass nouns which have just this kind of denotation. Take the noun ‘furniture’, for example. ‘Furniture’ is a mass noun:

(5a) # My sofa is a furniture.

(5b) The living room has plenty of furniture, but the office doesn’t have enough.

‘Furniture’ doesn’t pick out a furniture-y substance or stuff. Instead, it picks out particular tables and chairs, and collections of tables and chairs:<sup>8</sup>

(6a) That is furniture. (*Pointing at a desk*)

(6d) The only furniture I own is a desk and two chairs.

---

<sup>8</sup> What is a ‘collection’ of items of furniture? Some model the idea of ‘collecting’ individuals (items of furniture, or portions of stuff) together using sets (e.g. Schwarzschild 1992); others use some sort of ‘join’ or ‘sum’ operation (e.g. Link 1983). The details do not matter too much for our purposes.

Other examples of nouns like ‘furniture’ include *ammunition*, *cutlery*, *mail*, and *baggage*, which are true of bullets and shells, knives and forks, letters and parcels, and bags and suitcases.<sup>9</sup> Linguists call these ‘Object Mass Nouns’.

The Count Fundamentalist suggestion on the table sees ‘reason’ as a kind of Object Mass Noun:

[[*reason for Sally to go to the dentist*]] = a predicate that’s true of each of the reasons for Sally to go to the dentist, plus collections of reasons for Sally to go to the dentist.

Just like how “furniture in Sally’s house” picks out nothing over and above the items of furniture in Sally’s house, “Sally’s reason to go to the dentist” picks out nothing over and above her reasons to go. It’s a pleasingly elegant account.

But it’s hard to see how it could be right.

Let’s start with a small but significant point. If this particular count-first semantics were correct, then ‘reason’ would be a dual life noun (a noun with a mass and count form) whose mass form is an Object Mass Noun (picks out individuals). But this would buck a striking trend: as you survey various Object Mass Nouns, you’ll see that none of them are dual life nouns. None of *furniture*,

---

<sup>9</sup> Object Mass Nouns have a long and interesting history. See Chierchia (1998), Gillon (1992) and Bale & Barner (2009) for three examples of people who pay particularly close attention to them. One way of modelling the difference between mass nouns like ‘furniture’ and substance mass nouns like ‘water’ (roughly following the way Link 1983 distinguishes between plurals and mass nouns) is to give them both ‘join semi-lattice’ extensions, but have the former and not the latter bottom-out in minimal parts. While ‘furniture’ is true of particular tables and chairs and all the ways of collecting these individuals together, ‘water’ is true of portions of water, parts of these portions, parts of parts of these portions..., and all the ways of collecting these portions together.

*cutlery, mail*, etc. have a count-noun form – in each case you need to say *piece* of furniture, cutlery, mail, and so on. Some theories of the mass/count distinction take it as a datapoint that this holds universally: Bale & Barner (2009, p.227), for example, claim there is *no* dual life noun whose mass form picks out individuals. If the Count Fundamentalist semantics on the table were right, then ‘reason’ would seemingly be acting like no other dual life noun around.<sup>10</sup>

We should already be suspicious of the Count Fundamentalist’s suggestion, then. But things get worse. Here’s a robust generalisation: when Object Mass Nouns are embedded in comparatives, one can always interpret the comparative as making a claim about *number*.<sup>11</sup> Take something like:

(7) Last year, I received more mail than my brother.

There is clearly a reading of (7) according to which it’s true iff the *number* of pieces of mail I received was greater than the *number* of pieces of mail my brother received. (7) clearly has a true reading in a case where, say, I received three letters last year and my brother received a single 10kg package.<sup>12</sup> Or take:

(8) There is more furniture in the study than the living room.

---

<sup>10</sup> Yoad Winter (p.c.) thinks that he may have found a counterexample in ‘fowl’ (as in “that bird is a fowl!”; “we caught plenty of fowl!”). That may be right, in which case the above paragraph needs to be weakened – it’s *almost always the case* that Object Mass Nouns don’t have count forms. That’s still something of a mark against the Count Fundamentalist.

<sup>11</sup> See Barner & Snedeker (2005), Bale & Barner (2009).

<sup>12</sup> Perhaps it *only* has a true reading in such a case. But all we need for the argument in this section to hold is that it has *a* reading on which we measure by counting.

Suppose the office contains three light, small plastic chairs, and the sitting room contains one heavy, bulky sofa. Despite the difference in size/weight of the objects concerned, (8) still has a salient true reading. Contrast (9):

- (9) There is more reason for Tim to go to the dentist than there is reason for Sally to go.

We simply cannot evaluate (9) by counting Sally or Tim's reasons. Suppose I tell you a story:

Tim has three reasons to go to the dentist – he wants to ask about a new whitening treatment, he wants to get his teeth professionally cleaned, and he would enjoy the drive. Sally, by contrast, has one reason to go to the dentist – she has an excruciating toothache.

If 'reason for Sally to go to the dentist' were an Object Mass Noun which simply picked out Sally's reasons to go to the dentist, then (9) should have a true reading in the case described. After all, when Object Mass Nouns are embedded in comparatives you measure by counting, and Tim does indeed have more *reasons* to go to the dentist than Sally does. But clearly (9) has no true reading in this case. On the contrary, Sally has more reason to go to the dentist than Tim does (she has an excruciating toothache, whereas none of Tim's reasons are particularly weighty).

### 1.3 ‘Reason’ picks out reasons and weights?

The meaning of ‘reason for Sally to go to the dentist’ considered in the previous section did not contain any information about the weights of the various considerations counting in favour of Sally going to the dentist. What happens if we add weights back in?

There’s a few ways we might go about doing so. Here’s one way:

[[*reason for Sally to go to the dentist*]] = a predicate that’s true of pairs  $\langle p, d \rangle$  (where  $p$  is a reason for Sally to go to the dentist, and  $d$  is the degree to which  $p$  counts in favour of Sally’s going to the dentist) and collections of these pairs.

Now ‘reason for Sally to go to the dentist’ doesn’t pick out her reasons (a list of facts), it picks out her reasons complete with information about the weightiness of each of those reasons.

This might sound promising, but in fact we haven’t really made any progress. This semantics *still* sees ‘reason’ as an Object Mass Noun. It’s still like ‘furniture’ or ‘mail’ insofar as it’s a predicate of discrete, countable things – namely the ordered pairs themselves. Just as ‘furniture in my office’ picks out, say, three plastic chairs (and various ways of collecting them together), this approach sees ‘reason for Tim to go to the dentist’ as picking out the following three ordered pairs (along with the various ways of collecting the pairs together):

<Tim wants to ask about a new whitening treatment,  $d_1$ >

<Tim wants to get his teeth professionally cleaned,  $d_2$ >

<Tim would enjoy the drive,  $d_3$ >

But now we run into exactly the same problem as we did in the previous section. Comparatives with Object Mass Nouns can be evaluated by counting. So, if this semantics is right, when ‘reason for Tim to go to the dentist’ is embedded in a comparative like (\*), the comparative should have an interpretation which goes by counting reasons/weight pairs. But that’s clearly a bad prediction: one cannot hear reason-comparatives like (\*) as claims about number.

It’s the very *shape* of the Count Fundamentalist’s denotation for ‘reason’ that’s wrong. You cannot fix the problem simply by adding-in information about the weight of the agent’s reasons.

#### **1.4 Objection: agents always have infinitely many reasons**

Let’s take a second to respond to an objection the Count Fundamentalist might make. What rules out the counting interpretation in reason-comparatives like (\*), they might say, is the fact that whenever an agent has one normative reason to do something, they have infinitely many. It’s always impossible to count up an agent’s reasons, and that’s why only the weighing interpretation of the comparative is available.

It’s not obvious that agents do always have infinitely many reasons to act, but let’s concede this idea to the Count Fundamentalist for the sake of argument – I don’t think it will really help them.

Firstly, it seems to me that comparatives involving Object Mass Nouns still have counting-interpretations available when the domain is infinitely big. Suppose Bob has infinitely many large sofas in his living room, and Ann has infinitely many small chairs in hers. Consider the comparative:

(10) There is more furniture in Bob's living room than there is furniture in Ann's.

My first response is to think (10) is false – they have the same amount of furniture, as they each have infinitely many pieces. Getting a true reading of (10) seems to require some sort of coercion on my part. Contrast:

(\*) There is more reason for Sally to go to the dentist than there is reason for Tim to go.

(Recall: Sally's reasons to go to the dentist are stronger than Tim's). If 'reason for Sally to go to the dentist' were a predicate of infinitely many things, like 'furniture in Bob's living room', then we'd expect (\*) to also have two readings: an easily accessible false reading, and a true reading that requires some kind of coercion to access. But that's not what we find: (\*) *only* has a true reading, and it doesn't seem to take any coercion in order to glom onto it.

Here's an even more important thing to notice. It's clear that somehow or other we do often find ways of winnowing down a large set of reasons to something countable. It's totally commonplace



for people to number the considerations that count in favour of something (“there’s one reason for you to accept the offer, and two reasons for you not to”). And in these contexts – contexts in which it’s *not* true to say agents have infinitely many reasons – we still don’t measure reason-to-act by counting:

- (11) You have one strong reason to accept the offer but only two weak reasons to reject it, so (it seems) you have more reason to accept the offer than reject it.

Overall, it seems that the Count Fundamentalist cannot save their semantics by appealing to infinities.

### **1.5 A different Count Fundamentalist account?**

‘Reason for Sally to dance’ doesn’t itself denote a degree, or a weight; it picks out some  $x$  which can be measured to yield a degree, or a weight. But if we try and build that  $x$  out of reasons, we’re going to start generating unwanted counting interpretations of comparatives. A collection of reasons, or a collection of pairs of reasons and weights, or a collection of  $n$ -tuples of reasons/weights/agents/actions, or a collection of *whatever* – all of these collections can be counted

up. But, try as you might, you can't count up Sally's reason to act. So Sally's reason to act isn't identical to any of these things.<sup>13</sup>

There may be some more sophisticated way of getting a count-first semantics to work, on which 'reason for Sally to go to the dentist' neither picks out a degree nor a collection of things. I haven't thought of any. I think we've been shown enough to justify looking in a different direction.

## §2 What 'Reason' Might Mean

### 2.1 'Reason' as a predicate of states

The mass noun 'reason' doesn't pick out reasons. So what does it pick out? What do we mean when we say that someone has reason to do something?

So far as I can see, there are two directions we can go in at this stage. One is to model the mass noun 'reason' a bit like substance mass nouns (lava, iron, coffee), and have it be a predicate of

---

<sup>13</sup> *Objection:* The argument must go wrong somewhere. The denotation of 'iron' eventually bottoms out in minimal parts – there are iron-atoms, none of whose parts are themselves iron. The denotation of 'iron' is, therefore, a collection. But 'iron' comparatives don't go by number. Why can't we say, similarly, that 'reason' bottoms out in reasons, despite the fact that reason-comparatives don't go by number? *Response:* It's a hotly contested issue whether the denotation of 'iron' really does bottom-out in minimal parts. Chierchia (1998) says it does; Link (1983), Wellwood (2019), and many others say it doesn't. Regardless, those (like Chierchia) who believe the extension does bottom-out in minimal parts think that the question of what the minimal parts are is a vague matter – try as you might, you *can't* count up the smallest things that still fall under the extension of 'iron'. It's this indeterminacy at the minute level which ensures that iron-comparatives then don't go by counting. So while one could, if one wanted, model 'reason' along the lines that Chierchia suggests, this suggestion doesn't seem to be particularly friendly to the Count Fundamentalist: the minimal parts in the denotation wouldn't be *reasons*, which are clearly things we find ways of counting (as in "there are two reasons for you to accept the offer").

abstract reasony stuff.<sup>14</sup> When Sally has more reason to go to the dentist than Tim, that's because she has more of a special kind of normative gunk than Tim does.<sup>15</sup>

I have more sympathy for this way of going than you might have expected, but it's not the approach I take here. This isn't just because I find the idea of reason-gunk somewhat unnatural. The bigger issue is that we don't use the mass noun 'reason' anything like 'iron' or 'coffee'. Substance mass nouns like these display a kind of divisibility. The coffee in my cup is made up of coffee-parts: there's the coffee in the top half of the cup, the coffee in the bottom eighth, and so on.<sup>16</sup> This divisibility is reflected in the semantics for the noun – 'coffee' is standardly taken to pick out portions of mereologically-structured stuff, whose parts are all coffee.<sup>17</sup> But the idea that *reason for Sally to go to the dentist* picks out something imbued with this kind of mereological structure is really quite odd. We certainly don't use the mass noun 'reason' as if it admitted of parts and wholes:

---

<sup>14</sup> The way I'm imagining modelling 'reason' like a concrete mass noun, here, involves giving it a join semilattice denotation which either doesn't bottom-out in minimal parts, or bottoms-out in minimal parts which we can't count up. (See Link 1983, Chierchia 1998).

