Essays on Epistemic Structure

Rachel Elizabeth Fraser

D.Phil Philosophy

Linacre College

Trinity Term 2016

Word count: 64, 102.

Table of Contents.

Acknowledgements

The acknowledgements sections I enjoy reading the most are obnoxious, ungrateful and hubristic. In the preface to her 1966 play, Valerie Solanas wrote:

> I dedicate this play to ME, a continuous source of strength and guidance, and without whose unflinching loyalty, devotion, and faith this play would never have been written. Additional acknowledgements: Myself–for proof reading, editorial comment, helpful hints, criticism and suggestions and an exquisite job of typing. I –– for independent research into men, married women and other degenerates (quoted in Fahs, 2008).

Anyone who shares my tastes will be disappointed –– I lack Solanas' singular personality, as well as having rather more to be thankful for. Thanks go to Pamela Sue Anderson, Max Baker-Hytch, Soheila Benlevi, Matt Benton, Stephen Blamey, Paula Boddington, Joe Carlsmith, Sam Carter, Kam Chadha, Rey Conquer, Jon Courtney, Christina Dietz, Wolsey Fitzherbert, Hilary, Jack, and Steve Fraser, Mikkel Gerken, David Glick, Lydia Heaton, Catherine Holdsworth, Daniel James, Alex Kaiserman, Sarah Karacs, Nakul Krishna, Katharine Jenkins, James Kirkpatrick, Darius Lagencrantz, Rae Langton, Paul Lodge, Annina Loets, David Mathers, James Matharu, Tushar Menon, Andreas Mogensen,

Beau Madison Mount, Andrew Peet, Ian Phillips, Martin Pickup, Alex Roberts, Zoe Saunders, Johanna Schnurr, Robert Simpson, Amia Srinivasan, Jason Stanley, Robert Stearn, Eloise Stonborough, Olúfẹ́mi O. Táíwò, Hillary Taylor, Jacob Williamson, Tim Williamson, Stephen Wright and Andy Yu. I am particularly grateful to both my supervisors, Lizzie Fricker and John Hawthorne, and special thanks go to John for his continued supervision of me after his move to California.

I began to write this thesis in the autumn of 2012, at the same time as I began to train as a line-worker with Oxfordshire Sexual Abuse and Rape Crisis Centre (OSARCC). I am indebted to everyone I worked with at OSARCC for offering me fellowship, sustenance and camaraderie.

I wrote much of this thesis whilst in receipt of an AHRC studentship, and a Ruth and Neville Mott scholarship from Linacre College. I'm very thankful to the AHRC and Linacre College, and to Dr Cecily Crampin, Mrs Alice Crampin and Professor Michael Crampin. The final sections of this thesis were written whilst living in the house of Fram Dinshaw and Claudia Fitzherbert, and I remain grateful for their hospitality.

References

Fahs, B. 2008. The Radical Possibilities of Valerie Solanas. *Feminist Studies* 34 (3): 591–617.

Abstract

This thesis is made up of five chapters: one integrative chapter, and four discrete essays. Each of the four essays is concerned with some aspect of epistemic structure. The first essay, *Risk, Doubt and Transmission,* explores structural analogies between testimonial transmission and closure principles. The second, *Testimonial Pessimism,* traces the connections between moral and aesthetic testimony's putative shortcomings and the philosophy of language. The third essay, *Epistemic Ambivalence*, argues that the ideology of higher-order defeat is paradox prone, and as such, ought to be rejected. The fourth essay, *Tragedy and Rationality,* makes vivid some costs associated with endorsing an anti-akratic constraint on rationality.

Epistemology As Romance

Kant made epistemology Romantic

– Rorty

–

No, no, no, no! Come, let's away to prison.
We two alone will sing like birds i' th' cage.

– King Lear, Act V, Scene III

1. Rorty, Lear and The Fool.

Genre is a funny thing, and not much thought of by philosophers. Even those anglophone philosophers who are self-conscious about their style -- I am thinking, in particular, of Bernard Williams, for whom 'to discover the right style is to discover what you are really trying to do' -- tend towards formal conservatism (Williams, 1993).[1] Perhaps we should look upon this lacuna with relief: often when analytic philosophers do purport to attend to the stylistic tics of their discipline, it is with an all-too-hearty confidence in the clarity of its prose (and a snide contrast with the continental tradition hovering in the background). The stylistic homogeneity of contemporary analytic philosophy is, I suspect, mostly an

[1] I am indebted to Nakul Krishna for discussion of Williams.

unmysterious residue of its institutional placement, and the short history of the discipline ties it tightly -- perhaps more tightly than its humanistic companions -- to the university in its more technocratic, or what Collini calls its 'credentialising', incarnations (2012).

When philosophers do talk about genre, it tends to be in an expansionist spirit: we are urged -- quite rightly -- to be alert to the often covert operations of genre in the construction of a philosophy canon which excludes women, non-Europeans, and people of colour. Gardner draws our attention to 'non-standard philosophical forms': poem, allegory, epistle, and novel. These forms have, she claims, been more accessible to women than, say, the paper or the treatise, but we fail -- however deliberately -- to recognise these 'genres' as philosophical (2003). Nathaniel ~~Coleman~~ has shown us that despite enslaved Africans having argued 'from the moment they were abducted' that slavery was unjust, they fail, even now, to be recognised as *philosophers*:

> if they did write anything, it was only "stories,"' and so where '[l]iterature, history and politics have treated…persons enslaved-as-negro, and, more generally, other persons racialised-as-black, as artists, biographers and campaigners…philosophers have not tended to treat such persons as philosophers (2015).

My appeal to genre here is less worthy, rather more frivolous, and certainly more self-indulgent than those made by Gardner and ~~Coleman~~. A severe clipping of one's ambitions is, it seems, a near requirement of ever finishing a thesis, and my appeals to genre (in general) and to romance (in particular) are attempts at gilding circumspection. This is a thesis in epistemology, written in Oxford, and in the early years of the twenty-first century. Perhaps predictably then, it contains no essays on skepticism. Its otherwise-occupied essays do not even gaze in that direction. Insouciance is one way of honouring the anti-skeptical impulse -- Moorean 'furious common sense', as Cavell (1979) puts it, is another. But even a modest genealogy of the epistemological pieces I do my part to shift around the board -- they come, for the most part, with a heavy externalist flavour -- would show marks left on them by the skeptic's chisel, or the threat of it. If any anti-skeptical impulse is to be found, congealed, in the essays to follow, it is a bequest of recent intellectual history.

The skeptic, in this imagining, with his chisel in hand, is an artisan -- not a romantic hero -- and it is in this unromantic pose that he has carved out so many of the contours of contemporary epistemology. Rorty, though, distinguishes between narrow and broad forms of skepticism, and labels my artisan an envoy of the former. Skepticism, in its narrow 'professionalised' sense, articulates something like the worry that our ordinary ascriptions of

knowledge and justified belief are false. But in its broader forms skepticism is not really a 'worry' at all -- the language of 'the worry' being shot through with all the studied understatement and intellectual sang-froid of 'English' philosophy that Rorty derides. Rather, it is a 'longing' or a 'terror' that -- as so many longings and terrors do – appears to us in different guises, and, despite vivacity and urgency, resists articulation that is not impressionistic or couched in metaphors that we cannot hope to cash. Rorty's focus is on epistemic anguish, but his 'broad' skepticism comes in playful as well as terrorised varieties. Foucault records that his response to Borges' 'Chinese encyclopedia', which he regarded as 'breaking up all the ordered surfaces and all the planes with which we are accustomed to tame the wild profusion of existing things,' was laughter (1994). By Rorty's lights, at least, Bernard Williams is attempting to pry Descartes from the hands of the 'professionals' when he construes the project of Pure Enquiry as:

> A matter not just of overcoming limitations on enquiry and hence of occasional error, as understood within the framework of our outlook, but of overcoming any systematic bias or distortion or partiality in our outlook as a whole, in our representation of the world: overcoming it, that is to say, in the sense of gaining a standpoint (the absolute standpoint) from which it can be understood in relation to reality, and comprehensively related to other conceivable representations (1990).

This is epistemology as romance. The literary theorist Frederic Jameson tells us that:

> Romance as a literary form is that event in which the transcendental horizon -- the overall organisational category within which the various inner worldly experiences take place -- of my experience becomes precisely visible as something like an inner worldly object in its own right (1975).[2]

It discloses, or promises to disclose, a way, as Rorty puts it, 'to hold the world in one hand and our descriptions of it in the other and compare the two'. (Note: this is not quite the way Moore talks about hands.) If one is in the grip of romantic anguish, feeling oneself a caged animal, and one learns that, all the same, one's quotidian knowledge ascriptions are safe, then it is natural to respond that the language of 'knowledge' is just not the emancipatory rhetoric one took it to be. Sometimes it is hard not to feel about Rorty's 'English' philosophy the way Jill feels about Eustace in C. S. Lewis' *The Last Battle*: 'This was very good sense but, at the moment, Jill hated Eustace for saying it. He was fond of being dreadfully matter of fact when other people got excited' (2001).

Epistemology-as-romance may fetishise what Williams calls 'the absolute standpoint', but for too long this standpoint was the only one epistemologists cared about. Projects of

[2] I am indebted to Robert Stearn for discussion of Jameson.

Impure Enquiry are now well under way: projects which seek to emphasise, without denigrating, the 'socially situated' character of knowers and of knowing. Standpoint theory, in its Marxist, feminist, and Black-feminist flavours, tells us, variously, how 'material life not only structures but sets limits on the understanding of social relations', that 'in systems of domination the visions available to the rulers will be both partial and perverse', and that Black-feminist thought is properly approached as 'situated in a context of domination…not…a system of ideas divorced from political and economic reality' (Hartsock, 1987; Collins, 2001).