<sup>15</sup> Some (e.g. Weaver & Scharp) attribute something like this 'reason picks out stuff' picture to Fogal (2016). But Fogal is quite clear: "the term 'stuff' can be misleading, and is generally far less apt for so-called "abstract" mass nouns ('information', 'freedom', 'reason', 'advice', ...) than it is for more "concrete", substance-denoting mass nouns..." (Fogal, 2016, f.n.5). Nonetheless, as I read that paper, Fogal does seem to think that reason comes in quantities or amounts, and that stronger/weaker reasons are stronger/weaker in virtue of the different amounts of reason they provide. I'm not sure how much of this amount/quantity talk Fogal would be able to capture without falling back on something like the stuff picture.

<sup>16</sup> Contrast count nouns: if  $x$  is a cat, the bottom half of  $x$  isn't a cat. Contrast Object Mass Nouns too: if  $x$  is furniture, the bottom half of  $x$  needn't be furniture. There do seem to be limits on the divisibility of substance mass nouns. My coffee contains molecules of H<sub>2</sub>O. H<sub>2</sub>O molecules are not coffee. See Wellwood 2019 (p.43; p.188-9) for relevant discussion.

<sup>17</sup> See, e.g., Link 1983, Wellwood 2019.

- (12a) There is some coffee in my cup, and every part of it is coffee.
- (12b) ? There is some reason for Sally to go to the dentist, and every part of it is reason for Sally to go to the dentist.

None of this is insurmountable for an account which sees ‘reason’ as picking out a kind of stuff, of course. But it does make me think we should look elsewhere.

Thankfully, there’s another way forward. I’m going to sketch an analysis of ‘reason’ in terms of something which many philosophers and linguists are already committed to, and which semanticists have already been using to analyse abstract mass nouns. The method is inspired by Wellwood (2019) and a later paper from Cariani, Santorio and Wellwood (m.s.), who focus on the abstract mass noun ‘confidence’. On this way of going, ‘reason’ doesn’t pick out portions of stuff, it picks out ways agents that agents can *be* – ‘reason’ is a predicate of states.<sup>18</sup>

It’s best to introduce the idea of states by way of events. Alongside the usual things metaphysicians take to populate the world (e.g. properties and particulars) most philosophers are now happy to

---

<sup>18</sup> Here and through the rest of this chapter I speak as if it’s only agents who can be in these states. But nothing I say here rules out non-agents from being in these states too. (In fact, in the next chapter, I put forward a positive case for reasons for non-agents).

admit events in addition: bitings, ringings, stabbings, and singings. For the semanticist, the appeal of events is that they help give an elegant semantics for a sentence like the following:<sup>19</sup>

(13) Johnny mixed up the medicine.

According to Davidson and those following him, the truth-conditions of (13) can be given like so:

[[*Johnny mixed up the medicine*]] = TRUE iff

$\exists e$  [**mixing**( $e$ ) & **agent**( $e$ , Johnny) & **theme**( $e$ , the medicine)]<sup>20</sup>

‘**Mixing**( $e$ )’ is to be read “ $e$  is a mixing event”. The other bold-faced words represent ‘thematic roles’ – the idea being that events have a kind of internal structure which is tracked by the grammar. The expression ‘**agent**( $e$ , Johnny)’ is to be read “Johnny is the unique agent of  $e$ ”, and so on.

Neo-Davidsonians want to push this framework even further: quantification over event-like things shows up almost everywhere in natural language semantics.<sup>21</sup> For our purposes, the most important

---

<sup>19</sup> Much of the appeal of event semantics derives from its treatment of adverbial modification, and its explanation of the validity of ‘diamond’ inferences – that is, its explanation of why (13\*) entails (13) but not vice versa:

(13\*) Johnny slowly and carefully mixed up the medicine in the basement for three hours.

According to the Davidsonian, (13\*) is true iff  $\exists e$  (**mixing**( $e$ ) & **slow**( $e$ ) & **careful**( $e$ ) & **agent**( $e$ , Johnny) & **theme**( $e$ , the medicine) & **location**( $e$ , the basement) & **duration**( $e$ , three hours)). (13) follows from (13\*) by conjunction elimination, but (13\*) does not follow from (13), due to the invalidity of conjunction introduction.

<sup>20</sup> Davidson himself would have presented the semantics differently – he would have had argument places for the agent and patient – but these differences are not important for us.

<sup>21</sup> The extension from events to states is admittedly controversial. Parsons (1990, p.186) himself introduces states somewhat reluctantly. Maienborne (2007, 2011) argues we ought to treat ‘state verbs’ (*sit, stand, lie*) differently from stative verbs (*know, weigh*) and copular constructions. The former denote Davidsonian states, the latter denote what she calls ‘Kimian’ states – temporally bound property exemplifications.

things added to our repertoire by the neo-Davidsonian are *states*.<sup>22</sup> States are typically used to analyse stative verbs (e.g. *believe, recognise, own*)<sup>23</sup> and adjectives (e.g. *clever, blue*). Take (14):

(14) Bob is anxious about nuclear war.

A standard neo-Davidsonian analysis of the adjective ‘anxious’ would take it to be a predicate of anxiety-states, like so:

$$\begin{aligned} [[\textit{anxious}]] &= \lambda s. \mathbf{anxiety}(s) \\ &\approx \{s \mid \mathbf{anxiety}(s)\} \end{aligned}$$

The object of the anxiety is called the *theme* of the state:

$$\begin{aligned} [[\textit{anxious about nuclear war}]] &= \lambda s. \mathbf{anxiety}(s) \ \& \ \mathbf{theme}(s, \textit{nuclear war}) \\ &\approx \{s \mid \mathbf{anxiety}(s) \ \& \ \mathbf{theme}(s, \textit{nuclear war})\} \end{aligned}$$

And an anxiety-ascription like (14) amounts to saying that the subject is the holder of the relevant state:<sup>24</sup>

---

<sup>22</sup> I don’t want to take too much of a stand here on the metaphysics of states. All Cariani et al. require in the case of ‘confidence’, for example, is that the relevant entities have thematic participants and can be ordered. I can’t see why things wouldn’t work just as well with the semantics quantifying over, say, property exemplifications/instantiations (Kim 1993) or perhaps tropes (Moltmann 2013).

<sup>23</sup> For details on the stative/non-stative distinction, see Skow (2018). One rough but handy test (from Dowty p.55) for the difference is this: a verb phrase V is non-stative iff “one thing S did was V” is grammatical, and stative otherwise.

<sup>24</sup> Thematic roles present at logical form ‘**holder**’, ‘**theme**’ etc. bundle the sentence’s constituents into one long predicate of states –  $\lambda s. (\mathbf{anxiety}(s) \ \& \ \mathbf{holder}(s, \textit{Bob}) \ \& \ \mathbf{theme}(s, \textit{nuclear war}))$ , which is then existentially quantified over. For most sentences, the existential quantification is there at the level of logical form but unpronounced.

[[*Bob is anxious about nuclear war*]] = TRUE iff

$\exists s$  [**anxiety**(*s*) & **holder**(*s*, Bob) & **theme**(*s*, nuclear war)]

That’s an analysis of the adjective ‘anxious’. But we might also wonder about the mass noun–*anxiety*. What should we say about that? One could, if one wanted, take ‘anxiety’ to be a predicate of anxiety-stuff – some kind of abstract gunk that an agent possesses in some quantity when they’re anxious. But why would we? Now we’ve introduced states to our ontology we can treat the noun ‘anxiety’ just like we treated the adjective:

[[*anxiety*]] = [[*anxious*]] =  $\lambda s$ . **anxiety**(*s*)

And we can do the same for other nouns too – both for those that have adjectival forms (e.g. nouns like *wealth*, *heat*, *confidence*) and those nouns that bear less close relations to adjectives or verbs (e.g. nouns like *pleasure*, *despair*, *swagger*).<sup>25</sup> These nouns don’t pick out heat-stuff, or confidence-stuff, pleasure-stuff or swagger-stuff – they pick out states.<sup>26</sup>

I think we should give a similar states-based treatment to the mass noun ‘reason’. The noun is a predicate reason states:

[[*reason*<sub>+mass</sub>]] =  $\lambda s$ . **reason**(*s*)

---

<sup>25</sup> Cariani et al. give a states-based semantics for confidence/confident. They also mention ‘swagger’ as a noun that should receive a states-based analysis.

<sup>26</sup> Roger Schwarzschild (2022) has recently argued that *all nouns* are predicates of states.

$$\approx \{s \mid \text{reason}(s)\}^{27}$$

This semantics can be rendered compositional in standard neo-Davidsonian ways. When Sally has some amount of reason to go to the dentist, she is the holder of a reason-state whose theme is *that she goes to the dentist*.<sup>28</sup>

[[*There is reason for Sally to go to the dentist*]] = TRUE iff

$\exists s$  [reason(*s*) & holder(*s*, Sally) & theme(*s*, *that Sally goes to the dentist*)]

Giving a states-based theory of ‘reason’ is a way of taking the mass noun seriously – it avoids the Count Fundamentalist idea that ‘reason for Sally to go to the dentist’ picks out a degree, or a mere collection.<sup>29</sup> But the nice thing about this semantics is that it does so without positing quantities of

---

<sup>27</sup> Here I’m going to assume that ‘reason’ is a bit like the nouns ‘swagger’, ‘pleasure’ or ‘despair’, insofar as it’s a state-denoting noun which isn’t merely the nominalisation of some adjective. But we may nevertheless ultimately want to tie its semantics to a close-by adjective, and say that just as ‘anxious’ stands to *anxiety* and ‘wealth’ stands to *wealthy*, ‘reason’ bears a similarly close relationship to *reasonable*. It doesn’t seem too off the mark to say that one has reason to do something just when it’s to-some-degree reasonable for them to do it. But there are subtleties here. The mass noun ‘reason’ can take an objective/fact-relative sense which ‘reasonable’ seemingly cannot. If Sally has a painless cavity she’s unaware of, then she has reason to go to the dentist. But it needn’t be reasonable (not even to *some degree* reasonable) for her to go. Secondly, there seems to be a difference in the kind of scales that ‘reason’ and ‘reasonable’ each exploit. ‘Reasonable’ is – like *full*, *clean* – plausibly an *Absolute Gradable Adjective* (borrowing the terminology of Kennedy 2007): there’s a limit to how reasonable it could be for an agent to do something. After all, we do speak of an agent’s  $\Phi$ ing can be *perfectly* or *fully* reasonable; and it is a little marked to say it’s reasonable for  $S_1$  to  $\Phi$ , reasonable for  $S_2$  to  $\Phi$ , but more reasonable for the former to  $\Phi$  than for the latter to  $\Phi$ . By contrast, presumably there’s no limit to how much reason an agent can have to do something. Thanks to Jack Spencer for bringing this second point to my attention. I leave an analysis of ‘reasonable’ for another time.

<sup>28</sup> We could also take the theme of reason-states to be actions, rather than propositions – the semantics is compatible with either way of thinking about the objects of reason claims.

<sup>29</sup> Another (defeasible) way of arguing for a states-based analyses of a noun is to see whether it refers to something that can last through time (“my anxiety lasted the length of the exam”, “my confidence that it would rain only lasted as long as I could see the clouds”). And indeed, to my ears at least, we can get some of these going in the case of ‘reason’: e.g. “Sally’s reason to go to the dentist lasted for as long as she had toothache”, or “Tom’s reason to do the Macarena only lasted for as long as the song was playing”. One might fear that “Sally’s reason” and “Tom’s reason” are count nouns in these sentences, but we can seemingly force the relevant mass reading: “The slanderous newspaper articles and blog posts gave her plenty of reason not to show her face in public, and she knew that this reason she had not to show her face would last until her name was cleared”.



reason *stuff*. Instead, the mass noun ‘reason’ is a predicate of states, which many semanticists and philosophers were already committed to.<sup>30</sup>

The semantics for the comparative is a little more complicated, but for the sake of space I’ll skip the technicalities here. Most of the details can be borrowed directly from Wellwood and Cariani et al. Context serves to supply a *measure function* (represented in the logical form by a variable  $\mu_c$ ) which maps entities (e.g. pluralities, events, states) to degrees. A comparative like (\*) is true (in context  $c$ ) iff  $\mu_c$  maps Sally’s reason-state to a larger degree than Tim’s reason-state:

[[There is more reason for Sally to go to the dentist than for Tim to go]] = TRUE iff  
 $\exists s$  (**reason**( $s$ ) & **holder**( $s$ , Sally) & **theme**( $s$ , *that Sally goes to the dentist*) &  $\mu_c(s) > d$ )<sup>31</sup>

Some constraints need to be posited in order to rein-in the range of admissible values for  $\mu_c$  (i.e. the admissible ways of mapping states to degrees). Cariani et al. do so by imposing orderings on states: confidence/height/heat states can be ranked by *how much confidence/height/heat they are*. Then we just need to impose the constraint that any admissible  $\mu$  will respect this background ordering: whenever  $s_1$  has a smaller size than  $s_2$ ,  $\mu(s_1) < \mu(s_2)$ . Something like that strategy should work fine for ‘reason’ too (modulo a complication discussed below). Semantics enthusiasts should consult Wellwood (2019) and Cariani et al. for further detail.