Sometimes these moves towards a socialised epistemology are conceived of as complementary to 'traditional' approaches, as when Fricker admits that '[m]any philosophical questions may be best served by the traditional, maximally abstracted conception of the human subject' whilst recognising that 'confining oneself to that…philosophical repertoire incurs a needless impoverishment' (2007). More often they are presented as a more-or-less tacit rebuke to the 'idealised' epistemic agent they find in the pages of traditional epistemology, who strides and swaggers, Flashman-like, across social

reality, always indifferent to its peaks and troughs.[3] Flashman, of course, is a grotesque vision of the romantic hero, and insofar as a traditional epistemology does traffic in his ghosts, this represents a degenerate romantic impulse. Just as romantic heroes can be allegorical without being invulnerable, Charles Mills observes in *Ideal Theory as Ideology*, that models can abstract without idealising (2005). Traditional epistemology (when it is any good at all) is never quite a fantasy of invulnerability, however much it may neglect to heed the ways in which epistemic frailty is moulded by our social identity. And standpoint theory has its own romantics -- Williams' Descartes dovetails remarkably with Patricia Hill Collins:

> Each group speaks from its own standpoint and shares its own partial, situated knowledge. But because each group perceives its own truth as partial, its knowledge is unfinished. Each group becomes better able to consider other groups' standpoints without relinquishing the uniqueness of its own standpoint or suppressing other groups' partial perspectives. "What is always needed in the appreciation of art, or life," maintains Alice Walker, "is the larger perspective. Connections made, or at least attempted, where none existed before, the straining to encompass in one's glance at the varied world the common thread, the unifying theme through immense diversity." (2000).

[3] Something other – something stronger -- than *this* form of critique is, I take it, present in Dotson's *Conceptualizing Epistemic Oppression*, where the focus is not only on re-calibrating the vocabulary of epistemic theory so as to sensitise it to social reality, but on altering our epistemic forms of life -- those 'instituted social imaginaries' which normative social meanings, customs, expectations, assumptions, values, prohibitions, and permissions --the habitus and ethos -- into which people are nurtured from childhood' (Code, 2008). Dotson envisages our altering parts of the world that we use our current epistemic vocabulary to describe, rather than our simply changing the vocabulary.

Epistemic romanticism, then, like its literary counterpart, can come with a range of political resonances. This pluralism is of no help to me: this is to be a solidly *un*romantic thesis. Le Doueff sees 'the distant heirs of Socrates' in the figure of the Shakeseparean Fool and his 'sarcastic and corrosive utterances' (1991). But to the romantic, this thesis, however careful its mappings, may resemble more, Lear's caged insistence that 'we'll pray and sing and tell old tales'. So be it. (At last, a chance for some proper insouciance!)

2. Woolf vs Bennet.

Virginia Woolf (sometimes called a romantic modern) had her first novel, *Jacob's Room,* reviewed by the Edwardian writer Arnold Bennett. He wrote that he had 'seldom read a cleverer book than Virginia Woolf's *Jacob's Room*, a novel which has made a great stir in a small world. It is packed and bursting with originality, and it is exquisitely written. But the characters do not vitally survive in the mind' (1923). Woolf responded to this rebuke with rebukes of her own; but her purpose is not to dispute the centrality of character to the novel. All novels, Woolf says, 'deal with character, and…it is to express character -- not to preach doctrines, sing songs, or celebrate the glories of the British Empire, that the form of the novel, so clumsy, verbose, and undramatic, so rich, elastic, and alive, has been evolved' (1924). (To be sure, critics of bourgeois humanism will find as much to dislike in this breathy account of the novel as they do in analytic epistemology.) Woolf asked her 1924 audience to imagine the situation of the novelist in 1910. They had, she says, 'great difficulty' -- there was no English writer 'from whom they could learn their business'. She herself, she says, had tried asking her elders:

> How shall I begin to describe this woman's character? And they said, 'Begin by saying that her father kept a shop in Harrogate. Ascertain the rent. Ascertain the wages of shop assistants in the year 1878. Discover what her mother died of. Describe cancer. Describe calico. Describe -- ' But I cried: 'Stop! Stop!' And I regret to say that I threw that ugly, that clumsy, that incongruous tool out of the window (1924).

This is parody, of course, but something of the laboured, implacable process Woolf describes is reminiscent of the 'convoluted, ramshackle' definitions so familiar in post-Gettier epistemology (Williamson, 2007). And this approach to character, and how we, the readers, are supposed to get a grip on character, strikes Woolf as very odd. After all, she notes, all of us have 'some skill' in 'the art of character-reading' -- without it 'it would be impossible to live for a year without disaster'. But we don't, when we make these judgements -- even at their most reflective -- rely on the slew of details urged upon us by the ventriloquised Bennett.

I am better off with respect to my elders than Woolf was. Recent epistemology has increasingly sought not after definitions, but to cast light on *structural* features of knowledge, and its theoretical kin.[4] We cast light on the structure of knowledge, or

[4] Epistemologists sometimes riff on the vocabulary of metaphysicians by talking of 'epistemic kinds'; but my talk of epistemic structure ought not be read as importing the ideology of 'joint carving' into epistemology.

possibility, or naturalness, or rationality, or oppression by making explicit the theoretical roles they are to play and getting clear about the internal connections between the different claims we might make about these phenomena. In place of definitions we have what Sider calls 'rich characterization' (2011). The four essays in this thesis share a concern with *structural* features of some core epistemic phenomena, in particular knowledge, rationality, deduction and testimony: the first essay explores the structural commonalities between transmission theses in the epistemology of testimony and closure principles for deduction; the second, the significance of content for the epistemic behaviour of testimony.[5] The third essay argues that we ought to discard the ideology of higher-order defeat, arguing against a tradition that sees the defeasibility of knowledge as central to its structure, and the fourth shows how anti-akratic constraints on rationality generate diachronic instability. In the next few sections, I describe each of the essays in some detail, before marking out some of the ways in which their concerns overlap.

3. Risk, Doubt and Transmission ('Risk').

[5] The former is forthcoming in *Philosophical Studies,* the latter in *Knowledge, Belief, and God: New Insights in Religious Epistemology.*

Risk explores the structural analogies between closure principles and what epistemologists of testimony sometimes called transmission principles or theses.[6] Epistemologists who like closure principles for knowledge -- principles touted by their fans as encoding the thought that 'deduction is a way of extending one's knowledge' -- are invested in thinking knowledge has a certain 'reproductive' structure, or that competent deduction is a way of producing more knowledge. Similarly, epistemologists of testimony who endorse some variant on the claim that tellings, when they go well, transmit knowledge from speaker to hearer, endorse a certain sort of story about what I'm calling the reproductive structure of knowledge. But closure principles, however appealing, are plausibly in tension with other structural claims often made for knowledge -- in particular, with the thought that there are *anti-luck* constraints on knowledge, and the thought that knowledge has a *defeasible* structure. I show that, whatever the tensions between closure principles and anti-luck epistemologies, and between closure principles and the ideology of defeat, we can see these tensions too between transmission principles and, respectively, anti-luck epistemologies and the ideology of defeat. First I show how Lasonen-Aarnio's (2008) claim that there is a

[6] We need to be careful here: irritatingly, 'transmission thesis' is used in two quite different ways within the epistemology of testimony. At times it picks out a very large class of theories that aim to give a necessary and sufficient set of conditions for testimony that *p* from a knowledgeable speaker to produce knowledge that *p* in a hearer; at others, it is used with a much narrower remit, to pick out theories which analyse testimony as a way for speakers to 'pass on' their justification to hearers (see, for example, Wright 2015). It is the former, coarser sense that interests me.

tension between single premise closure and anti-luck constraints on knowledge can be extended to make trouble for transmission theses. Second I show how Schechter's claim that there is a tension between single premise closure and the thought that knowledge is defeasible can be extended to make trouble for transmission theses. This matters, I say, for the trajectory of the epistemology of testimony. As things stand, many, if not most, epistemologists of testimony endorse some sort of transmission thesis. I end the paper by sketching the consequences of this trouble for the dialectic in the epistemology of testimony.

4. Testimonial Pessimism ('Pessimism').

Recent epistemological history has inclined towards a hearty 'testimonial optimism', keen to stress the division of epistemic labour and how ubiquitously we depend upon the words of others. Contemporary gumption has largely replaced older traditions of vigorous gloom:

> The floating of other men's opinions in our brains makes us not one jot the more knowing, though they happen to be true. What in them was science, is in us but opiniatrety; whilst we give up our assent only to reverend names, and do not, as they did, employ our own reason to understand those truths which gave them reputation…In the sciences, every one has so much as he really knows and

> comprehends: What he believes only, and takes upon trust, are but shreds; which however well in the whole piece, make no considerable addition to his stock who gathers them. Such borrowed wealth, like fairy-money, though it were gold in the hand from which he received it, will be but leaves and dust when it comes to use (Locke, 1997).[7]

Dismay at this pitch is unusual in contemporary philosophy, but it finds something like a muted mouthpiece in some of the literature on moral testimony. 'Borrowed wealth' might be all very well when it comes to bank opening times and battle dates, but we must, the thought often goes, be self-made men when working out how we ought to act and live. Sliwa (2012) tells us that much of the literature takes it that there is something defective about basing one's moral beliefs on testimony, Mogensen (2015) that 'although we can happily rely on the word of others when it comes to the geography of India or the mating habits of penguins, it seems we can't be quite so willing to accept testimony on moral questions'. There are other pockets of Lockean despondency: for example, Meskin notes that even realists about aesthetic properties widely accept that 'aesthetic testimony cannot

[7] It would be wrong to conclude that we can draw a simple contrast between contemporary sanguinity and lugubrious early moderns. Locke's own views marry diffuse strains; elsewhere in his Essay he writes 'in things that happen indifferently, as that a bird flies this way or that; that it should thunder on a man's right or left hand, etc. when any particular matter of fact is vouched by the concurrent testimony of unsuspected witnesses, there our assent is unavoidable… a man cannot avoid believing it, and can as little doubt of it, *as he does of the being and actions of his own acquaintance*, whereof he himself is a witness' (my emphasis); Jonathan Barnes in 1980 called 'second-hand' beliefs 'scavengings…a rotten way of acquiring beliefs and…no way at all of acquiring knowledge'.

provide…justification or knowledge' (2004). And in light of traditions that tie mathematical knowledge to the ability to articulate a *proof*, mathematical beliefs based only on testimony can seem to lack the epistemic gilding required for knowledge (Chisholm, 1966, Easwaran, 2011, Williams, 1972). Contemporary 'pessimists', then, defend a cluster of different theses that stress, in different ways, testimony's epistemic shortcomings. They typically focus on particular *kinds* of testimony that look (perhaps systematically, perhaps peculiarly) defective along some important axis of evaluation. In other words, pessimism tends to come in domain-specific flavours. My aim in this second paper is, in part, to suggest a move away from this way of packaging testimony: instead of starting with particular cases (or species) of what look like defective testimony, one can start by casting about for positions in the philosophies of mind and language which *predict* pessimist intuitions. Once one can show that testimonial pessimism looks to flow from a given commitment, one can endorse the position along with its pessimistic upshot, or use the connection to draft an error theory for pessimist intuitions: 'we are tempted by pessimism insofar as we are tempted by this other philosophical position. But the position is wrong, and we should reject it, along with its pessimistic implications'. My project is not to license Lockean gloom, but to establish a theoretically significant connection between pessimist readings of testimony, and two specific commitments one might have in the philosophy of language. I do not aim to say

what I think we ought to *do* with this connection; that is, I aim to remain agnostic on whether we should take the connections I sketch to give us a way of vindicating pessimism, or whether they are better read as part of an error theoretic programme.