---

<sup>30</sup> Unlike the stuff-analysis, the states-analysis needn’t give reason-to-act any kind of mereological structure. But that’s not to say that we couldn’t give the states such a structure if we wanted to. Pasternak (2019) does so in order to analyse comparatives involving nouns that pick out attitudes (‘hatred’, ‘desire’).

<sup>31</sup> Here,  $d$  is the extension of the than-clause.

## 2.2 An outstanding puzzle

I close this section with discussion of an interesting outstanding puzzle about reason-comparatives (a puzzle, to be clear, for any semantics of the mass noun ‘reason’ – not just a puzzle for the states-based analysis). Take a look at the following:

*Wealth.* *A* has \$2000 in cash, but he also has \$1500 in debt. *B* has \$1000 in cash and no debt. *Who has more wealth?*<sup>32</sup>

*Help.* *A* and *B* are each finishing their dissertations. *A*’s little sister gave *A* extensive written comments. But she also occasionally logged-in to *A*’s computer to delete huge chunks of text. *B*’s little brother brought *B* coffee a couple of times. *Who had more help finishing their dissertation?*

*Support.* Two candidates are running for president. The first candidate, *A*, is adored by many, but loathed by even more. (She’s loathed slightly more than she’s loved). Most people don’t have an opinion either way on candidate *B*. The few people who do have an opinion on *B* like him a little bit. *Which candidate has more support?*

*Evidence.* *A* has three state-of-the-art, highly reliable thermometers. One of her thermometers says the temperature is 20 degrees, and the other two say it’s 15 degrees. *B*

---

<sup>32</sup> Thanks to Abigail Thwaites and Justin Khoo for both suggesting the analogy with wealth.

has one cheap, somewhat unreliable thermometer, which says the temperature is 20 degrees. *Who has more evidence that the temperature is 20 degrees?*

In each case, there are two equally good answers to the question posed. First answer – *A* has more, because they have the good-stuff in greater quantities than *B* does. Call that the ‘gross’ reading.<sup>33</sup> Second answer – *B* has more, because once you take the bad-stuff *A* has into account, *B* comes out on top. Call that the ‘net’ reading. In each case, the comparative “*A* has more such-and-such than *B* does” has both a true and a false reading.

*Reason-comparatives* display these two readings, just like comparatives involving *wealth*, *help*, *support*, and *evidence*:

*Reason.* Sally has an excruciating toothache. But she’s also injured her hand, and needs to go to the emergency room on the other side of town. Rob has no problems with his teeth, but is due a check-up. *Who has more reason to go to the dentist?*

There are clearly two equally good answers to this question. A gross answer: Sally has more reason to go, because she has an excruciating toothache whereas Rob is merely due a check-up. A net answer: Rob has more reason to go, as he should go to the dentist whereas Sally shouldn’t. The comparative “Sally has more reason to go to the dentist than Bob does” has both a true and a false reading.

---

<sup>33</sup> One could call this a ‘*pro tanto*’ reading, but it seems odd to speak of *pro tanto* help, or *pro tanto* wealth.

A full theory of reason-comparatives will eventually need to explain where gross vs. net readings come from. Two approaches spring to mind, though I don't develop either of them in detail here.

One approach would be to posit shifts in the meaning of the mass noun involved, by appealing to some sort of polysemy. Perhaps 'reason' is ambiguous between  $\text{reason}_{\text{gross}}$  and  $\text{reason}_{\text{net}}$ , 'support' is ambiguous between  $\text{support}_{\text{gross}}$  and  $\text{support}_{\text{net}}$ , and so on.<sup>34</sup> There wouldn't be any need to give radically different semantic entries for the two senses of these nouns. We could continue to say that 'reason', for example, is a predicate of states – we'd just need to distinguish between two kinds of states that an agent can be in (a gross reason state and a net reason state). Sally's *gross* reason state is greater than Rob's, but Rob's *net* reason state is greater than Sally's. Perhaps too there are certain systematic things we can say about how the sizes of the net reason states relate to the sizes of the gross ones (e.g. the size of an agent's net reason state targeting  $\Phi$ ing is equal to the difference between the size of the gross-state targeting  $\Phi$ ing and the size of the gross-state targeting  $\sim\Phi$ ing).

An alternative would say that the different readings are explained by shifts in the *way* we're measuring, rather than shifts in *what* we're measuring. We can measure a wealth-state directly or we can measure it by comparing it to the size of a debt-state. Similarly, we can measure a reason-state targeting  $\Phi$ ing directly (giving us its gross size), or we can measure it by comparing it to the gross size of the reason-state targeting  $\sim\Phi$ ing (giving us its net size). In some ways this is more

---

<sup>34</sup> Broome (2013) posits a similar ambiguity between *pro tanto* reasons and *pro toto* reasons.

attractive than positing ambiguities in ‘reason’, ‘wealth’, ‘help’, ‘support’, and ‘evidence’ but its formal implementation turns out to be somewhat fiddly.<sup>35</sup>

I suspect that there’s lots of interesting further work to be done on this question of how and when ‘net’ vs ‘gross’ readings of comparatives arise. But I put the question aside for now. It’s time to move on to the philosophical consequences of the semantics.

### §3 Why This Matters

When I talk about Sally’s reason to go to the dentist, I’m not talking about the reasons there are for her to go. When I say she has more reason to go to the dentist than Tim, I’m not making a claim about the strength of the considerations that count in favour of her various options. Instead, I’m measuring a particular kind of state that Sally is in.

Now, for the final section of the chapter, I want to return to reasons. Assuming that something like the above states-based semantics is right, what does it tell us about reasons and their fundamentality?

---

<sup>35</sup> Recall that Cariani et al. posit a background ordering on the domain of states, and rein in the values that the measure function  $\mu_c$  can take by requiring that whenever  $s_1 < s_2$  (according to the background ordering),  $\mu_c(s_1) < \mu_c(s_2)$ . If we want to keep this constraint (and we do need some constraint on the values of  $\mu$ ) then in order to allow that  $\mu$  sometimes maps reason-states to degrees representing gross-size and other times maps reason-states to degrees representing net-size, we’d need to posit some sort of contextual flexibility in these background orderings. I think this idea could be made to work (we would have need to posit some flexibility in the ordering to accommodate other context-sensitive dimensions of ‘reason’ anyway), but the details would need to be worked out carefully.

For all I've said so far, this may all look like a kind of semantic book-keeping – a useful exercise to work through, perhaps, but an exercise without any deep philosophical consequences. After all, these are distinct claims:

*Count Fundamentalism*: The best semantics for sentences involving the mass noun 'reason' will understand them in terms of reasons.

*Reasons First*: Facts about normative reasons explain all other normative facts.<sup>36</sup>

Count Fundamentalism is a semantic claim – it's a claim about the meaning of sentences featuring a particular normative expression. But Reasons First is a metaphysical claim – it's a claim about the relative fundamentality of two kinds of normative fact. And notice that these two claims appear to be perfectly consistent:

- (i) The best semantics for sentence S is not given in terms of Xs.
- (ii) If S is true, then, necessarily, its truth is explained by facts about the Xs.

Alan is popular. His popularity is explained by, or grounded in, various facts (e.g. the fact that his friends Bob, Carl, Dave... and Ziad each like him). Perhaps it's even necessary that, if Alan is popular, his popularity is explained by facts about his friends. But on the face of it at least, it doesn't follow from this metaphysical claim that the best *semantics* of the sentence "Alan is

---

<sup>36</sup> Note: One recent self-proclaimed Reasons Firster, Mark Schroeder, wants to understand Reasons First in such a way that it's compatible with the mass sense of 'reason' being normatively fundamental (2021, p.40-1).

popular” will invoke friends or the liking relation. (Perhaps instead the best semantics will say that the sentence is true iff the degree to which Alan is popular exceeds some contextually salient standard,<sup>37</sup> or iff he’s the holder of some popularity-state).

One might have expected that the Reasons Firster could resist a slide from semantics to metaphysics in just the same way. Couldn’t they just accept the semantic claims I made in the previous two sections (i.e. concede that the meaning of ‘Sally has reason to go to the dentist’ doesn’t invoke reasons) but insist that Sally’s being in a reason state targeting going to the dentist is *explained* by facts about her reasons to go to the dentist and what they favour, just as the fact that Alan is popular is explained by facts about his friends and their liking of him?

I think this manoeuvre is surprisingly difficult for the Reasons Firster to pull-off. In this particular case, the path from semantics to metaphysics is shorter than you might expect.

Here’s why. With dual life nouns, one normally expects that either the meaning of the count form can be given in terms of the mass form, or that the meaning of the mass form can be given in terms of the count form. Normally, the reduction runs from count to mass. What is *a beer*? It’s a portion (or kind) of beer. What is *a medicine*? It’s a kind of medicine. What is *a silence*? It’s a period of silence. What is *a hair*? It’s a length of hair. Sometimes, perhaps, the mass form is derived from the count form. Plausibly, the mass noun ‘chicken’ is a predicate of matter – perhaps specifically edible matter – from *a chicken*. The important point is that whether the semantic analysis goes one

---

<sup>37</sup> C.f. Kennedy 2007.

way or the other (from count to mass or from mass to count), with dual life nouns we always expect to find such a reduction.<sup>38</sup>

The lesson of §1 was that the meaning of the mass noun ‘reason’ can’t be given in terms of the meaning of the count noun. So, with the above generalisation in view, we should expect to find an analysis which runs the other way – we’d expect to be able to cash out the meaning of the count noun ‘reason’ in terms of reason states too.

What might an analysis of reasons in terms of reason look like? One can get quite far here by looking to an insight from Daniel Fogal (2016) and his later joint work with Olle Risberg (2023a). Fogal notices that for some dual life nouns (e.g. *lights, pleasure, sorrow*), the noun’s count form appears to denote the thing that *produces* the mass form.<sup>39</sup> For Fogal, a marker for this kind of noun is that they appear in ‘generational’ environments:

(15a) My scoring of the goal gave me great pleasure.

(15b) My scoring of the goal was a pleasure.

---

<sup>38</sup> This idea – that we can always analyse the count form of a dual life noun in terms of its mass form (or vice versa) – would seem to drop out of many views which say the mass/count distinction is in some sense structural. Roughly speaking, on these kinds of views, mass nouns are *transformed* into count nouns when they’re embedded within certain environments. Bale & Barner have it, for example, that nouns come out of the lexicon neither mass nor count: count noun marking maps denotations that don’t denote individuals to denotations that do; mass noun marking is merely an identity function. On a picture like this, for any dual life noun, the denotation of the count form can be seen as the result of applying some ‘singularizing’ function to the denotation of the mass form. (c.f. Wellwood 2019, p.95-7; Borer 2005, p.101-109)

<sup>39</sup> A similar point is independently made by Grimm (2014). Grimm isolates a class of what he calls ‘psych nouns’, and suggests that many of these nouns denote *states* in their mass form, and the ‘stimuli’ of those states in their count form.



Crucially, we see that ‘reason’ can appear in similar kinds of generational sentence. The facts we call reasons are seemingly just those facts we say give us reason:

(16a) The fact that Sally had a toothache gave Sally reason to go to the dentist.

(16b) The fact that Sally had a toothache was a reason for Sally to go to the dentist.

(17a) The fact she was in intense pain provided Sally with reason to go.

(17b) The fact she was in intense pain was a reason for Sally to go.<sup>40</sup>

Reasons provide agents with reason – that’s the insight from Fogal (and later Fogal & Risberg), and I think there’s surely something right about it.