I begin by presenting a reading of medieval mystics as proto-pessimists, and show that one reading of these authors has their pessimism flowing from tacit commitments to particular (controversial) views in the philosophy of language: 'emotionism', and what I call 'strong' views of the *de re*. These readings suggest a general connection between pessimist commitments and, respectively, emotionist commitments, and 'strong' readings of *de re* content. I go on to show how these suggestions can do interesting work by sketching the upshot of the proposed connections for moral and aesthetic testimony.

5. Epistemic Ambivalence ('Ambivalence').

Some pictures have our first-order epistemic attitudes generating tight constraints on which higher-order epistemic attitudes we are entitled to, or vice versa. Other pictures have the two operating more discretely. Contemporary epistemology approaches the project of modelling these relations using a number of different guises: the ideologies of epistemic

akrasia (can it be rational to believe that *p* whilst rationally believing that you ought not to believe that P?), improbable knowing (might I know that *p* whilst it is very improbable on my evidence that I know that P?), and *higher-order defeat.*[8] It would, I think, be a mistake to run all these ways of talking together -- the connections between rationality, knowledge and defeasibility are not obvious, and are certainly not uncontested. Nevertheless, we can draw a rough distinction between pictures that excuse dissonance between our first and higher order attitudes, and those that do not, and call them, respectively, *ambivalent* and *integrationist* epistemologies.

In *Ambivalence* I look in detail at the ideology of defeat, and one species of it in particular: higher-order defeat. The paper identifies the ideology of higher-order defeat as one artefact of integrationist epistemology, before arguing that that it is paradox prone, and as such, ought to be abandoned. I end the paper by locating the phenomenon of epistemic ambivalence as continuous with the ambivalence felt in other domains when we consider the *situated* character of various sources of value in our lives, and sketch the prospects for

[8] See, for example, Greco (2014), Horowitz (2014), Lasonen-Aarnio (2014), Littlejohn (2015), Owens (2002), Wedgwood (2011), Williamson (2011).

extending the paper's arguments against the ideology of higher-order defeat to other iterations of integrationist epistemology.

6. Tragedy and Rationality ('Rationality').

We all know that the world is an unfair place. Those who act as they should not often end up better off than those who have done nothing wrong. As Thrasymachus insisted to Socrates, injustice is often more profitable than justice (1981). But fewer of us are convinced that the world has a tragic structure. Tragedies happen when we cannot avoid acting as we must not: Agamemon, a 'previously guiltless man', finds himself in a tragic situation when he must either sacrifice Iphigenea, or leave his army becalmed, and 'both courses involve him in guilt' (Nussbaum, 1986). The tragedian insists that, whatever Agamemnon does, he will have done something he shouldn't. The anti-tragic retort is that, however things appear, there must be some way for Agamemnon to avoid acting as he should not act. Here, then, are the core anti-tragic commitments:

> ANTI-TRAGIC: If you are in a permissible state at t_1, then at t_1, you can avoid acting impermissibly between t_1 and t_2.

FORBIDDEN: If you are in a permissible state at t_1 and an impermissible state at t_2, you have done something impermissible.

Of course, imperatives come in different flavours: there are oughts of rationality, as well as oughts of ethics, and rationality itself is polylithic: it is standard to distinguish between epistemic and prudential rationality. To be an epistemic anti-tragedian is to endorse distinctively epistemic versions of ANTI-TRAGEDY and FORBIDDEN:

ANTI-TRAGIC$_e$: If you are in an epistemically permissible state at t_1, you can, at t_1, avoid acting epistemically impermissibly between t_1 and t_2.

FORBIDDEN$_e$: If you start off in a permissible belief state at t_1 and end up in an impermissible belief state at t_2 then you have done something epistemically impermissible.

The project of this paper is to show, and to make vivid, the tensions between the anti-tragic impulse in epistemology, and the claim that rationality has an anti-akratic structure.

Anti-akratic constraints on rationality, popular with contemporary epistemologists, see dissonance between first and higher-order epistemic attitudes as a species of rational failure. Here is one fairly standard statement of the constraint:

> Anti-Akratic Requirement (AAR): Any rationally permitted state in which I ought to have high confidence that, right now, I ought to have low confidence that p, is a state in which I do not, right now, believe that p.

I argue that there are substantial tensions between anti-akratic and anti-tragic constraints on rationality. This should be a startling conclusion. Both constraints seem natural outgrowths of a picture of rationality as action guiding and tightly bound to the regulation of practical and theoretical reason. Defenders of anti-akratic constraints stress that akratic states license 'bad reasoning and irrational action' (Horowitz, 2013). And if there are cases in which, through no fault of my own, there is no way for me to avoid acting irrationally, it is hard to see how rationality can be action guiding. And Rinard (m.s.) notes that 'one central aspect of rationality is its guidance-giving role. Rationality is an ideal to which we aspire; the requirements of rationality are dictums to which we strive to conform.' But if rationality can

tell me that there is nothing I am permitted to do, I cannot 'coherently aspire to conform to the requirements of rationality'.

7. Connections.

The thesis is most naturally approached as two paired essays: one pair has testimony as its shared focus, the other has integration and ambivalence as its epistemic themes. This geometry can bear a little complication: the transmission theses I put pressure on in *Risk* standardly exploit the ideology of higher-order defeat attacked in *Ambivalence*. This suspicion of higher-order defeat makes a difference to the dialectical force we ought to associate with the case against transmission theses: if higher-order defeat is a suspect piece of ideology, the fact that transmission theses sit badly with its applications is not a straightforward cost for the former. Nevertheless, transmission theorists, when motivating and defending their favoured views, often make liberal and explicit use of the ideology of defeat. Insofar as transmission theses have the appeal they do *because* they can be modified and weakened with defeat clauses, the tension should worry their proponents.

The dialectical situation is a little more complex when it comes to the tensions between transmission theses and anti-luck epistemology. One tempting response is this: 'if the transmission theorist is no worse off than the fan of closure, so much the better for the transmission theorist!' If we hang on to closure principles despite their vulnerabilities, that transmission theses have those same vulnerabilities is no death blow.

These remarks are plausible from a distance, but mottle on closer inspection. The classic safety theoretic response to Lasonen-Aarnio's argument -- the response I am inclined to take most seriously -- appeals to the tightness of the connection between our modal and our epistemic judgements. To think her putative counterexamples to closure are genuine, you need to endorse certain judgements about which cases are close to which other cases. And you can resist doing this if you are happy to be suitably mercenary in your evaluations of closeness. But even mercenaries can have principles; there are systematic differences between deduction and testimony, and one might expect these differences to bleed into our judgements as to which cases count as close to others. Inference, we might think, is a *mental* event; testifying, a *social* one. Testifiers can be spatially and temporally remote from their testifiers; inferences look bound to have a more 'local' character. There may be asymmetries along a different axis too: if the conceptual connections between knowledge and inference

are systematically tighter than those between knowledge and testimony, we might expect to see different response patterns in judgements of closeness made when deductions are at issue compared to when testimony is the focus. We can sharpen this thought by contrasting the following two (rough) claims:

(a) Being inclined (under appropriate circumstances) to use *p* as a premise in one's reasoning (theoretical or practical) is partly constitutive of one's affirming that *p.*

(b) Being inclined (under appropriate circumstances) to testify that *p* is partly constitutive of one's affirming that *p.*

The first claim is significantly more compelling than the second. A disposition to draw certain inferences from *p* can look like an important part of having an attitude of affirmation towards that content. ('If you don't think it's coloured then you don't *really* think it's red.') But it seems one might well believe, even *know* some proposition and yet not be disposed to tell that proposition, even if we idealise away from worries an agent might have about not being believed, thought foolish, or having their credibility. undermined[9] Perhaps one may

[9] A piquant example of how a desire to remain credible regulates our speech: in 1750 Lord Chesterfield wrote to his son, warning him not to acquire the reputation of a teller of marvellous tales: 'Had I really seen anything

know that *p* whilst being systematically unable to articulate that *p* (think of a case in which one knows a *de re* proposition but one's environment contains nothing that could be the right sort of referent of the right sort of demonstrative). Suffering from epistemic injustice or from certain forms of what Langton calls 'illocutionary silencing' may put one in a position in which one is not recognised as, and thereby *fails to be,* a testifier (1993).[10] One's status as believing, or inferring, looks less dependent on one's social position than one's status as a testifier; being able to resist Lasonen-Aarnio's challenge to closure principles does not mean being able to resist my corresponding challenges in the epistemology of testimony.

Of course, there are ways of pushing back against these proposed asymmetries. Those who think seriously about epistemic injustice may propose that there is no sharp line between inner and public life -- that being treated as a testifier is part of what allows us to be believers in the first place.[11] Fricker seems to be suggesting something like this when she says:

---

so very extraordinary as to be almost incredible, I would keep it to myself rather than, by telling it, give any one body room to doubt for one minute of my veracity.'

[10] When an agent experiences illocutionary silencing, they are able to utter words, but they are unable to perform what Austin called an illocutionary speech act, such as warning, promising,consenting, or marrying (Langton, 1993).

[11] I am indebted to Nakul Krishna for conversation on silencing and epistemic injustice.

> A mental state cannot count as a belief at all unless it has a reasonable life expectancy. It must be the sort of thing that one is disposed to assert not only now but in the future too. And Williams's proposal is that engagement in mutually reliant, and so mutually trustful, dialogue with others is the chief impetus for this process by which the mind becomes settled. This is because if my interlocutor asks me a question, and given that I come to the exchange in the frame of mind of someone with an interest in ongoing trust with such an interlocutor (in due course she may be able to tell me something I need to know), then her question calls on me to ask myself how the world is in order that I may answer truthfully. This creates a pressure for me to avoid fantasy in my thinking (most specifically to avoid desires slipping through as beliefs to produce wishful thinking), and to tell her something I believe to be true, thus contributing to the steadying of my mind. We might say, then, that trustful conversation with others is the basic mechanism by which the mind steadies itself. Such dialogue pressurizes the subject into having attitudes of belief towards only those propositions that merit it. It draws the subject away from assertoric caprice and towards doxastic stability (2007).

Naomi Scheman gestures similarly when she tells us that: 'mental phenomena are socially constructed, meaning that their "holding together" as particular states, events, or processes is dependent on their being embedded in a context of social practices' (2011). Presumably, she thinks of testimony as one of these 'social practices'. Edward Craig, albeit with a different set of motivations, has 'suggested a tight connection between knowledge and testimony: the concept of knowledge, as we operate it in everyday practice, is tied to informants' (1999). If these lines of thought can be made good, and the asymmetries gestured towards above turn out to be merely apparent, and the vivid structural analogies between closure and

transmission principles do not turn out to conceal more subtle differences in their mechanics, then transmission principles look more likely to stand or fall along with closure.

The dialectical impulse of *Pessimism* might seem opposed to that of *Risk*. And considered under one aspect, this is so. Pessimist instincts look poised to make trouble for transmission theses -- I cannot, it seems, pass on to you my knowledge that meat-eating is wrong as I can pass on my knowledge that Edinburgh is in Scotland. Such blockages run counter to the spirit of transmission theses. But *Pessimism* does not straightforwardly vindicate the pessimist instinct, and it develops a reading of pessimism compatible with transmission theses. For transmission theses standardly give us conditions under which testimonially based beliefs count as knowledge; *Pessimism* reads the failure of moral (and aesthetic) testimony (if there is one) as a failure to *produce* testimonially based belief.

Other ways of thinking about these two essays see their juxtaposition differently. I have already noted that recent epistemological history has seen a turn away from attempts to analyse knowledge by producing non-trivial lists of necessary and sufficient conditions for some agent A's knowing some proposition *p.* Those who hope, nevertheless, to stake out an important territory for knowledge in epistemology typically seek to do so by arguing that

knowledge systematically structures our practical and epistemic lives. Knowledge has, for example, been proposed as a norm of assertion, action, and practical reason (Hawthorne & Stanley, 2008; Williamson, 2000). But *content* tends to be bracketed or ignored in these discussions, with knowledge being supposed to structure our practices in ways that are insensitive to what *sort* of knowledge is in play. Attempts to systematically embed knowledge in our evaluative practices are undermined if the best treatments of testimonial pessimism appeal to 'domain specific' norms for belief. Hills (2009, 2013) for example, has argued that our suspicion of moral testimonially based belief indicates that understanding, rather than knowledge, lies at the core of moral epistemology; similar views can be found in Hopkins (2007), McGrath (2009), Nickel (2001). *Pessimism* suggests a way of making sense of pessimist instincts that does not mute the theoretical significance of knowledge. Insofar as *Risk* can be read as one more lesson as to the failings of post-Gettier epistemology, *Pessimism* can be read as a defence of knowledge-first approaches.

The interplays between *Ambivalence* and *Rationality,* and their common dialectical orientation, are more obvious. The first introduces the distinction between integrationist and ambivalent epistemologies, and argues against one version of the former. *Rationality* traces some costly consequences associated with integrationist dicta. Towards the end of

Ambivalence, I sketch the prospects for extending my case against the ideology of higher-order defeat so that it speaks against other incarnations of the integrationist project, such as the anti-akratic constraints which are the focus of *Rationality*. Together they push us towards an epistemology friendly to ambivalence.

Preambles are often pleasant -- they allow for elaborate and carefully planted series of hedges, excusing the avenues unexplored and the perspectives unconsidered. At some point, though, the work must be left to fend for itself. I leave mine to do so, now.

References.

Barnes, J., 1980. Socrates and The Jury. *Proceedings of The Aristotelian Society sup. vol.* 54 (1): 193–206.

Bennett, A., 1923. Is the Novel Decaying? *Cassell's Weekly.* 28 March.

Chisholm, R., 1966. *Theory of Knowledge.* Englewood Cliffs, New Jersey: N.J.,Prentice-Hall.

Code, L., 2008. Advocacy, negotiation, and the politics of unknowing. *Southern Journal of Philosophy* 46 (S1): 32–51

~~Coleman~~, N.A. T., 2015. *How Philosophy Was "Whitewashed,* trans. Bowersox, J. & James, D. http://sgrp.typepad.com/sgrp/2015/08/how-philosophy-was-whitewashed-an-interview-with-dr-nathaniel-adam-tobias-coleman-by-aaron-salzer-of-scienceorfat-t.html.

Collini, S., 2012. *What Are Universities For?* London: Penguin Books.

Collins, P. H., 2000. *Black Feminist Thought: Knowledge, Consciousness and The Politics of Empowerment.* New York; London: Routledge.

Craig, E., 1999. *Knowledge in a state of nature.* Oxford: Oxford University Press.

Dotson, K., 2014. Conceptualizing Epistemic Oppression. *Social Epistemology* 28 (2): 115-138.

Easwaran, K., 2009. Probabilistic Proofs and Transferability. *Philosophia Mathematica* 17 (3): 341-362.

Foucault, M., 1994. *The order of things: an archaeology of the human sciences.* New York: Vintage Books.

Fricker, M., 2007. *Epistemic Injustice: Power and the Ethics of Knowing.* Oxford: Oxford University Press.

Gardner, C. V., 2003. *Women philosophers: genre and the boundaries of philosophy.* Oxford: Westview.

Greco, J., 2014. A Puzzle About Epistemic Akrasia. *Philosophical Studies* 167 (2): 201–219.

Harding, S., 1987. *Feminism and Methodology: Social Science Issues.* Bloomington: Indiana University Press.

Hartsock, N., 1987. The Feminist Standpoint: Developing the Ground for a Specifically Feminist Historical Materialism. In eds. Harding, S. & Hintikka, M. B., *The feminist standpoint theory reader: intellectual and political controversies. New York: Routledge, pp. 35–54.*

Hawthorne, J. & Stanley, S., 2008. Knowledge and Action. *Journal of Philosophy* 105 (10): 571–590.

Hills, A., 2009. Moral Testimony and Moral Epistemology. *Ethics* 120 (1): 94–127.

- 2013. Moral Testimony. *Philosophy Compass* 8 (6): 552–559.

Hopkins, R., 2007. What is wrong with moral testimony? *Philosophy and Phenomenological Research* 74 (3): 611–634.

Horowitz, S., 2014. Epistemic Akrasia. *Noûs* 48: 718–744.

Jameson, F., 1975. Magical Narratives: Romance as Genre. *New Literary History* 7 (1): 135–163.

Langton, R., 1993. Speech Acts and Unspeakable Acts. *Philosophy and Public Affairs* 22 (4): 293–330.

Lasonen-Aarnio, M., 2008. Single Premise Deduction and Risk. *Philosophical Studies* 141 (2): 157– 173.

– 2014. Higher Order Evidence and The Limits of Defeat. *Philosophy and Phenomenological Research* 88 (2): 314–345.

Le Doueff, M., 1991. *Hipparchia's Choice,* trans. Selous, T.. New York: Columbia University Press.

Lewis, C. S., 2001. *The Chronicles of Narnia*. London: Collins.

Littlejohn, C., 2015. Stop Making Sense: A Puzzle About Epistemic Normativity. *Philosophy and Phenomenological Research.* doi: 10.1111/phpr.12271.

Locke, J., 1997. *An Essay Concerning Human Understanding.* London: Penguin Books.

Majumdar, R. & McLaurin, A., 1997. *Virginia Woolf: the critical heritage.* London: Routledge.

McGrath, S., 2009. The Puzzle of Pure Moral Deference. *Philosophical Perspectives* 23 (1): 321–344.

Mills, C., 2005. 'Ideal Theory' As Ideology. *Hypatia* 20 (3): 165–184.

Mogensen, A., 2015. Moral Testimony Pessimism and The Uncertain Value of Authenticity. *Philosophy and Phenomenological Research* 92 (3). doi:10.1111/phpr.12255.

Nickel, P. 2001. Moral Testimony and Its Authority. *Ethical Theory and Moral Practice* 4 (3): 253–266.

Owens, D., 2002. Epistemic Akrasia. *The Monist* 85 (3): 381–397.

Rinard, S., m.s. Reasoning One's Way Out of Scepticism. http://philpapers.org/archive/RINROW.pdf

Rorty, R., 1981. From Epistemology to Romance: Cavell on Skepticism. *The Review of Metaphysics* 34 (4): 759–774.

Scheman, N., 2011. *Shifting Ground: Knowledge and Reality, Transgression and Trustworthiness.* Oxford: Oxford University Press.

Schechter, J., 2013. Rational Self-Doubt and the Failure of Closure. *Philosophical Studies* 163 (2): 428-452.

Shakespeare, W., 1994. *King Lear.* London: Penguin Popular Classics.

Shapin, S., 1994. *A Social History of Truth: Civility and Science in Seventeenth Century England*. Chicago and London: University of Chicago Press.

Sider, T., 2011. *Writing The Book of the World.* Oxford: Oxford University Press.

Sliwa, P., 2012. In defense of moral testimony. *Philosophical Studies* 158 (2): 175–195.

Stanhope, P. D., 1775. Earl Of Chesterfield, *The Life of the late Earl of Chesterfield… in a series of letters to his son.* Philadelphia: John Sparhawk.

Walker, A., 1983. *In Search of Our Mother's Gardens*. New York: Harcourt Brace Jovanovich.

Wedgwood, R., 2011. Justified Inference. *Synthese* 189 (2): 1–23.

Williams, B., 1972. Knowledge and Reasons. In ed. von Wright, G. H., *Problems in the Theory of Knowledge.* The Hague pp.1-11.

- 1990. *Descartes: The Project of Pure Enquiry.* London: Penguin.
- 1993. *Morality: An Introduction to Ethics.* Cambridge: Canto.

Williamson, T., 2000. Knowledge and Its Limits. Oxford: Oxford University Press.