But care is needed when it comes to implementing this insight within our semantics. The mass noun ‘reason’, for Fogal, seems to be a predicate of *amounts* of reason – amounts of *pro tanto* normative support. There’s a quantity of reason for Sally to go to the dentist which is produced by the fact she has a toothache, and there’s a distinct quantity of reason for Sally to go to the dentist which is produced by the fact that she’d enjoy the drive. By contrast, on the theory I sketched in §2, the mass noun ‘reason’ doesn’t pick out quantities or amounts of reason, and there’s no distinguishing the reason for Sally to go to the dentist that ‘comes from’ her toothache from the

---

<sup>40</sup> Weaver & Scharp (p.145) object to Fogal on the grounds that generational talk breaks down in cases where there are reasons of which the agent is unaware. I do not share their judgements. Suppose Sally has a painless but dangerous cavity in her tooth which she doesn’t know about. There is *a* reason for her to go to the dentist. There is also *reason* for her to go to the dentist. The cavity is what *gives her* reason to go to the dentist. It’s true that she’s unaware she has this reason, but she has it nonetheless.

reason for Sally to go to the dentist that ‘comes from’ her desire to drive. Sally’s (mass noun) reason to go to the dentist is a single thing – a state she’s in, which has a certain size.

Nonetheless, we can adapt a version of the Fogalian insight to the states-based framework. Sometimes we’re in a state because the world stands thus and so. I have support (I am in a support-state) because of my *supports* – various things that give me support. I have purpose (I am in a purpose-state) because of my *purposes* – various aims or goals that give me purpose. I have reason to  $\Phi$  (I am in a reason-state targeting  $\Phi$ ing) because of my various *reasons to  $\Phi$*  – various facts that give me reason.

Semantically, the simplest way to cash this idea out is to have the count-noun ‘reason’ be a predicate of sources of reason states, giving us truth-conditions like so:

[[*The fact that Sally has a toothache is a reason for her to go to the dentist*]] = 1 iff

$\exists s$  [**reason**(*s*) & **holder**(*s*, Sally) & **theme**(*s*, Sally goes to the dentist) & **source**(*s*, *that Sally has a toothache*)]<sup>41</sup>

What is it precisely for some worldly thing (e.g. a fact) to be the source of a state? This is a hard question, in part because source-talk seems appropriate for a whole range of state/count-noun pairs.

---

<sup>41</sup> Two important technical notes. First: standardly, thematic roles like ‘**holder**’ correspond to functions. If  $x \neq y$ , then it can’t both be the case that **holder**(*s*, *x*) and **holder**(*s*, *y*). As I’m thinking of it, ‘**source**’ is a predicate rather than a function – many facts can be the source of a single reason state. If one wanted to, one could preserve the letter of uniqueness by having ‘**source**’ map a state to a set of facts, each of which are reasons. (The count noun ‘a reason’ would then be a predicate of *elements* of sources). Second note. Some linguists (e.g. Parsons 1995) use ‘source’ to pick out a thematic role that helps analyse sentences like “the bird flew from Morocco to Spain” – here Morocco has the ‘source’ role in the event. I suspect this is merely a case of overlapping terminology.

Just as reasons are the source of reason, it's tempting to say too that pleasures are the sources of pleasure, sorrows are the sources of sorrow, surprises are the sources of surprise, supports are the sources of support, helps are the sources of help, purposes are the sources of purpose, justifications are the source of justification, and warrants are the sources of warrant. It would surely be a mistake to say that the nature of the 'source' relation is the same in each of these cases. *Warrants* don't stand to *warrant* in the way that *surprises* stand to *surprise*. For each kind of state, there's room for theorising about the metaphysics of the relation at issue.

Certainly in the case of 'reason' the relevant kind of relation will need to be non-causal: a fact might give me reason to  $\Phi$ , even though I'm unaware of it, or causally isolated from it.<sup>42</sup> The source relation will also need to be *gradable*. Normative reasons can be better or worse – more or less weighty. The fact that Sally has a toothache is a good reason for her to go to the dentist. The fact that she wants to flip through the magazines in the waiting room isn't. The source relation will need to be cashed out in such a way that the former is *more of a source* of her reason-state than the latter.

However the details are filled out, this kind of states-based analysis of the count noun 'reason', would clearly be bad news for the Reasons Firster. Recall – the Reasons Firster wanted to explain (or ground) facts about reason states in terms of facts about reasons. This is a lost cause if *all it is*

---

<sup>42</sup> The fact that the square root of 81 is 9 gives me reason to write '9' on my maths exam. But this fact clearly has no kind of causal influence on me (or anything, for that matter). It is not the source of a reason-state in the same way that going on a rollercoaster might be the source of a pleasure-state.

to be a reason is to be the source of some reason state.<sup>43</sup> To see why, just take Sally's toothache. In virtue of what is this toothache a reason for Sally to go to the dentist? Our answer needs to appeal to reason states: Sally's toothache is a reason for her to go to the dentist, because this fact is one of the sources of the reason state she's in – that's just all it is for a consideration to be a reason. On the states-based analysis it's facts about *reason states* that explain facts about *reasons*, rather than the other way around.<sup>44</sup>

## §4 Conclusion

Reason-talk is not a kind of reasons-talk. When you say that Sally has more reason to go to the dentist than the cinema, you're not making a claim about the weightiness of her reasons. Instead, you're making a claim about Sally and the relative sizes of two states she's in.

---

<sup>43</sup> Here's a shorter argument against Reasons First, which doesn't go by way of explanation or grounding. Reasons First normally has a kind of *primitivism* built into it (Fogal p.83): the property of being a reason cannot be analysed in terms of other normative notions. A semantics for the count noun 'reason' in terms of reason states would clearly violate this primitivism, as this semantics would itself be an analysis of what it is to be a reason in terms of another normative notion.

<sup>44</sup> Sure, the facts *which are* reasons may explain/ground reason states. But the Reasons Firsters wanted to say something much stronger: that facts *about* reasons – i.e. facts of the form *p is a reason for S to Φ* – explain or ground reason states. (See Fogal 2016, f.n.19 and Schroeder 2021, p.4). To help make this clearer, consider 'ought'. Plausibly, facts about what you ought to do are at least partly explained by your reasons. (If you ought to Φ, and *p* is a reason for you to Φ, then *p* helps explain why you ought to Φ.) This can be accepted by anyone – Reasons Firsters like Schroeder or staunch anti-Reasons-Firsters like Broome. Part of what's at issue between the two of them is the explanatory status of facts *about* reasons (facts of the form *p is a reason for S to Φ*), and specifically the question of whether facts about what you ought to do are explained by facts of this form. If it turned out that we could *analyse* '*p is a reason for S to Φ*' in terms of ought – that *all it is* for *p* to be a reason for S to Φ is for it to explain why *S* ought to Φ (c.f. Broome 2013) – then that would settle the explanatory status of facts about reasons. On a Broomean style analysis, it's facts about oughts that help explain facts about reasons, rather than vice versa. A similar dialectic is going on above: the states-based analysis can (and should) concede that *reasons* help explain facts about reason states; what they should deny is that *facts about reasons* explain facts about reason states. Instead, facts about reason states explain facts about reasons.

There is still lots of interesting work to be done. On the semantic side of things, there's still the question of how to formally model 'net' vs 'gross' readings of comparatives involving mass nouns. On the metaphysical side, the most pressing remaining question concerns the nature of the 'source' relation. Sally is in a reason-state targeting going to the dentist. She also has a toothache. The latter is a source of the former. But what precisely is the nature of the source relation? Why is the fact *that she has a toothache* a source of the reason state, but not the fact *that she has blue eyes*?

Even with these questions outstanding, we have come a long way. Spend a few weeks engrossed in contemporary ethical theory and you may well come away thinking that it's reasons – considerations that count in favour of certain things to certain degrees – which sit at the bottom of our normative speech, driving our talk about what we ought to do, should do, must do, what it would be good to, admirable to do, and so on. If the arguments in this chapter are sound, then this picture goes wrong very early on. Forget 'ought', 'must' and 'good' – normative reasons aren't even driving our talk about *what there's reason to do*.

## Chapter 3: Reasons for Non-Agents

Normative reasons are a bit like forces.<sup>1</sup> Just as there are pushes and pulls determining how things in fact behave, there are normative pushes and pulls determining how we *should* behave – facts that exert a kind of normative pressure.<sup>2</sup>

Suppose Larry is deciding whether or not to dance at a party. There are considerations that count in favour of his dancing (e.g. *that it will be fun*) which have a certain kind of normative push. These are the reasons for Larry to dance. There are also considerations that count against his dancing (e.g. *that he is likely to fall over*) which exert their own countervailing force. These are the reasons for Larry not to dance. These normative pushes and pulls interact, and come together to determine what Larry, all things considered, ought to do.

A physical force like gravity is exerted on everything in the natural world. But the kind of normative force exerted by reasons is normally taken to have a much narrower domain. Suppose I go on a walk in the countryside and pick up a rock lying in a field. There are various reasons *why* the rock has the shape it does (to do with facts about the planet's surface, erosion, and so on). But

---

<sup>1</sup> Fogal & Risberg (2023a, p.70-1) develop the analogy with forces in some detail.

<sup>2</sup> Normative reason claims are standardly contrasted with *explanatory* or *motivating* reason claims. We make an explanatory reason claim when we cite some consideration as an explanation, in a sentence of the form *p is a reason why q* ('the fact that I ate bread is a reason why there are crumbs'), *p is a reason that q* ('the fact that I ate bread is a reason that there are crumbs') or – when *x* is a noun phrase – *p is a reason for x* ('the fact that I ate bread is a reason for the crumbs'). We make a motivating reasons claim when we identify a consideration which an agent acted on, in a sentence of the form *S's reason for Φing was that p* ('my reason for eating bread was that I was hungry') or *p is the reason for which S Φd* ('that I was hungry was the reason for which I ate bread').

these are explanatory, not normative reasons. We wouldn't normally think that there's a normative reason *for* the rock to take that particular shape, or any other shape.

Similarly, suppose I throw the rock at a window, causing the window to shatter. My throwing of the rock is a reason *why* the window shattered. But, again, it seems odd to think that my throw provided the window with a normative reason *for* shattering – that my throw somehow counted in favour of its shattering, or helped make shattering the thing which the window ought to do.

How should we explain the fact that there are reasons for Larry to dance, but there aren't reasons for windows to shatter or for rocks to have one shape rather than another? Here's a suggestion: it only makes sense to ascribe normative reasons to *agents* – more or less, things like you or me. Call this idea Agentialism.

**(Agentialism)**       $p$  cannot be a normative reason for S to  $\Phi$  unless S is an agent.

What exactly is an agent? I want to avoid taking too much of a stand on this question, partly because by the end of this chapter it should feel even muddier than it already does. But for the sake of argument (and to make Agentialism as plausible as possible) it does no harm to cast the net widely, and include anything with the capacity for some kind of mental life: humans, aliens, animals, and so on. Rocks, windows, electrons and national parks are not agents.

Agentialism is widely held.<sup>3</sup> Here is Maria Alvarez in her popular introduction to normative reasons:

A reason for  $\Phi$ ing is a reason for *someone* to  $\Phi$ , that is, for someone to do those things that can be done for reasons: act, believe, want, feel an emotion, or whatever else [that] can be done for a reason. (Alvarez, p.10, emphasis added)

And here is T. M. Scanlon endorsing Agentialism early in *What We Owe to Each Other*:

The rudimentary observation that a reason is a consideration that “counts in favor of” something points toward a question, “In favor of what?” and hence toward an important distinction, between those things for which reasons, in the sense I have in mind, can sensibly be offered or requested and those for which they cannot. It makes no sense to demand a reason, in this sense, for an event in the world that is unconnected with any intentional subject. I might ask, “Why is the volcano going to erupt?” But what I would be understood to be asking for is an explanation, a reason why the eruption is going to occur, and this would not (at least among most contemporary people) take the form of giving the volcano’s reason for erupting. (p.18)

---

<sup>3</sup> There are a couple of important exceptions. Stephen Finlay (2014, p.88-9) takes ‘*p is a reason for...*’ to be a predicate of states of affairs which needn’t involve an agent. Likewise, John Broome (2015) writes that a sentence such as “there is reason for the lights to come on at dusk” (p.81) ascribes the ‘reason property’ to a state of affairs (the lights being on at dusk). However, he claims that obtainings of the reason property will typically be ‘owned’ by an agent (e.g. whoever it is who’s in charge of the lights). He does leave it open that there could be unowned reason-obtainings, but writes “there is a case for thinking that ‘There is reason for the lights to come on at dusk’ has a special meaning in this case, and does not report an obtaining of the reason property. It might be evaluative rather than strictly normative, for example.” (p.85). We’ll come back to the suggestion that claims about reasons for non-agents are merely evaluative in §2.2.