- 2007. *The Philosophy of Philosophy.* Oxford: Oxford University Press.
- 2009. Probability and Danger. *The Amherst Lecture in Philosophy* 4: 1–35. <http://www.amherstlecture.org/williamson2009/.
- 2011. Improbable Knowing. In ed. Dougherty, T., *Evidentialism and its Discontents*. Oxford: Oxford University Press, pp. 146-164.

Woolf, V., 1924. *Mr Bennett and Mrs Brown*. London: Hogarth Essays.

Wright, S., 2016. The Transmission of Knowledge and Justification. *Synthese* 193 (1): 293–311.

Abstract - Risk, Doubt and Transmission.

Despite their substantial appeal, closure principles have fallen on hard times. Both anti-luck conditions on knowledge and claims that knowledge is defeasible look to be in tension with natural ways of articulating single-premise closure principles (Lasonen-Aarnio, 2008; Schechter, 2013). The project of this paper is to show that plausible theses in the epistemology of testimony ('transmission theses') face problems structurally identical to those faced by closure principles. First I show how Lasonen-Aarnio's claim that there is a tension between single premise closure and anti-luck constraints on knowledge can be extended to make trouble for transmission theses. Second, I show how Schechter's claim that there is a tension between single premise closure and the thought that knowledge is defeasible can be extended to make trouble for transmission theses. I end the paper by sketching the consequences of this trouble for the dialectic in the epistemology of testimony.

Risk, Doubt and Transmission.

Introduction.

As it is with words, so it is with knowledge: one can *do things* with both. Closure principles tell one sort of story about how knowledge behaves when we use what we know as premises in deductive arguments. In the epistemology of testimony, *transmission theses* (which come in both reductionist and anti-reductionist flavours) tell stories about how knowledge behaves when we tell others what we know.[12] Despite their substantial appeal, closure principles have fallen on hard times: both anti-luck conditions on knowledge and the defeasibility of knowledge look to be in tension with natural ways of formulating closure principles (Lasonen-Aarnio, 2008; Schechter, 2013). The project of this paper is to show that transmission theses in the epistemology of testimony face problems structurally similar to those faced by closure principles.

[12] I use here the lexicon of reductionism and anti-reductionism. Alternative lexicons (as opposed to distinctive taxonomies) have been proposed in which, for example, 'fundamentalism' takes the place of 'anti-reductionism' (Graham, 2006). I prefer to avoid any suggestion of religiosity.

Lasonen-Aarnio has argued that there is a tension between single premise closure principles and ANTI- LUCK:

(ANTI-LUCK) If I am sufficiently lucky to have a true belief that *p*, then I do not know that *p.*[13]

Schechter has argued that there is a tension between single premise closure and DEFEAT:

(DEFEAT) A subject's knowledge that *p* can be lost when misleading evidence is acquired, even where the subject persists in believing that *p.*

In the first part of this paper I argue that the tension charted by Lasonen-Aarnio between single premise closure and ANTI-LUCK is replicated in the epistemology of testimony. Second, I show that the tension charted by Schechter between single premise closure and DEFEAT is also replicated in the epistemology of testimony.

1.1 Preliminaries.

Suppose we begin with the optimistic thought that deductions, when they go well, extend our knowledge.[14] Fleshed out a little, the thought is this: deduction, *when everything goes*

[13] Lasonen-Aarnio does not exploit the ideology of 'belief forming methods' in her statement of ANTI-LUCK. Neither Lasonen-Aarnio's arguments nor my own depend on the omission of this ideology.

well with it, is an operation which takes us from tokens of knowledge to tokens of knowledge. This thought will only count as non-trivial if we can say something more about what it is for a deduction to go well than just that deductions go well when they have a knowledge token as their output. But we do seem to be able to say a bit more than this. We seem, for example, to have some appropriately independent grasp of the distinction between competent and incompetent deductions, and we can tie the notion of a deduction going well to the idea of a deduction being competent. Even so, we can tell more or less demanding stories about what counts as a deduction going well. The closure principle I will be working with in this paper (modelled on Schechter's) tells a relatively demanding story:

Single-premise closure (SPC):

Necessarily, if S

(i) knows that *p*,

(ii) comes to believe that *q* solely on the basis of competent deduction from *p*,

[14] If endorsing a closure principle is being optimistic, it is the sort of optimism that comes in degrees: some closure principles make more modest claims than others. Both Williamson and Hawthorne formulate more optimistic theses about deduction than I do here -- their preferred formulations make no reference to defeat (Hawthorne, 2003; Williamson, 2011). Note that to refuse to endorse a closure principle is not to deny the epistemic significance and interest of deduction. One might, for instance, think that *generic* statements about deduction capture the procedure's behaviour more aptly than universally quantified claims with attached necessity operators.

(iii) retains knowledge that *p* throughout and,

(iv) S does not have a defeater for the claim that the deduction was competent,

then S knows that *q*.

Friends of closure principles claim that they codify the intuitive thought that deduction is a way of extending one's knowledge (Williamson, 2000). Transmission theses in the epistemology of testimony look to codify another intuitive thought: that tellings, when they go well, are a way of *sharing* one's knowledge.

Just as we might be lax or strict about what counts as a deduction going well, we can be lax or strict about what counts as a telling going well. Reductionists and anti-reductionists – at least on natural ways of glossing their dispute – disagree as to how strict we ought to be about what counts as a telling going well. But the purpose of this paper is not to worry about how to best distinguish reductionist and anti-reductionist claims. Rather, I am interested in a certain structure that is shared by typical statements of reductionism and anti-reductionism. Typically, such statements (purport to) give us non-trivial, non-circular

characterisations of circumstances under which accepting a speaker's testimony will be knowledge producing.[15] Call theses with this general form *transmission theses*.[16]

For the sake of sharpness I will begin by tracing the interaction of ANTI-LUCK and DEFEAT with one transmission thesis. I will then argue that we ought to expect similarly structured interactions regardless of which transmission thesis we go for. The statement of transmission I'll be working with is relatively strict about what counts as a telling going well:

Knowledge-Transmission (KT):

Necessarily, if

(i) S knows that *p,*

(ii) A comes to believe that *p* solely on the basis of competent understanding of S's testimony that *p,* and

[15] More precisely, they have the form:

> Necessarily, if S knows that *p* and A comes to believe that *p* solely on the basis of competent understanding of S's testimony that *p*, and θ obtains, then A knows that *p*;

where we have some grip on what it is for θ to obtain that is suitably independent of our grip on what it is for acceptance of testimony to be knowledge producing.

[16] Some people use 'transmission thesis' in a much narrower way than I do, to pick out, say, just theses which deny that accepting testimony from an ignorant speaker can be knowledge producing. See, for example, Faulkner (2011), Wright (2015). Note too that given my account of transmission theses, theories which claim to be neither reductionist nor anti-reductionist count as transmission theses. See, for example, Lackey (2008), Faulkner (2011).

(iii) A has good, undefeated evidence (a) that *A's understanding was competent* and (b) that *S is a reliable testifier,*[17]

[17] It has been suggested that condition (iii) builds in a quasi-inferential structure to the epistemology of testimony. If this is so, it ought not to be surprising that the arguments which Lasonen-Aarnio and Schechter develop with inferential processes in mind can be applied to the epistemology of testimony. Of course, it *might* be the case that the epistemology of testimony has an inferential or quasi-inferential structure, but if the arguments of this paper depend upon importing a quasi-inferential structure into the epistemology of testimony, the arguments will lack the bite they purport to have. However, it is not the case that endorsing condition (iii) commits one to the thought that the epistemology of testimony has a quasi-inferential structure. All condition (iii) requires is that a hearer *be in possession of a certain kind of evidence*; it does not say anything about the *role* this evidence is to play. In particular, one can think that taking a speaker's word for it that *p* can be knowledge producing only when certain propositions are in one's evidence without thinking that those propositions are relied upon as premises, or as something premise-like, by the hearer when they form their belief that p. An example might help make this point vivid. Suppose one thinks that the epistemology of perception does not have an inferential structure: when I see a goldfinch, and come to believe on the basis of this goldfinch-appearance that a goldfinch is over there, the process whose output is the belief that there is a goldfinch does not, in any interesting or important way, map onto inferential processes. With this in mind, consider the following case, BIRDWATCHING:

> Ella enjoys birdwatching, In order to birdwatch, Ella relies on a pair of binoculars. If these binoculars are not cleaned regularly, they produce very misleading visual impressions. Ella often forgets whether or not she has cleaned the binoculars, but she keeps a note, in her diary, of when she last cleaned the binoculars: so long as they have been cleaned in the last week, the binoculars are perfectly reliable, On Saturday, Ella goes birdwatching, and spots a goldfinch through her binoculars. 'A goldfinch is over there', she thinks to herself.

Plausibly, Ella *knows* that there is a goldfinch only if she has checked in her diary that she cleaned the binoculars within the last week. That is, Ella appears to need *positive evidence* that her binoculars are not misleading if she is to come to know, on the basis of her visual impression, that *a goldfinch is over there*. But one might be in favour of this constraint, and think that binocular aided perception is no more plausible a candidate for having a quasi-inferential structure than perception unaided by binoculars. Imposing the constraint that Ella have checked when she last cleaned her binoculars does not force us to think that when Ella is looking through the binoculars, she relies, however tacitly, on the claim that her binoculars are clean as a premise in an inference to the conclusion that there is a goldfinch. We can think that perception is just as 'direct' when aided by binoculars as when it is not so aided, but that there are additional background constraints that must be satisfied if said direct process is to be knowledge producing where Ella is relying on binoculars. Similarly, one can think that that taking a speaker's word for it that *p* can be knowledge producing only when certain propositions are in one's evidence without thinking that those propositions are relied upon,

then A knows that *p*.[18]

Two points should be made about this formulation. First, competent understanding of a telling requires a hearer to appreciate the illocutionary force of the speech act. Second, exactly how best to understand what (who) counts as a reliable testifier is contested.[19] For the purposes of this paper, I will take a speaker to be reliable at some context of utterance *c* iff in *c* they have a disposition to tell only true propositions, and this disposition is not easily *finked* or *masked* (see Bird, 1998 and Lewis, 1997).[20]

Before making my argument, I want to say a little about its dialectical upshot. It is not clear what friends of closure ought to make of the objections posed by Lasonen-Aarnio and

---

even tacitly, as the premises or quasi-premises in an inference to p. Thanks to Ralph Wedgwood and Timothy Williamson for making these points salient to me.

[18] I have been talking about knowledge and knowledge producing processes. Both ANTI-LUCK and DEFEAT can be formulated as theses concerning justification, rather than knowledge. Similarly, both transmission theses and closure principles can be formulated as theses about justification, as well as articulated in terms of knowledge. In this paper I will be talking primarily about the interaction between closure principles and transmission theses formulated as theses about knowledge. But each of the arguments considered here can be run, *mutatis mutandis*, as an argument concerning the justification-centred incarnations of these theses.

[19] For discussion of how to construe speaker reliability see Fricker (2015), Goldberg (2008), Lackey (2015).

[20] Such reliability may or may not be rooted in a disposition to refrain from forming false beliefs. As Lackey has pointed out, a speaker may be a reliable testifier with respect to whether or not *p* even when their own beliefs as to whether or not *p* are hopelessly mistaken. In her case CREATIONIST TEACHER, she envisages a testifier whose testimony in the classroom is guided by the scientific consensus concerning the theory of evolution, but whose own beliefs are not so guided (Lackey, 2008).

Schechter. Similarly, it is not obvious what epistemologists of testimony ought to make of the structurally analogous objections presented here.

There might be ways of reworking closure principles so that they avoid the problems posed by Lasonen-Aarnio and Schechter without losing their theoretical power and interest. Similarly, there may be ways of reworking transmission theses so that they retain their significance whilst avoiding the problems mapped in this paper. Working out whether or not this is so is beyond the scope of this paper. My project is to point out that there are significant pressures on transmission theses that are not noted in the literature, rather than looking in detail at what happens to the dialectic once this pressure is applied.

2.1 Anti-Luck Epistemology.

Beliefs can be risky in two different ways. One kind of risk corresponds to the objective chance of a believed proposition being false. Suppose I believe that Hillary Clinton will win the Democratic nomination. This belief might both *be true* and be *such that there is an objective chance of 0.2 of Hillary Clinton not winning the Democratic nomination.* Another kind of risk that attaches to a belief corresponds to the objective chance that something will go wrong in the process of my acquiring it. Suppose at $_{t1}$ I embark on the process of deducing that *p & q* from the known proposition *p* and the known proposition *q*.

At $_{t1}$ it is true that I shall perform a competent deduction. This is consistent with there being, at $_{t1}$, a small objective chance that I shall deduce something from the premises *p*, *q*, that they *do not entail* (Lasonen-Aarnio, 2008).

It is a familiar point that worries about the first sort of risk are enough to make trouble for multi-premise closure principles. It has sometimes been thought that single premise closure is immune from worries about risk accumulation, because where *p* entails *q*, $P(q) \geq P(p)$. But we have distinguished between two sorts of risk that beliefs may be subject to. Lasonen-Aarnio exploits this distinction to argue that single premise closure principles are also vulnerable to worries about risk accumulation. After giving an account of her argument, I will show how to generate a structurally similar argument against transmission theses.

ANTI-LUCK can be unpacked in different ways. For the sake of vivacity, Lasonen-Aarnio begins by showing us what happens if we adopt the *most-worlds safety* (MWS) reading of ANTI-LUCK, before showing that the argument generalises:

MWS: A knows that *p* only if A has no relevant false beliefs in at least n% of close worlds.

For the sake of argument, let $n = 90$.

Key to Lasonen-Aarnio's case is the point that deductions, in order to count as competent, do not 'require infallibility, or zero chance of error' (2008). As she notes, if only infallible deductions were competent, actual agents would not be in a position to perform competent deductions (2008). She proposes a weaker constraint on competence: that A competently deduces *q* from *p* just in case:

(i) *p* entails *q,*

(ii) A's actual deduction is successful and

(iii) in at least m% of the close worlds in which A deduces a relevant proposition *q** from a relevant proposition *p**, a conditional with *p** as its antecedent and *q** as its consequent is true.

For the sake of argument, let m (like n) = 90.[21]

Before rehearsing Lasonen-Aarnio's argument, we need to say something about the relationship between competent deduction and falsity. It is possible to competently deduce something true from a false proposition: I might competently deduce (the true proposition) that I am less than eight feet tall from the false proposition that I am seven feet tall. Given

[21] Given that Lasonen-Aarnio proposes that a deduction of *q* from *p* is competent 'just in case' conditions (i) - (iii) are met, it is most natural to read her as proposing (i) - (iii) as necessary *and* sufficient conditions rather than as merely necessary conditions.

that there are no close worlds in which I am eight or more feet tall, it is also possible for competent deduction from a false belief to give me a *safely* true belief. Similarly, it is possible to incompetently deduce a safely true proposition. Suppose I believe (truly) that some mammals exist. Suppose, too, that I am inclined to deduce the conjunction that some mammals exist and some fish exist from the proposition that some mammals exist. This is certainly not a competent deduction: that some mammals exist does not entail that mammals *and* fish exist. But the conjunction is safely true: worlds where there are no fish are far away. These are not complications that Lasonen-Aarnio ignores. She does not (falsely) take it that incompetent deductions give rise to only false beliefs. Rather, she carefully engineers her cases so that all the incompetent deductions available to an agent give rise to false beliefs. Lasonen-Aarnio gives us the following case:

> BASKETBALL:
>
> I am applying to join a basketball team, and I am asked which of the following intervals my height will fall into at t_3: [150 cm, 151 cm] or [151 cm, 152 cm] or …. or [219 cm, 220 cm]. I decide to perform a deduction to answer the question. I know that, right now, at $_{t1}$, to the nearest tenth of a centimetre, I am 175.9 cm tall. In other words, I know now that at $_{t1}$ my height is between 175.85 cm and 175.94

cm. Suppose that at $_{t1}$ I am given a hormone such that there is a 0.1 chance that the hormone will take effect before t_3. If it takes effect it will cause me to grow at least one centimetre between $_{t1}$ and t_3. If the hormone doesn't take effect, then my height will not change between $_{t1}$ and t_3. In fact, it will not take effect.

Let us set things up so that our schematic letters have the following values:

p: My height at t_3 will be between 175.85 cm and 175.94cm.

q: My height at t_3 will be in the interval [175cm, 176cm].

*q**: My height at t_3 will be in the interval [174cm, 175cm].

Suppose that the objective chance of 0.1 at $_{t1}$ that the hormone will take effect is sufficiently low for me to know at $_{t1}$ that *p*. Suppose I competently deduce from *p* that *q*: in 90% of close worlds, including in the actual world, I deduce that q from *p*. But in 10% of close worlds I deduce *q** instead. Thus, I deduce something from *p* in every close world. For me to know at $_{t1}$ that q, it must be that in at least 90% of close worlds I have no relevant false belief. There being a 0.1 chance that my height will be at least one centimetre greater than 175cm at t_3 means that there is a 0.1 chance that my height will not be in the interval [175cm, 176cm] at t_3. Assume with Lasonen-Aarnio, that this means that in exactly 90% of

close worlds my height at t_3 is in the interval [175cm, 176cm] and in 10% of close worlds in a different interval. Does this mean that I have a relevant false belief in no more than 10% of close worlds? No: 10% of the worlds in which I make a deduction are worlds in which my deduction does not go well and I deduce from *p* some proposition *q** such that condition (iii) is not met. Lasonen-Aarnio argues that unless the set of worlds in which my deduction goes badly is a subset of the set of worlds in which the hormone takes effect before t_3, the proportion of close worlds in which I have a relevant false belief will be 19%. Only worlds in which my belief that *p* is true and my inference is successful will be worlds in which the proposition I deduce is true:

	Truly believes p	Falsely believes p
Inference good	81% of close worlds	9% of close worlds
Inference bad	9% of close worlds	1% of close worlds

Lasonen-Aarnio notes that the general structure of her argument does not depend on reading off MWS from the ANTI-LUCK intuition. Here is her thinking. Let *risk* be a placeholder for something that knowledge is only tolerant to in small quantities. Any view

that makes knowledge tolerant to some non- zero degree of risk whilst setting *some* limit to how risky knowledge can be will, in her words 'struggle' with SPC. However we spell out risk and its connections with knowledge and deduction, if beliefs that are a little risky can be knowledge constituting, and deductions can be competent even when they are a little risky, these two sorts of risk will look prone to aggregate when we competently deduce a proposition from a known premise, in such a way as to make trouble for SPC.

One thing that will vary with these different ways of cashing out ANTI-LUCK is the best way to make precise what counts as a competent deduction. Recall the framework used in tandem with the MWS reading: A competently deduces q from p just in case (i) p entails q (ii) A's actual deduction is successful and (iii) in at least m% of the close worlds in which A deduces a relevant proposition q^* from a relevant proposition p^*, a conditional with p^* as its antecedent and q^* as its consequent is true. For a reliabilist framework, Lasonen-Aarnio replaces (iii) with (iii'):

> (iii'): in a high ratio of cases in which A deduces a relevant proposition q^* from a relevant proposition p^*, a conditional with p^* as its antecedent and q^* as its consequent is true.

Suppose I incompetently deduce *q* from the known premise p, where *p* is false in some close worlds. Lasonen-Aarnio envisages the worlds in which the deduction goes badly being evenly distributed across worlds in which *p* (and my belief that p) is true and worlds in which *p* (and my belief that p) is false. But if what it is for a deduction from *p* to go badly is for me to deduce some *q** such that it is false that [not-p or q*] then a deduction from *p* cannot go badly in worlds in which *p* is false. If the proposition *if p then q** is true whenever *p* is false, then any inference from *p* in a world in which *p* is false will count as a good inference. And in that case, in a case like BASKETBALL, all the bad inferences will be in worlds in which *p* is true and only 80% of close worlds will be close worlds in which I truly believe that *p* and my inference is successful:

	Truly believes p	Falsely believes p
Inference good	80% of close worlds	10% of close worlds
Inference bad	10% of close worlds	0% of close worlds

The point is not that if the consequences I have outlined really are consequences of Lasonen-Aarnio's proposed constraint on competent deduction, her argument fails to work.