Not only is Agentialism widely presupposed by philosophers, it also plays important roles in their theorising about normative reasons. It is increasingly popular to give an analysis of normative reasons in terms of facts about the standards of good practical reasoning. According to ‘Reasoning Views’,<sup>4</sup>  $p$  is a reason for  $S$  to  $\Phi$  just when (roughly speaking)  $S$  would be doing good practical reasoning if they reasoned from the belief that  $p$  to a motivation to  $\Phi$ . Taking Kieran Setiya’s (2014) analysis as our exemplar:

**Reasoning View:** The fact that  $p$  is a reason for  $S$  to  $\phi$  just in case  $S$  has a collection of psychological states,  $C$ , such that the disposition to be moved to  $\phi$  by  $C$ -and-the-belief-that- $p$  is a good disposition of practical thought, and  $C$  contains no false beliefs. (Setiya, p.222)

The Reasoning View says that whether some consideration  $p$  counts as a reason for  $S$  to  $\Phi$  depends on facts about  $S$ ’s psychological states and facts about the standards of good practical reasoning. Non-agents (rocks, windows, electrons, cities) have no psychological states, and practical reasoning just isn’t the kind of thing that they can do. On the face of it, if the Reasoning View is true, then Agentialism is true too – there are no reasons for non-agents.<sup>5</sup>

This chapter argues that Agentialism is false. There are normative reasons for non-agents to be certain ways and do certain things. It looks as though we must therefore reject the Reasoning View,

---

<sup>4</sup> I borrow this terminology from Samuel Asarnow (2017), who presents a kind of Reasoning View. See also Kieran Setiya (2014) and Jonathan Way (2017).

<sup>5</sup> We’ll make this argument more watertight in §3.1.

and any other picture of normative reasons which ties them essentially to distinctively agential capacities.

In §1, I argue that there are reasons for non-agents. It is perfectly sensible to talk of reasons for artifacts, plants and insects to be one way rather than another, or do one thing rather than another.

In §2, I respond to three ways that the Agentalist might try and resist my argument. It's here that I respond to the objection that 'reason for' is ambiguous between a sense that can be applied to non-agents and a sense which can only be applied to agents. In §3, I unpack the consequences of the collapse of Agentalism. The Reasoning View is most likely false, as are some of the weaker ways we might try and link normative reasons with distinctively agential capacities. I close with discussion of a new question that the falsity of Agentalism opens up: *which* things, precisely, are pushed and pulled by normative reasons, if not agents?

There'll be some philosophers who insist that this chapter's thesis is false by definition: they'll say that 'normative reason' is a term of art, and that it has a stipulated meaning which is reserved for a very distinctive phenomenon (e.g. Rossian *prima facie* duties). Very well. In that case, this chapter isn't aiming to show that there are 'normative' reasons for non-agents. What I'm arguing is that there are reasons for non-agents to be some ways rather than other ways, and do certain things rather than other things; that these reasons aren't merely motivating or explanatory; and that they frequently move with claims about what the non-agent concerned ought to be like or do. These reasons sound to me like they're deserving of the label 'normative', but I'm not interested in policing how philosophers use their terms.

## §1 Against Agentialism

### 1.1 Initial cases – reasons for non-agents to be certain ways

You have chickens. Last year, you fenced-off a bit of your yard to keep them safe and stop them from wandering around in your vegetable garden. But it's become clear that the fence is too short. The chickens jump over it, wreak havoc, and get chased by your cat.

These sentences are true:

- (1a) There is a good reason for you to keep your cat indoors for the time being.
- (1b) One of the reasons for you to keep your cat indoors is that the chickens can jump over their fence easily.

These are claims about *your* normative reasons. But notice we can also say (of the fence):

- (2a) There is a good reason for the fence to be a few feet taller.
- (2b) One of the reasons for the fence to be taller is that the chickens are currently able to jump over it easily.

The sentences in (2) are certainly not claims about motivating reasons. The fence cannot do anything at all, let alone be motivated to do one thing rather than another. They're not explanatory reason claims either. The infinitival phrase "the fence to be taller" does not pick out an actual

event, fact, state of affairs or anything else which is apt for being explained. Instead, on the face of it, the sentences in (2) are genuine normative reason claims, just like the sentences in (1). After all, the sentences fit with a number of clearly normative things we want to say about the fence and its height. The fence *ought* to be taller. It would be a *better fence* if it were taller.

If (2a) and (2b) are true, then there are reasons for fences to be some ways and not others. Fences are not agents. So, we have here a case against Agentialism.

Further cases are easy to generate. Your local pub used to have its bathrooms outside next to the beer garden. Patrons kept getting lost, so at the bottom of each menu it says “BATHROOM IN BEER GARDEN”. But recently, the pub has undergone a renovation and they’ve moved their bathroom indoors, next to the bar. These are true:

(5a) Now that the bathrooms are indoors, there’s no longer any reason for the menus to say ‘BATHROOM IN BEER GARDEN’.

(5b) There is more reason for the menus to say ‘BATHROOM NEXT TO BAR’ than ‘BATHROOM IN BEER GARDEN’.

Menus are not agents, but there are reasons for them to say some things and not others.

A final case. Sarah believes that *Titanic*’s long runtime only serves to make the film worse. Sarah believes:

(6) There is no good reason for *Titanic* to be so long.

It's not that Sarah believes there is no enlightening explanation as to *why* the film is so long (perhaps she understands perfectly well why it has the length it does). Instead, her belief has a certain kind of normative content. She believes there's no good reason *for* the film to be so long – that there's no consideration which justifies its length.

## 1.2 An early objection – normative reasons target *doings*

Fences, menus and movies are all powerless to do anything. A fence can be made taller, but it can't make itself taller. A restaurant owner can write something different on a menu, but a menu can't change what's written on it. For *Titanic* to have been shorter, James Cameron would have had to direct it differently.

This motivates an early objection. Surely (the thought might go) the kind of normative reason claim which ethicists are deploying, and which something like the Reasoning View is analysing, targets the performance of actions. But all the sentences above target things *being* a certain way – they all use stative verbs (*be taller, say such and such, have shorter duration, etc.*).<sup>6</sup> So something fishy must be going on.

---

<sup>6</sup> Recall the rough but handy test for the difference between stative and non-stative verbs I mentioned in the previous chapter (found in Dowty p.55): a verb phrase V is non-stative iff “What S did was V” is grammatical, and stative otherwise.

The objection is tempting, but we need to move carefully. Suppose I have an important package being delivered before 12pm today, and I'll need to sign for it. This is clearly true:

- (7) The fact I'll need to sign for the delivery is a reason for me to be at my house this morning.

No philosopher would think twice about agreeing that (7) is a claim about normative reasons, I take it. There's nothing fishy here. But the verb phrase '... be at my house' doesn't pick out an action – it's stative. A rock can be at my house without doing anything at all. So we should not dismiss the sentences in §1.1 merely on the grounds that they do not target the performance of actions.<sup>7</sup>

Secondly, even if it were true that reason ascriptions tend to target the performance of actions, merely pointing this out doesn't really amount to a defence of Agentialism. What the Agentialist needs is some Agentialist-friendly explanation of our inclination to accept the sentences in §1.1. And it's not obvious what such an explanation might look like.

One Agentialist strategy would be to lean hard on the idea that 'reason for' is ambiguous. Perhaps there are two normative senses of 'reason': one which applies to considerations that count in favour

---

<sup>7</sup> It's worth noting that the fact there are reasons to be certain ways is consistent with the spirit of something like the Reasoning View, if not the letter. After all, it seems that practical reasoning could culminate in an intention or motivation to be at home. (One would think that we can intend not merely to *do* particular things at particular times, but, more generally, to *be* thus and so).

of an agent doing something, and one which applies to considerations that count in favour of the things being thus and so.

I discuss ambiguity further in §2, but we can already make an initial case against this suggestion. Here's a standard test for detecting ambiguity. When you have an expression that's ambiguous (like 'bank'), it sounds odd – 'zeugmatic' – to force two senses of the word together within a single sentence, especially if one of the two instances of the word is unpronounced:

(3a) ? Joan deposited money at her bank, and John fished at his.<sup>8</sup>

This test does indeed detect an ambiguity between the normative sense of 'reason' and, say, the explanatory sense. The following is marked:

(3b) ? The fact that you're hungry is a reason for the noise your stomach is making and for you to make a sandwich.

By contrast, we can easily conjoin agential and non-agential normative reason ascriptions. There is nothing objectionable or zeugmatic about the following:

(4) The fact that the chickens keep escaping is a reason for the fence to be a few feet taller and for you to keep your cat indoors for the time being.<sup>9</sup>

---

<sup>8</sup> See discussion of the 'Identity Test' in Zwicky & Sadock p.18-20. For further discussion, see Sennett.

<sup>9</sup> Thanks to Kenneth Black for suggesting I look to this test.

Here, we're seemingly using one and the same sense of 'reason for' to pick out a consideration that both counts in favour of an agent doing something, and in favour of the fence being a certain way. We'll look more at the hypothesis that 'reason for' is ambiguous later, but we have above good *prima facie* evidence that the sentences in (2) use 'reason for' in the very same sense it's used when we talk about what agents have reason to do.

An alternative Agentialist explanation of the sentences in §1.1 posits widespread ellipsis, or some other kind of gap between the surface and logical form of the relevant sentences. One could insist that all the talk of reasons for things to be the case is really shorthand for talk of actions being performed by a salient agent.<sup>10</sup> For example, perhaps when we assert (2a), while we say that there's reason for *the fence* to be taller, really we're talking about *you*:

(2a\*) There is good reason for *you to see to it that* the fence is taller.

Clearly, if (2a) simply meant (2a\*) then we'd have no counterexample to Agentialism. But this is not a good way of going either. Take (6):

(6) There is no good reason for *Titanic* to be so long.

What (6) is saying is that there is *now* no good reason for *Titanic* to be so long. It clearly does not mean the same thing as:

---

<sup>10</sup> Finlay raises and dismisses this suggestion (p.89).



(6\*) There was no good reason for James Cameron to make *Titanic* so long.

Suppose that back in 1997, a sadistic naval historian forced Cameron to include as much historical detail as possible. (6\*) is false, but (6) is (perhaps) true.<sup>11</sup>

Despite these obstacles, some will continue to insist that when we talk about reasons for things to be thus and so, we're not using 'reason' in quite the same way as when we talk about reasons for things to be done. I disagree, but thankfully we don't need to settle that question here. There are further cases which end up making this issue moot. Plenty of non-agents *are* in fact capable of doing things. Clocks tick, sirens blare, pistons pump, toasters toast, sunflowers turn, mosquitos bite, spiders spin, and espresso machines make espressos.<sup>12</sup> And looking at such cases only serves to reinforce our argument against Agentialism.

### 1.3 Reasons for non-agents to *do* certain things

Sometimes there are good reasons for a non-agent to do what it does, and sometimes there are good reasons for it to do something different.

---

<sup>11</sup> This might encourage the alternative Agentialist paraphrase:

(\*) There is no good reason for anyone to see to it that *Titanic* is so long.

But it's hard to see how anyone *could*, now, see to it that *Titanic* has some particular length or other. (How would one go about doing so?) If that's right, then this paraphrase will make (\*) trivially true, when in fact it makes a non-trivial claim about a movie.

<sup>12</sup> I am not saying here that non-agents can perform actions in the thick sense that distinguishes an intentional raising of one's hand from a twitch. But these things can clearly do actions in a way that a fence cannot.

## *Satnav*

Satellite navigation systems ('satnavs') do things: they display instructions on their screen and read out commands. Alan's satnav is sophisticated: it uses data about the traffic to help calculate what's expected to be the quickest route to his destination. One morning, Alan is using his satnav to drive to work. He approaches a junction: he could go left, which would take him on a slightly longer route, or he could take the main road straight on. As it happens, for the last few hours there's been standstill traffic a mile ahead of him on the main road. Once you're on the main road, you can't turn off it. Here are two true claims about Alan's satnav:

- (7a) The fact that Alan would hit standstill traffic if he went straight is a reason for Alan's satnav to direct Alan to the left.
- (7b) The fact that Alan would get to his destination quicker by turning left than by going straight is a reason for the satnav to direct Alan left.

These are apparently normative reasons for a non-agent (a satnav) to do one thing rather than another.

Here's another version of the case. Alan is driving along the same road and gets to the same junction. To his surprise, the satnav directs him left, along the slightly longer route. Later, once he arrives at work, he asks his colleagues who took the main road whether they encountered any traffic. They didn't – the main road was clear all morning. Alan can say, truly:

(7c) There was no good reason for the satnav to send me left at that junction. (So there's probably something wrong with it...)

### ***Smoke alarm***

It's 12pm, and Betty has burned her lunch. The kitchen is thick with smoke. Strangely enough, the smoke detector does not go off. Betty concludes that the smoke alarm must be defective.

Later, she tells her partner Carl that they need a new alarm:

(8a) Betty: I think the alarm is broken – it didn't go off earlier.

Carl: Well, was there any reason for it to go off?

Betty: Yes – there was smoke everywhere.

This dialogue is perfectly ordinary. Betty is committed to:

(8b) There was a reason for the smoke alarm to go off.

Betty and Carl have a neighbour called Diane. Diane never burned her lunch, so her kitchen was never smoky. This is true:

(8c) At 12pm, there was more reason for Betty's smoke alarm to go off than there was for Diane's smoke alarm to go off.

## *Flytrap*

Venus flytraps use fine hair-like protrusions to detect when an insect is in its trap. Once its hairs have been stimulated twice in quick succession, a well-functioning, healthy Venus flytrap snaps its trap shut.

Suppose you're watching a Venus flytrap. A small beetle is walking within the trap, and brushes up against two of the trap's hairs in quick succession. Oddly enough, the trap remains open. The beetle flies away unharmed. You can truly say:

- (9a) The fact that the beetle brushed up against those hairs meant that there was a reason for the plant to shut its trap.
  
- (9b) There was more reason for the plant to have closed *that* trap (where the beetle was) than *that* one (which never got stimulated).

In conclusion, there are reasons for satnavs, smoke detectors and plants to do some things and not others. Contrary, to Agentialism, there are reasons for non-agents.

## §2 Objections and Replies

### 2.1 The Epistemic Diagnosis

The fact that there was traffic up ahead was good evidence that the satnav would send Alan left. The smokiness of Betty's kitchen was good evidence that her smoke alarm would go off. The fact that the Venus Flytrap's protrusions were stimulated was good evidence that it would close its trap.

When  $p$  is evidence for  $q$ ,  $p$  is a reason to believe that  $q$ . In the above cases, then, the facts cited as reasons for S to  $\Phi$  are also reasons to believe that S *will*  $\Phi$ . For example:

(8a\*) There was a reason (for Betty) to believe the smoke alarm would go off.

(8b\*) There was more reason to believe that Betty's smoke detector would go off than that Diane's would go off.

So here's a plausible sounding debunking of the argument in the previous section. All the relevant sentences are false, as written. When we hear them as true it's because we're charitably reinterpreting them as claims about what we have reason to believe, as in (8a\*) and (8b\*). But, of course, it's perfectly consistent with Agentialism that the smokiness of the kitchen is a reason to *believe* the smoke alarm will go off. Call this the **Epistemic Diagnosis**.

Let's start with a small worry. The Epistemic Diagnosis risks overgenerating. It certainly isn't always the case that when *p is a reason to believe that S will  $\phi$*  is true, we're inclined to hear *p is a reason for S to  $\phi$*  as true. Suppose my friend puts a spinning top in an opaque box, starts it spinning and closes the lid. Ten seconds later, she opens it up and tells me it's still spinning. We are not inclined to accept:

- (11) The fact my friend said "the top is still spinning fast" is a reason for the top to spin for at least the next five seconds.

But if the Epistemic Diagnosis is right about why we're inclined to accept the original data, then we would expect to be able to get a true reading of (11) too. After all, my friend's testimony is a reason to *believe* the top will spin for at least the next five seconds.

Second response. The Epistemic Diagnosis is hopeless when it comes to explaining some of the first cases we looked at – the sentences which employed stative verbs. For example, take:

- (2a) The fact that the chickens keep escaping is a good reason for the fence to be taller.

We can easily imagine that the fact cited isn't a reason to believe the fence *will* be taller – just suppose that I'm a sadistic chicken owner, who's less likely to increase the height of the fence, conditional on my chickens being able to jump over it. The Epistemic Diagnosis will need to be accompanied by a different explanation of the original cases.

Third and final response. Suppose the smoke detector is currently malfunctioning (perhaps, for example, it is being set off by steam from a kettle). Betty can say, truly:

(12a) There is no good reason for the smoke detector to be alarming right now.

According to the Epistemic Diagnosis, we hear (12a) as true because this is true:

(12b) There is no good reason to believe that the smoke detector is alarming right now.

But (12b) is false. Betty can plainly hear the detector blaring away.

I conclude that that the Epistemic Diagnosis is a poor explanation of our inclination to accept the various reason-ascriptions I presented in the previous section.

## **2.2 The Ambiguity Diagnosis**

An alternative Agentalist strategy is to posit an ambiguity in ‘reason for’ between the sense that’s employed when we ascribe reasons to non-agents and the sense which is the target of analysis in ethics.

We’ve already seen one piece of data that counts against this. We can run together agential and non-agential reason ascriptions:

(14) The fact the chickens keep escaping is a reason for the fence to be a few inches taller, and for you to keep your cats indoors for the time being.

(15) The fact that the fly is stimulating those hairs is a reason for the plant to close its trap, and thus for you to keep your fingers out of the way.

However, to be fair, the acceptability of (14) and (15) does not decisively refute the Ambiguity Diagnosis. Firstly, while these kinds of tests are good for detecting homonyms like ‘bank’, they’re less probative when we’re dealing with cases of polysemy (that is, cases where there are distinct meanings which are closely related). For example, ‘speech’ is polysemous – it can pick out what is said (a kind of abstract object, perhaps), or it can pick out some particular speaking event. This polysemy is not detected by our test, as Viebahn points out (p.754):

(16) John’s speech was well-crafted but [John’s speech was] barely audible.

(16) is perfectly acceptable, despite the two occurrences of ‘speech’ apparently having different referents.<sup>13</sup> Secondly, even if the test were probative in this case, all the acceptability of (14) and (15) shows is that the sense of ‘reason’ at issue when we ascribe reasons to non-agents can also be used to ascribe reasons to agents. But it’s consistent with this that there’s *another* sense of ‘reason’

---

<sup>13</sup> See Viebahn 2018 for further cases. To be clear, there are many cases of polysemy which our test will correctly pick up on. ‘Dish’ is polysemous: it can pick out the meal, or it can pick out the thing a meal is served on. This polysemy is detected by our test: “The dish was delicious and made of pottery” induces the zeugma effect. ‘Cap’ can pick out the thing on a human head, or on the head of a pencil. “Bob’s cap was fabric and the pen’s was plastic” is similarly marked.



– a sense which doesn't appear in (15) or (16) – which can only be sensibly ascribed to actions performed by an agent.

Point taken. But what the Ambiguity Diagnosis lacks at this point is any positive motivation for thinking that there *is* such an ambiguity. Without such a motivation, any appeal to ambiguity will inevitably look *ad hoc*.

So far as I can see, the Agentialist's best hope for motivating an ambiguity is to look to other normative vocabulary, find some putative ambiguity there that is well motivated, and use that to motivate a corresponding ambiguity in 'reason'. For example, the Agentialist could look to 'ought'. Take a sentence like:

(17) Larry ought to dance.

Some philosophers take (17) to have two readings, corresponding to two distinct senses of 'ought'.<sup>14</sup> One sense of 'ought' relates Larry to the action *dancing*. We can call this the *relational* sense of 'ought'. It's the relational sense of 'ought' which we use to give advice and which, in normal circumstances, moves with judgements about praiseworthiness and blameworthiness.<sup>15</sup>

---

<sup>14</sup> See Schroeder (2011), and the references therein.

<sup>15</sup> Schroeder calls it the 'deliberative' sense, and others call it the 'agential' sense. This terminology is awkward in the present context, as it would seem to prejudge that non-agents, e.g. smoke alarms, cannot be subject to this sense of 'ought'. So I choose to muddy the terminological waters yet further by calling the agential/deliberative sense of *ought* 'relational'.

The other sense of ‘ought’ is *evaluative*. (17) supposedly has a second reading which would more naturally be expressed as *it ought to be the case that Larry dances* or *it would be best if Larry dances*. Suppose Larry’s dancing always spices up the party. If a party is getting dull, you might assert (17), in order to say that the party would be better if it had Larry’s dancing in it. This can be true even if you wouldn’t advise Larry to dance, or even if Larry wouldn’t be blameworthy for not dancing.

So says a certain popular view about ‘ought’. I’m somewhat sceptical that the distinction between the different readings of (17) is best captured via an ambiguity in ‘ought’<sup>16</sup>, but let’s accept the ambiguity for now, assume it’s well motivated, and see how far we get in defending Agentialism.

The Agentialist I’m imagining says that corresponding to the relational sense of ‘ought’, there’s a relational sense of ‘reason for’, relating an agent with an action. And it’s *this* sense of ‘reason for’ which they think is central in ethics, and it’s *this* sense which the Reasoning View is analysing. But (says the Agentialist) there’s also a distinct evaluative sense of ‘reason for’ which moves with the evaluative ‘ought’ – perhaps a one-place operator over states of affairs or propositions. And it’s this evaluative sense of ‘reason for’ which we’re employing when we say there is a reason for the smoke alarm to ring.

This response looks appealing. But I’m sceptical that it will really be of any help.

---

<sup>16</sup> See Chrisman (2012) for a case in favour of a uniform semantics.

Let's start with a small but important point. It's surprisingly difficult to get reason-talk to piggyback off the evaluative 'ought' in the way that the objection imagines. Suppose it ought to start raining in Nevada (which is currently undergoing a drought) and stop raining in Germany (which is flooding). Neither of these have a true reading:

(18) There's a good reason for it to start raining in Nevada.

(19) There's more reason for it to rain in Nevada than Germany.

Or return to Larry again, whose dancing would spice up the party. The world would be best if it contained Larry's dancing, but we can well imagine that nonetheless it isn't advisable for Larry to dance (perhaps, for example, he's finalising some important business deal). The truth of the evaluative ought-claim isn't enough by itself to license any corresponding normative reason talk – these sound false:

(20) There is good reason for Larry to dance.

(21) There is more reason for Larry to dance than for him not to.

That's one problem for the Agentialist's proposal. Here's another: we can seemingly talk about reasons for non-agents to do things even when the world would be no better if they did it. Terry the torturer keeps his victims in the basement. He likes to keep his victims sleep-deprived, as part of his torture regimen. To help him do so, he sets up a smoke alarm in his basement. Every night, he goes down into the basement and burns toast beneath the smoke alarm – setting it off and waking up the prisoners.

Suppose we're watching Terry do his usual sadistic routine. His prisoners are fast asleep, for now. The toast burns, and the smoke floats up towards the detector. If there were any justice in the world, the alarm wouldn't make a peep. Nonetheless, this is still true:

(22) There is a reason for Terry's alarm to go off.

After the fact, we can even conjoin a reason-ascription to the negation of the evaluative claim:

(23a) It ought to have been the case that the alarm didn't ring, but there was a reason for it to do so.

(23b) It would have been best if the alarm hadn't rung, but there was a reason for it to do so.

Overall, it looks very unlikely to me that our inclination to accept the sentences in §1 is explained by the fact that 'reason for' is tracking some evaluative property which moves with the evaluative ought.

### **2.3 The Anthropomorphism Diagnosis**

Here's one final Agentialist defence. Sometimes we ascribe mental states to things which (arguably, at least) don't have them:

- (24) The satnav wants me to turn left here.
- (25) The computer is confused (because you're trying to open too many programs at once).
- (26) The jellyfish believed it was under attack.

The Agentalist may well insist that it's our disposition to anthropomorphise things which explains our inclination to ascribe reasons to non-agents. When we say there's a reason for the smoke alarm to ring, we're (inaccurately) treating the alarm as a person.

It's *possible*, I suppose, that the kind of reason-talk I'm pointing to is founded on this kind of mistake. I don't have any argument which decisively rules this possibility out. But in order for the worry to have much dialectical force, we need some argument for thinking that this *really is* what's going on. (The alternative explanation, that we accept these sentences because they're true, certainly seems simpler).

Regardless, there is at least one good direct argument against the Anthropomorphism Diagnosis. With agents, we can slide between possessive and non-possessive reason-ascriptions straightforwardly:

- (27a) There's a reason for Larry to dance.
- (27b) Larry has a reason to dance.