Even if they are consequences of her proposed constraint, risk will accumulate so as to make trouble for SPC. Rather, one might think that the conclusion we ought to draw from the above, is that there is a certain triviality built into her proposed account of competent deduction when it is glossed this way: the proposed condition is too easily met to do the work of capturing what counts as a competent deduction. This concern replicates even if we gloss ANTI-LUCK in, say, reliabilist terms rather than using the ideology of safety. Suppose, as Lasonen-Aarnio proposes, that a reliabilist gloss of competent deduction of *q* from *p* will require a high ratio of cases in which some relevant proposition *q** is inferred from some relevant proposition *p** to be such that in those cases the material conditional *if p* then q** is true. Such a constraint will be met whenever *p** is false in a high ratio of cases, arguably trivialising it.[22]

There is some wiggle room for Lasonen-Aarnio, who suggests that '[t]he conditional [of condition (iii)]…could either be a logical entailment or a material implication' (2008). One reason for preferring the latter reading, she says, is that it would let us give a unified account of deductive inference and inference more generally. But maybe we ought to prefer the former reading: the problem above -- if it is one -- can be avoided by having the

[22] I take myself to be making good on what is noted in passing by Hawthorne and Lasonen-Aarnio, that worries about objective chance that arise for closure principles also arise for testimonial (and preservative memory) transmission principles (Hawthorne and Lasonen-Aarnio, 2009).

conditional be an entailment rather than a material conditional. If this means we lack the prospect of a unified account of inference, so much the worse for unified accounts. Importantly for my purposes, whether or not we can resolve these puzzles about how to understand competent deduction does not affect the prospects for generating a case against transmission theories. The trouble (if there is any) is generated by the conditionals of (iii) and (iii'). But we don't – as I show later – need to use a conditional when proposing an account of competent speech-understanding.

2.4 Testimony and Anti-Luck Epistemology

My project here is not to say that we should be persuaded by this case against SPC, or to trace its consequences for other ways of articulating closure principles. My project is to argue that if we are persuaded by this case against SPC, we ought also to reject KT.[10] To see this, we must get more precise about what counts as competent understanding. If it is unduly restrictive to require that competent deductions be infallible, it is unduly restrictive to require that competent speech-understandings be infallible.

Suppose we apply to speech-understanding the same standards and epistemic framework as Lasonen-Aarnio applies to deduction. Say A competently understands S's testimony that *p* only if (i) *in fact* A's interpretative process was successful and (ii) in m% of close worlds in

which A interprets S as having testified to some relevant proposition p^*, S did testify to some relevant proposition p^*, where m=90. Lasonen-Aarnio argues that the objective risk of a proposition's falsity and the risk of performing a faulty deduction can combine in such a way as to prevent deductions from being knowledge producing. I argue that objective risk and the risk of faulty speech-understanding of testimony can similarly combine in such a way as to prevent acceptance of testimony being knowledge producing. Consider the following case:

ASTRONAUT:

Suppose that NASA have just launched a one-man mission to Mars. If the mission is successful, Caroline Wood, the astronaut on board, will become the first person to land on Mars. The mission has a 0.9 chance of success, but if it fails, then the astronaut on board will die well before they reach Mars. Suppose the president asks the administrator of NASA what the full name of the first person to land on Mars will be. The administrator of Nasa tells him a proposition that he knows: that the first person to land on Mars will be an astronaut named Caroline Wood. (I suppose here, as we supposed above, that propositions with a 0.1 chance of falsity can be known.) The president competently understands the administrator, and has good evidence that his own understanding is competent and that the administrator is a reliable testifier.

If KT is true, the president knows that the first person to land on Mars will be an astronaut named Caroline Wood. But that the president competently interpreted the administrator's testimony is consistent with there being *some* close worlds in which the administrator tells the president that the first person to land on Mars will be an astronaut named Caroline

Wood and the president interprets the administrator as having told him (the false proposition) that the first person to land on Mars will be an astronaut named Carl Linwood.

Just as we had to say something about the relationship between competent deduction and falsity, we must say something about the relationship between competent speech-understanding and falsity. I may be told that p, and misinterpret the telling, taking myself to have been told that p^*. p^* may be a true proposition, whether or not p is a true proposition. But in ASTRONAUT, that the president has asked the administrator for the *full name* of the astronaut on the mission constraints which misinterpretations are available to the president. If he takes the administrator to have told him that the first person to land on Mars is named something other than Caroline Wood, the proposition he takes the administrator to be telling him is false.

Let us set things up so our schematic letters have the following values:

p: The first person to land on Mars will be an astronaut named Caroline Wood.

p^*: The first person to land on Mars will be an astronaut named Carl Linwood.

Assuming the same relationship between objective chance and modal space as that assumed by Lasonen-Aarnio, in 90% of close worlds in which the administrator tells the president

that *p*, *p.* Suppose too that we make the simplifying assumption that S tells A that *p*, and that A accepts the proposition that they *take* themselves to have been told by S, in all close worlds. (This is the analogue of the assumption made in BASKETBALL, that I deduce something from *p* in every close world.) Were there no chance at all of A failing to competently understand S's testimony then in at least 90% of close worlds, A would avoid relevant false belief. But it is compatible with A's understanding of the testimony being competent that in 10% of close worlds, A *misinterprets* S's testimony. If worlds in which A misinterprets S's testimony are uniformly distributed across worlds in which *p* and worlds in which *not-p,* A will avoid relevant false belief only in 81% of close worlds:

	Speaker truly believes p	Speaker falsely believes p
Audience understanding good	81% of close worlds	9% of close worlds
Audience understanding bad	9% of close worlds	1% of close worlds

Thus we have a case in which KT predicts that the president knows the proposition that the first person to land on Mars will be an astronaut named Caroline Wood, but ANTI-LUCK

requires that the president not know this proposition, as he fails to avoid relevant false belief in sufficiently many worlds.

It should be clear that, just as Lasonen-Aarnio's case does not depend on the anti-luck intuition being cashed out as a most-worlds safety view, nor does this parallel case against KT. Once again, let risk be a placeholder for something that knowledge is tolerant to only at minimal levels. For the non-sceptic, speakers can, presumably, know and tell audiences propositions that are a little risky. And competent understandings of speakers can also be a little risky. So long as these two sources of risk can aggregate, then KT will be in trouble.

2.5 Generalising Again.

We have a case against KT. But what I want is something more general: a case against *transmission theses*; not just one example of a transmission thesis. But there are the makings of a general problem here.

It should be obvious that weakening KT so that a hearer doesn't need good undefeated evidence that a speaker is reliable will not be a way of avoiding the sorts of worries sketched above. But what if we *strengthen* KT? I cannot show that there is *no* way of adjusting KT so that it meshes with ANTI-LUCK whilst retaining its theoretical interest. As I said above,

the project of this paper is to direct attention towards an unnoticed pressure on transmission theories, not to look in detail at the aftermath of its application. What I can do is look at the resources provided by the standard ideologies embedded in the epistemology of testimony literature, and show that they don't offer us a satisfactory way of reformulating KT.

As it stands, KT requires that a hearer have good undefeated evidence of a speaker's reliability. A natural way of strengthening KT is to require, say, that hearers *know* or be *certain of* a speaker's reliability. But if we are ever to have testimonially based knowledge that *p* where there is some objective chance of p's being false, then we need either (i) an epistemology of testimony that falls short of requiring knowledge or certainty of a speaker's reliability, or (ii) an epistemology of testimony that does not rule out our knowing or being certain that a speaker who tells us propositions which have some non-zero objective chance of being false is a reliable speaker. The alternative is a theory that rules out any testimonially based knowledge about the future. But so long as we have a theory that allows me to be certain of a speaker's reliability, even where he tells me some proposition that concerns the future, it seems that we will have cases in which risk aggregates in such a way as to frustrate transmission theorists.

The general structure of the problem is this. Any constraint we propose must either be such that it can be met in a case where a hearer is told a true proposition with a non-zero chance of falsity, or such that it cannot. If the latter, then we have a theory that is too strong (it seems to rule out any testimonially based knowledge of the future). If the former, then the constraint looks badly placed to rule out problematic cases of risk aggregation.

Take instead the suggestion that we require hearers to be sure that speakers are sincere, or that hearers have no defeater for the proposition they are told. So long as I can be sure that speakers who tell me propositions about the future are sincere -- as any decent account of sincerity must allow -- sincerity-centred constraints will fail to dispel the prospect of risk aggregation. Similarly, unless I have a defeater for *any* risky proposition, it seems that we will still get cases in which transmission theses will predict my having testimonially based knowledge whilst ANTI-LUCK rules it out.

Similar problems will afflict any attempt at fixes that invoke the ideology of competent understanding. For any way of spelling out what counts as a 'good enough' epistemic position (knowledge, certainty) with respect to the claim that one has competently understood a given speech-act, being in this sort of position will either require that a given interpretation was not at all risky, or it will not come with this constraint built-in. If it does

the former, then it will be much too strict: only risk free interpretations will get to be knowledge producing. But if it does the latter, then the structure of the initial problem will remain intact.

The generality of the problem can be made particularly vivid by thinking through the implications of requiring hearers who have been told *p* to know that they have been told *p* by a knowledgeable speaker. Let *q* be the proposition that S knows *q**. Suppose that I know that q, where both *q* and *q** are false in exactly 10% of close worlds. In this 10% of close worlds, S nevertheless believes and tells me that *q**. In the actual world, if I know that *q*, I am presumably in a position to competently deduce that *q**, and thus, if we like SPC, to acquire knowledge of the proposition testified to on the basis of deduction. But so long as we allow for the fallibility of competent deduction, I might competently deduce that *q** from the known premise that *q* despite there being some close worlds in which my deduction is unsuccessful. And if this is so, there are two different sorts of risk attached to my eventual belief that *q**: the risk that *q* is false (as it is in some close worlds) and the risk that my deduction goes wrong (as it does in some close worlds). And these two sorts of risk can aggregate. Unless we have some way of reconciling SPC with ANTI-LUCK, even

coming to believe a told proposition on the basis of deduction, rather than on the basis of mere acceptance of another's testimony, may fail to be knowledge-producing.

To reiterate: there may be ways of adjusting KT that manage what these suggestions cannot. Nevertheless, it ought to be clear that tensions between ANTI-LUCK and KT are not the consequence of some idiosyncrasy of the latter's formulation. If we are persuaded that there is a tension between KT and ANTI- LUCK, we should be cautiously pessimistic about the prospect of coming up with a transmission thesis that is not in tension with ANTI-LUCK.