But with non-agents, the possessive sentences sound unnatural:

(28) ? Betty's smoke alarm has more reason to go off than Diane's does.

(29) ? The Venus Fly Trap has a reason to close its trap.

And when the verbs concerned are stative, the possessive is completely unavailable:

(30) # The fence has a reason to be a few feet taller.

(31) # The menu has no reason to say "BATHROOM IN BEER GARDEN".

Agents *have* reasons. If it were true that our non-agential reason-ascriptions were a result of us mistakenly treating these things as if they were agents, then we'd expect to be able to say that smoke detectors *have* reasons to alarm. Instead, what we find is that we don't think of non-agents as things that have reasons, as you or I have them. We're not inclined to do quite the kind of personification that the Anthropomorphism Diagnosis imagines us doing.

## §3 Implications

### 3.1 Implications for theories of normative reasons

Let's begin by picking up a thread which we began in the introduction. Reasoning Views give an account of what it is to be a normative reason in terms of the standards of practical reasoning: reasons are premises in good pieces of practical thought.

The falsity of Agentialism is bad news for Reasoning Views. On the face of it, the fact that there's a reason for the smoke alarm to go off cannot be explained by appeal to the properties of good practical inferences. Practical reasoning is something we agents do. A theory which attempts to explain why it is that the smokiness of the room counts as a reason for the *alarm* to ring by looking to the properties of good practical thought is plainly looking for an explanation in the wrong place.<sup>17</sup>

It's worth going through this more slowly, for the sake of thoroughness. Let's start by once again taking Setiya's view as our paradigm:

**Reasoning View:** The fact that  $p$  is a reason for S to  $\phi$  just in case S has a collection of psychological states, C, such that the disposition to be moved to  $\phi$  by C-and-the-belief-that- $p$  is a good disposition of practical thought, and C contains no false beliefs. (Setiya, p.222)

---

<sup>17</sup> It's worth stressing that this challenge is distinct, and more intractable, than a related worry for the Reasoning View. One doesn't need to look beyond the domain of agents to come up with cases where  $p$  is a reason for a S to  $\Phi$ , even though the S couldn't reason from the belief that  $p$  to a motivation to  $\Phi$ . One simple kind of example looks to cases in which the subject wouldn't be able to form the belief that  $p$  because they lack the relevant concepts. Take Steve the Surgeon, who lives before the discovery of germ theory. Plausibly, the fact that his hands are covered in bacteria is a reason for Steve to wash his hands before operating on his next patient. But Steve lacks the relevant concepts to even believe that his hands are covered in bacteria, let alone the capacity to reason from that belief to an intention to wash them. This kind of case is somewhat uncomfortable for the Reasoning View, but probably not fatal. After all, we can imagine that if Steve *did* have the relevant concepts, and so was able to believe that his hands were covered in bacteria, then it would indeed be good practical reasoning for him to reason from this belief to a motivation to wash his hands. But the Reasoning View cannot appeal to the same kinds of counterfactuals to explain the existence of reasons for non-agents. We cannot imagine giving a smoke alarm the belief *that it is smoky*, and imbuing it with the ability to do practical reasoning. A smoke detector simply isn't the kind of thing that could have attitudes or do reasoning. Thanks to Kieran Setiya for discussion.

The fact that there's smoke is a reason for the smoke detector to ring. According to the Reasoning View, this is true just in case:

The smoke alarm has a collection of psychological states, *C*, such that the disposition to be moved to ring by *C*-and-the-belief-that-there's-smoke is a good disposition of practical thought, and *C* contains no false beliefs.

The smoke alarm has no psychological states: *C* is empty. So – at least as written – if the Reasoning View wants to capture the fact that the smokiness of the room is a reason for the alarm to ring they'll need to say:

The disposition to be moved to ring by the belief that there's smoke is a good disposition of practical thought.

It's not clear whether this even makes sense – it certainly isn't true. Practical thought doesn't give the single-premise inference 'there's smoke, so ring!' the thumbs-up. (Nothing capable of practical thought could perform reasoning which corresponds to that inference, as nothing capable of practical thought is capable of being moved to *ring*, in the relevant sense).

So, if I'm right that there are reasons for non-agents, then the Reasoning View is false. But the collapse of Agentialism also has consequences for a host of weaker principles. Here's a big one:



**Internalism:** If  $p$  is a reason for  $S$  to  $\Phi$ , then  $S$  is capable of being motivated to  $\Phi$  by the belief that  $p$ .<sup>18</sup>

If Internalism is true, then the fact *it will be fun* is a reason for Larry to dance only if Larry could be motivated to dance on the basis of this consideration. As written, Internalism presupposes Agentialism twice over: only agents can have beliefs, and only agents can be motivated by a consideration to do one thing rather than another. So, the counterexamples we gave to Agentialism are also counterexamples to Internalism. The fact that my kitchen is thick with smoke is a reason for my smoke detector to start alarming. But my smoke detector is not capable of being motivated to start alarming by a belief that my kitchen is smoky.

Here's another principle that will have to go, if Agentialism is false. There's an important sense in which I cannot twitch my arm or sneeze *for* a reason. I might twitch or sneeze because of some electrochemical goings-on in my body, but these movements aren't motivated in the right kind of way to be cases of acting for a reason. Here's a perhaps tempting thought: it follows from the fact that one couldn't twitch or sneeze *for a reason* that there cannot be *reasons for twitching* or *reasons for sneezing*. That's certainly the view we saw Alvarez express in the introduction: normative reasons are always reasons "to do those things that can be done for reasons: act, believe, want, feel an emotion, or whatever else [that] can be done for a reason." (Alvarez, p.10). This gives us:

**Acting For:** There is a reason for  $S$  to  $\Phi$  only if  $S$  could  $\Phi$  for a reason.

---

<sup>18</sup> The above definition is drawn from (Setiya, 2012, p.4). The huge internalism literature begins with Williams (1979). For a survey, see Finlay & Schroeder (2017).

Acting For appears to entail Agentialism. On the face of it, a non-agent like a satnav or a plant cannot act for one reason rather than another. Betty's alarm can't go-off *in order to* alert her to the fire – the fire merely causes it to do so. The ringing of Betty's smoke alarm is caused by the smoke in a similar way to how my twitch is merely caused by various electrochemical goings-on. Nonetheless, I've argued, there are reasons for Betty's smoke alarm to do one thing rather than another. So Acting For looks to be false.<sup>19</sup>

To be clear – none of the considerations in this chapter directly speak against various restrictions of these principles. We could weaken them, like so:

**Internalism restricted:** If  $p$  is a reason for S to  $\Phi$ , then: either (i) S is capable of being motivated to  $\Phi$  by the belief that  $p$ ; or (ii) S is a non-agent.

**Acting For restricted:** If  $p$  is a reason for S to  $\Phi$  then either: (i) S could  $\Phi$  for a reason; or (ii) S is a non-agent.

These may be true. But not only are these principles less interesting (in virtue of being disjunctive), but I also take it that these kinds of claims look less plausible than they did before. Part of the appeal of these sorts of principles was that they flowed from a particular picture of *what it is* to be

---

<sup>19</sup> This would have important consequences when it comes to theorising about normative reasons for agents. Could there be normative reasons for someone to sneeze? Perhaps. It would depend on how we fill out the details of a non-agential theory of normative reasons, but the idea that there's a good reason for someone to sneeze when, say, there's something irritating their nasal cavities has some plausibility to it. Some people suffer from the Photic Sneeze Reflex: they sneeze (sometimes compulsively) when exposed to bright light. Part of what makes this condition interesting, I take it, is that in these cases agents sneeze when there's no good reason for them to do so.

a normative reason – a picture that linked reasons up with distinctively agential capacities (motivation, intentional action, practical reasoning and so on). Without an Agentalist picture of what it is to be a normative reason, it becomes less clear why we should want to endorse one of these principles in the first place.

### 3.2 Implications for the scope of normativity

To what kinds of things are normative reason ascriptions sensibly made? The Agentalist's answer was *agents*. We've seen where that answer goes wrong. But what is the right answer?

Cast your mind back to the rock found in a field. Rocks in a field simply have the shapes they do, and there's typically nothing normative to say about it:

(24) ? There's more reason for the rock to have this shape than there is for it to have a slightly rounder one.

But some rocks are special – they're imbued with functions or purposes, and there are standards for their goodness or badness. Suppose I want a new paperweight. I find a rock in a field, and grind one side of it down until it's flat, so that it can sit nicely on a stack of papers. If someone asks me whether there's a good reason for my rock to have a flat base, I can say yes: rocks which don't have flat bases wobble and roll around, and don't make for good paperweights. Now there's a good reason for my rock to have the shape it does. There's more reason for it to have this shape than any other.

This kind of thought suggests that one needs some sort of teleological standard in the background to get reason-ascriptions going. In all the cases we looked at in §1, the non-agents concerned have functions, aims or purposes, which set some kind of standard for the object in question.<sup>20</sup> A well-functioning, non-defective smoke alarm rings just when it's exposed to smoke. So, if it's exposed to smoke, then that's a reason for it to ring. A well-functioning, non-defective Venus flytrap shuts its trap just when it's stimulated. So, if it's stimulated, that's a reason for it to shut its trap. And so on.

An appeal to functions, purposes or standards can get us quite far. Scanlon, in the passage quoted in the introduction of this chapter, claims that there are no reasons for volcanos to erupt. I'm inclined to agree. This does seem quite an odd thing to say:

(25) ? The fact that there'd been an earthquake a few years earlier was a reason for Vesuvius to erupt.

It's not like the earthquake counted in favour of the volcano's erupting, or made erupting the thing that the volcano ought to do. Similarly, I'm hesitant to accept:

(26) ? The fact that the temperature is rising is a reason for the ice shelf to drop into the ocean.

---

<sup>20</sup> Thomson (2008) develops something in the vicinity of these thoughts into an analysis of defectiveness, and later an analysis of 'ought'.

(27) ? The fact that it was hit hard by a stone was a reason for the window to smash.

Our hesitancy to accept these sentences is explained if reason-talk is only licensed when there's some standards for how the thing involved functions well – how it ought to be, or how it ought to respond to stimuli. It's no part of the function of volcanos that they erupt, or of ice shelves that they melt. It's not as though a volcano which fails to erupt in response to an earthquake is a defective or bad volcano. An ice shelf that fails to drop in response to rising temperatures is not a defective or bad ice shelf.

But there are perhaps edge-cases where appealing to teleology might not do all the work we need it to. Windchimes are (one might think) imbued with a certain kind of function – a good windchime produces a pleasant chiming sound when blown by the wind. Suppose I set up some windchimes on my front porch. The wind blows, causing the windchimes to knock together and produce a chiming sound. Was the fact it was windy a reason for my windchime to chime? My first inclination is to say *no*, though at this point my judgements (and the judgements of my informants) are very fragile. I'm inclined to say that the wind *caused* the chimes to chime, but it wasn't a reason *for* the chimes to chime. It's not as though the windiness counted in favour of chiming, or made chiming the thing which the wind chime ought to do.

Compare this to a case in which I set up an electronic wind detector on my front porch, which plays a chiming sound when it detects the wind rising above a certain speed. In this case, it really does seem that the fact it was windy was a reason for my detector to chime. The windiness made chiming the thing that the detector ought to do.

One possible explanation of what's going on would continue to appeal to teleology. There's plenty of room to doubt whether the windchimes really do have the *function* of chiming when it's windy, or whether this is just a case of them being reliably caused to do so. Sure, they were built with a certain purpose in mind, and I put them outside with the intention that they behave in a certain way (with the intention that they make a pleasant sound when the wind blows). But this sort of process doesn't seem to be generally sufficient to give something a function or a purpose. If I want to play a cruel prank on my neighbours, I might put a wineglass on the edge of my roof, intending it to fall and smash when the wind blows. But if I do so I haven't thereby succeeded in endowing the wineglass with the function of smashing when it's windy. (It seems to me that this would be so even if I made the wine glass specially, in order to play precisely this prank).

Another explanation (compatible with the first) would emphasise the different kind of relation that the windchimes and wind detector each stand in to what they do.<sup>21</sup> The windchimes chime because they're blown around by the wind – they're *moved*, and so they chime. By contrast, when the wind detector chimes, it seems to be true to say that it moved *itself* – some sensitive internal process occurred within the detector, which culminated in one part of the detector sending a signal to some other part of it, which then played a chiming sound. The difference between the windchimes and the wind detector might, then, correspond to the kind of difference Michael Thompson alludes to in this memorable passage:

---

<sup>21</sup> Thanks to Jack Spencer and Brad Skow for discussion on this point.

How does a bird's progress out of the stadium, where it has been mistaken for a fast ball, differ from a progress out in search of better food? It is usual to say that, in the one case, the bat moves the bird, and that, in the other, the bird *moves itself*. (Thompson, p.44, emphasis in original)<sup>22</sup>

Both the bird hit by the bat and the bird on the search for better food *move*. But they stand in different relations to their movement – only the latter moves itself. Similarly, both the windchimes and the wind detector *chime*. But they plausibly stand in different relations to their chiming – only the latter moves itself to chime. Reason-talk may track a distinction between full-blooded *do*-ers, like wind detectors (or satnavs, Venus flytraps, animals and humans) and objects which merely do things in virtue of having things done to them, like the windchimes.<sup>23</sup>

Either way, there are interesting lessons to be learned from looking at cases of this sort. Our task is to trace the shape of a line drawn in nature, between those things to which reasons can be sensibly ascribed and those things to which they cannot. We saw the Agentalist draw the line in the wrong place: they drew it sharply around all and only the agents. Now we're attempting to feel out where the line is really drawn. Cases like the one involving the windchimes are interesting because they're cases in which we're bumping up against the murky outer limits of sensible reason-ascriptions – we're approaching the point at which normative talk generally starts to break down. It's not surprising that our judgements themselves become unstable at these outer limits. Whatever

---

<sup>22</sup> Note that Thomson ultimately thinks appealing to the reflexive 'moves itself' doesn't help pin down what he's interested in (roughly, an account of life).

<sup>23</sup> The principle in the background can't be something quite as simple as 'for  $p$  is a reason for  $S$  to  $\Phi$  to be true,  $S$  has to be a self-mover'. At the very least, we'd need to restrict the principle to non-stative ' $\Phi$ 's. There are reasons for fences to be tall. Clearly a fence cannot move itself to be taller than it is.

we ultimately want to say about an edge-case like this (whether we want to appeal to self-movement, to a merely telic explanation, or something else completely), the theory we posit should respect the kind of instability we're seeing. A good theory will not attempt to trace sharp lines where there are really none to be found.

## **§4 Conclusion**

There are reasons for non-agents to be certain ways: for chicken coop fences to be tall and for menus to display accurate information. There are also reasons for non-agents to do certain things: for satnavs to send drivers on routes that avoid traffic and for Venus flytraps to close around insects when stimulated. Accordingly, we should not build a theory of what it is to be a reason out of distinctively agential materials. The force provided by normative reasons pushes and pulls more than just us agents.



## Bibliography

- Alvarez, M. (2010). *Kinds of Reasons: An Essay in the Philosophy of Action*. Oxford: Oxford University Press
- Asarnow, S. (2017). The Reasoning View and Defeasible Practical Reasoning. *Philosophy and Phenomenological Research* 95(3):614-636
- Bale, A. & Barner, D. (2009). The Interpretation of Functional Heads: Using Comparatives to Explore the Mass/Count Distinction. *Journal of Semantics* 26(3):217-252
- Barner, D., & Snedeker, J. (2005). Quantity judgments and individuation: Evidence that mass nouns count. *Cognition*, 97(1):41–66
- Blumberg, K. & Hawthorne, J. (2021a). Inheritance: Professor Procrastinate and the Logic of Obligation, *Philosophy and Phenomenological Research*. Advanced online publication. <https://doi.org/10.1111/phpr.12846>
- Blumberg, K. & Hawthorne, J. (2021b). Wanting what's not best, *Philosophical Studies* 179(4):1275-1296
- Borer, H. (2005). *Structuring Sense Volume I: In Name Only*. Oxford: OUP.
- Broome, J. (2013). *Rationality Through Reasoning*. Malden, MA: Wiley-Blackwell
- Broome, J. (2015). 'Reason versus ought', *Philosophical Issues*, 25.
- Cariani, F. (2011). 'Ought' and Resolution Semantics, *Nous*, 47(3):534-558.
- Cariani, Santorio & Wellwood. (2018). 'Confidence Reports'. (ms.)
- Chierchia, G. (1998). Plurality of mass nouns and the notion of 'semantic parameter'. In: *Events in Grammar*. Edited by Rothstein. Kluwer. pp. 53-103
- Chrisman, Matthew. (2012). "'Ought' and Control', *Australasian Journal of Philosophy* 90(3):433-451
- Dorr C., Nebel, J. & Zuehl, J. (2022). The case for comparability, *Nous*. Advanced online publication. <https://doi.org/10.1111/nous.12407>
- Dowty, D.R. (1991). The Semantics of Aspectual Classes of Verbs in English. In: *Word Meaning and Montague Grammar*. Studies in Linguistics and Philosophy, vol 7. Springer, Dordrecht.
- Finlay, S. (2014). *Confusion of Tongues: A Theory of Normative Language*. New York: OUP USA

- Finlay, S. & Schroeder, M. (2017). Reasons for Action: Internal vs. External, *The Stanford Encyclopedia of Philosophy* (Fall 2017 Edition), Edward N. Zalta (ed.).
- Fogal, Daniel. (2016). Reasons, Reason, and Context, in Errol Lord & Barry Maguire (eds.), *Weighing Reasons*
- Fogal, D. & Risberg, O. (2023a). Explaining normative reasons, *Nous* 57(1):51-80
- Fogal, D. & Risberg, O. (2023b). The weight of reasons. *Philosophical Studies*.  
<https://doi.org/10.1007/s11098-023-01974-y>
- Gillon, B. (1992). Towards a Common Semantics for English Count and Mass Nouns. *Linguistics and Philosophy*, 15(6)
- Grimm, S. (2019). Individuating the Abstract. *Proceedings of Sinn Und Bedeutung*, 18:182-200.
- Hedden, B. & Muñoz, D. (Forthcoming). Dimensions of value, *Nous*.
- Iatridou, S. & Zeijlstra, H. (2013). Negation, Polarity, and Deontic Modals, *Linguistic Inquiry* 44(4): 529-568
- Jackson, F. & Pargetter, R. (1986). Oughts, options, and actualism, *Philosophical Review*, 95 (2):233-255.
- Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy* 30(1):1 - 45
- Kiesewetter, B. (2015). Instrumental Normativity: In Defense of the Transmission Principle, *Ethics*, 125(4):921-946.
- Kim, Jaegwon, (1993). Events as Property Exemplifications. In *Supervenience and Mind*. Cambridge University Press, Cambridge.
- Kiss, T., Pelletier, F.J., Husić, H. & Poppek, J. (2017). Issues of Mass and Count: Dealing with ‘Dual-Life’ Nouns. In *Proceedings of the 6th Joint Conference on Lexical and Computational Semantics*, p.189–98
- Kolodny, N., & MacFarlane, J. (2010). Ifs and oughts, *Journal of Philosophy*, 103:115-143.
- Kratzer, A. (1991). Modality. In von Stechow & Wunderlich (eds.), *Semantics: An International Handbook of Contemporary Research*, 639–650. Berlin: de Gruyter.
- Link, G. (1983). The logical analysis of plurals and mass terms: A lattice-theoretical approach. In R. Bauerle, C. Schwarze, & A. Stechow (Eds.), *Meaning, use, and interpretation of language*. de Gruyter. Berlin. 302–323

- Maienborn, C. (2007). On Davidsonian and Kimian states. In: I. Comorovski & K. von Heusinger (eds.). *Existence: Semantics and Syntax*. Dordrecht: Springer, 107–130.
- Maienborn, Claudia. (2011). Event semantics. In Maienborn, von Heusinger & Portner (eds.), *Semantics; Volume 1*. (HSK Handbook series), Berlin, New York: Mouton de Gruyter.
- Moltmann, F. (2013) *Abstract Objects and the Semantics of Natural Language*, Oxford: Oxford University Press.
- Muñoz, D. & Spencer, J. (2020). Knowledge of Objective ‘Oughts’: Monotonicity and the New Miners Puzzle, *Philosophy and Phenomenological Research*, 103(1):77-91.
- Nebel, Jacob M. (2019) ‘Normative Reasons as Reasons Why We Ought’. *Mind* 128(510):459–484
- Parfit, D. (2011). *On What Matters: Volume One*. Oxford: Oxford University Press.
- Parsons, T. (1990). *Events in the Semantics of English: A Study in Subatomic Semantics*. MIT Press
- Parsons, T. (1995). Thematic Relations and Arguments. *Linguistic Inquiry*, 26(4):635-662
- Pasternak, R. (2019). A lot of hatred and a ton of desire: intensity in the mereology of mental states. *Linguistics and Philosophy*, 42(3):267-316
- Phillips-Brown, Milo. (2021). What Does Decision Theory Have to Do with Wanting? *Mind*, 130(518), 413-437.
- Portmore, D. (2019). *Opting for the Best: Oughts and Options*. New York, NY, USA: Oxford University Press.
- Scanlon, T.M. (1998). *What We Owe to Each Other*. Belknap Press of Harvard University Press.
- Schlenker, P. (2009). Local Contexts, *Semantics and Pragmatics*, 2(3):1-78.
- Schroeder, M. (2008). ‘Having Reasons’, *Philosophical Studies*, 139(1):57-71
- Schroeder, M. (2011). Ought, Agents, and Actions, *Philosophical Review*, 120(1):1-41
- Schroeder, M. (2021). *Reasons First*. Oxford: Oxford University Press.
- Schwarzschild, R. (1992). Types of Plural Individuals. *Linguistics and Philosophy* 15(6):641-675
- Schwarzschild, R. (2022). ‘Pure Event Semantics’. Available online: <https://ling.auf.net/lingbuzz/006888>

- Setiya, K. (2012). Introduction: Internal Reasons. In Kieran Setiya & Hille Paakkunainen (eds.), *Internal Reasons: Contemporary Readings*. MIT Press
- Setiya, K. (2014). What is a Reason to Act? *Philosophical Studies* 167(2):221-235
- Skow, Bradford. (2019). *Causation, Explanation, and the Metaphysics of Aspect*. (OUP)
- Smith, M. (1994). *The Moral Problem*. Oxford: Blackwell Publishers.
- Sennet, A. Ambiguity. *The Stanford Encyclopedia of Philosophy* (Summer 2023 Edition). Edward N. Zalta & Uri Nodelman (eds.).
- Snedegar, J. (2012). Contrastive Semantics for Deontic Modals. In Martijn Blaauw (ed.), *Contrastivism in Philosophy: New Perspectives*. Routledge.
- Snedegar, J. (2017). *Contrastive Reasons*. Oxford: Oxford University Press.
- Thompson, M. (2008). *Life and action: elementary structures of practice and practical thought*. Cambridge, Mass.: Harvard University Press
- Thomson, J. J. (2008). *Normativity*. Open Court.
- Viebahn, E. (2018). Ambiguity and Zeugma. *Pacific Philosophical Quarterly*, 99(4):749-762
- Von Fintel, K. (1999). NPI Licensing, Strawson Entailment, and Context Dependency, *Journal of Semantics*, 16(2):97–148.
- Von Fintel, K. (2012). The best we can (expect to) get? Challenges to the classic semantics for deontic modals. Paper for a session on Deontic Modals at the Central APA, February 17, 2012. <http://mit.edu/fintel/fintel-2012-apa-ought.pdf>.
- Von Fintel, K. (2021). How weak is your *want*? (m.s.). doi:10.5281/zenodo.5123513.
- Von Fintel, K. & Iatridou, S. (2008). How to say ought in Foreign: The composition of weak necessity modals. In Guéron & Lecarme (eds.), *Time and modality*, 115–141.
- Von Fintel, K. & Iatridou, S. (2023). Prolegomena to a theory of X-marking. Forthcoming in *Linguistics and Philosophy*. URL: <http://mit.edu/fintel/fintel-iatridou-2023-x.pdf>
- Way, J. (2017). Reasons as Premises of Good Reasoning. *Pacific Philosophical Quarterly*, 98(2)
- Weaver, Bryan R. & Scharp, Kevin. (2019). *Semantics for Reasons*. (OUP)
- Wellwood, Alexis. (2019). *The meaning of 'more'*. (Oxford Studies in Semantics and Pragmatics)

Williams, B. 1979. 'Internal and External Reasons' In Ross Harrison (ed.), *Rational Action*. Cambridge University Press. pp. 101-113.

Zwicky, A. & Sadock, J. (1975). 'Ambiguity Tests and How to Fail Them'. *Syntax and Semantics* 4:1-36.