3.1 Defeat Epistemology.

To think that knowledge is defeasible is not just to think that knowledge can be lost, or even just to think that knowledge can be lost in particular ways; say, as the result of acquiring misleading evidence. Here is a very trivial way in which acquiring misleading evidence might rob me of knowledge that *p*: *acquiring misleading evidence leads me to stop believing that p.* What I am calling *defeat epistemology* does more than just countenance this possibility. Here are some plausible thoughts:

(1) To count as believing that p, there is some non-zero threshold over which one's degree of confidence that *p* must be.

(2) What one's degree of confidence *ought* to be depends on one's evidence.

(3) Whether one counts as knowing that *p* depends (partly) on whether one's degree of confidence that *p* is appropriate given one's evidence.

And here is an observation: an agent's evidence may vary diachronically. Thus, at different times, different degrees of confidence that *p* may be appropriate. Just as I can acquire new evidence that speaks in p's favour, so I can (or so the orthodox position has it) acquire new evidence that speaks against *p.* Combined with thoughts (1)-(3) this observation gives us a

picture according to which knowledge that *p* may be lost when I acquire new evidence, *regardless of whether I respond to this new evidence by altering my degree of confidence that p.* A defeat epistemology is any epistemology that can accommodate this phenomenon. The project of this part of the paper is to show that any tension between SPC and DEFEAT is replicated in the interaction of KT and DEFEAT.[23]

[23] Note that the plausibility of thoughts (1) – (3) is not independent of one's views on evidence. For example, if one thinks that an agent's evidence is just those propositions that said agent knows (that is, that E=K) then whenever *p* is a proposition that I know, the epistemic probability that *p* will (on my evidence) be 1. Where *q* is some proposition that is entailed by a proposition that I know, the epistemic probability that *q* will (on my evidence) be 1 (Williamson, 2000). If we go for this sort of story about evidence, then there may well turn out to be cases where I ought (or so it seems) to have a low degree of confidence in propositions with an epistemic probability of 1. And such a picture does not sit easily with (2). Whether and how such an account might model an agent's encountering evidence against propositions that they know is not clear: if E=K and I know that *p* then there is no proposition *q* such that P(p/p&q) <1. (Williamson, 2000). If these sorts of considerations are enough to make you unsympathetic to the sort of picture codified in (1) - (3) much of the tension charted in this paper will fail to vex you. But even disinterested observers need not be uninterested. Note that puzzles with the same sort of structure as those produced by the Williamsonian picture do emerge on other, more standard views of evidence, such as those that assign probability 1 to all logical truths: it seems plausible that there are circumstances in which I ought to have a low degree of confidence in some logical truth, and if *p** is a logical truth, where P(*p**) = 1 there will be no proposition *q** such that P(*p**&q*) < 1. (I thank an anonymous reviewer for this point.) Nevertheless, the Williamsonian picture seems to be in a deeper tension with defeat epistemology than those views which assign probability 1 to logical truths. Endorsing the latter view does not stop us from thinking that there are lots of cases in which I know propositions whose probability is less than 1. And if *that* is the case then, so long as we assume there is some threshold such that the probability of *p* on my evidence needs to be greater than *n* (where 0<n<1) for me to know that p, we can easily model the defeasibility of knowledge of propositions that don't have probability 1. But for the Williamsonian there are no known propositions with a probability of less than 1.

3.2 Putting defeat epistemology to work.

Schechter argues that DEFEAT makes trouble for single premise closure principles using the following case:

> SCHECHTER: Consider a very long sequence of competently performed simple single premise deductions, where the conclusion of one deduction is the premise of the next. Suppose that Schechter knows the initial premise, but has no other evidence about the intermediate or final conclusions. Suppose that he comes to believe the conclusion, q, (to a very high degree) solely on the basis of going through the long sequence of deductions. He should – assuming he realises that he has performed a long sequence of deductions – think it likely that he has made a mistake somewhere in his reasoning (Schechter, 2013).[24]

Just as one might have very good evidence that some of one's beliefs are false (whilst having just as much evidence for each proposition one believes, as for any other proposition one believes) one might have very good evidence that some of the single premise deductions one performs go wrong, whilst having just as much evidence that one has gone wrong with

[24] Schechter runs his argument in terms of justification rather than knowledge. I have modified his cases so as to reflect this paper's focus on knowledge.

respect to *this* single premise deduction, as one has that one has gone wrong with respect to any other. In cases like SCHECHTER one has very good evidence that one has made a mistake somewhere in one's reasoning. If Schechter's degree of confidence in the conclusion is sufficiently high as to allow for belief in the conclusion, his belief will fail to count as knowledge, for his degree of confidence will be inappropriately high given his evidence (Schechter, 2013).

My project here is not to say that we ought to be persuaded by Schechter's case. Rather, I argue that if we are persuaded by his argument against SPC, then we ought to reject KT as well. For just as we can get long sequences of deductions, we can get long testimonial chains:

TESTIMONIAL CHAIN:

There is a testimonial chain consisting of speakers S_1 through S_n. S_1 knows that *p.* S_1 tells S_2 that *p* and S_2 tells S_3 that p, and so on, until S_{n-1} tells S_n that *p.* In each case, the hearer *in fact* competently understands the speaker. Each member of the chain (save S_1) comes to believe that *p* based solely the testimony of the preceding member of the chain, and has good evidence both that they are a competent interpreter of

speech and that the preceding member of the chain is a reliable testifier. Let S_n's knowledge of his situation be as follows. He knows:

(i) that he is situated at the end of a long testimonial chain consisting of *n* members,

(ii) that whatever proposition S_1 told S_2, S_1 knew that proposition at the time of the telling,

(iii) that for all *x* where $x>1$, S_x knows the proposition they tell S_{x+1} only if it is the proposition that S_1 told S_2,

(iv) that for all *x* where $x>1$, S_x believes the proposition that they tell S_{x+1} solely on the basis of S_{x-1}'s speech.

Note that this last condition means that for all x , where x>1, if S_x knows the proposition that they tell S_{x+1}, S_x's belief in that proposition is solely based upon competent understanding of S_{x-1}'s knowledgeable telling of that proposition. This is analogous to the idealising assumption implicit in SCHECHTER that Schechter knows all of his intermediate and final conclusions *only if* each of his deductions bestows knowledge.

Just as one might have good evidence that, sometimes, one goes wrong when performing single premise deductions, one can have good evidence that sometimes oneself and others

go wrong when interpreting the speech of others: words can be misheard, sentences incorrectly disambiguated, and context dependent constituents of sentences misinterpreted. Thus, if *n* is a sufficiently large number, S_n should think it likely that a misunderstanding has occurred somewhere in the chain.

Let us sharpen the structural similarities between SCHECHTER and TESTIMONIAL CHAIN. In SCHECHTER we have a series of competent deductions $D_1 \ldots D_{n-1}$ and in TESTIMONIAL CHAIN we have a series of competently understood tellings $T_1 \ldots T_{n-1}$. On Schechter's evidence (i) it is just as likely that one has made a mistake with D_1 as with any other deduction, and (ii) it is very likely on one's evidence that one has gone wrong with *one* of the deductions $D_1 \ldots D_{n-1}$. In the case of the testimonial chain, let us suppose that S_n's evidence is such that S_n should think it very likely that *some* member of the chain has failed to competently understand the previous speaker, but that every speaker is just as likely to have made a mistake of interpretation as any other speaker. If we judge in SCHECHTER that to come to believe that *q* solely on the basis of competent deduction D_{n-1} is to have an inappropriately high degree of confidence that q, given one's evidence, then we should judge that for S_n to come to believe that *p* solely on the basis of T_{n-1} is for S_n to have an inappropriately high degree of confidence that p, given her evidence. Either we must deny

that this interferes with S_n's having knowledge that *p* – which is tantamount to denying DEFEAT – or we must reject KT. For given that each of $S_2 \ldots S_n$ competently understands the telling they receive and S_n competently understands a telling from a knowledgeable speaker, if S_n comes to believe that *p* solely on the basis of this telling, she comes, according to KT, to know that *p.*

This is perhaps a little fast. KT will not have S_n as knowing that *p* if S_n lacks good undefeated evidence for the claim that S_n's understanding was competent or for the claim that S_{n-1} is a reliable testifier. Someone who wants to keep both DEFEAT and KT might argue that in TESTIMONIAL CHAIN S_n lacks the relevant good, undefeated evidence. But the prospects for such a line of argument are not good. On S_n's evidence, (i) there is a 0.k epistemic probability (where k has some non-zero value) that S_n has misinterpreted T_{n-1}, (ii) it is just as likely that a mistake has been made in interpreting T_{n-1} as it is likely that a mistake has been made in interpreting any other telling and (iii) it is very likely that a mistake has been made in interpreting *one* of the tellings $T_1 \ldots T_{n-1}$.

It should be easy to see that S_n might have good undefeated evidence for the claim that her understanding of S_{n-1}'s testimony was competent. We have said that k has some non-zero value. There must be *some* non-zero value such that k having that value is compatible with

S_n having good undefeated evidence that her understanding of S_{n-1}'s testimony was competent. Suppose one fails to have good undefeated evidence for the claim that one's understanding of some piece of testimony T_x was competent whenever the epistemic probability on one's evidence that one has failed to competently understand T_x is greater than 0. Practically no hearers (if any) ever have the sort of evidence on which their having failed to competently understand a telling is *that* improbable. The transmission theorist ought to want to avoid the unattractive commitment that practically no hearers, if any, are ever in a position to acquire testimonially based knowledge. If they are to do this, they must allow for the possibility of k having some non-zero value and yet S_n having good undefeated evidence for the claim that their understanding of S_{n-1}'s testimony was competent.

It is a little more difficult to see why S_n needn't lack the right sort of undefeated evidence for the claim that S_{n-1} is a reliable testifier. We have specified that S_n knows how many speakers are in the testimonial chain. We might think that if S_n's cognisance of the length of the testimonial chain means she ought to think that a misinterpretation has occurred somewhere in the chain, then it means that S_n ought also to think it likely that a misinterpretation occurred somewhere in the chain between S_1 and S_{n-1}. And if S_n ought to think this, then

she should think S_{n-1} an unreliable testifier, thus failing to have good undefeated evidence for the claim that S_{n-1} is reliable. But this is not quite right.

So far we have been talking quite imprecisely about likelihoods. Let's be more precise. Suppose that S_n ought to think that a misinterpretation occurred in a testimonial chain only if the (epistemic) probability on Sn's evidence that a misinterpretation occurred in a testimonial chain is greater than 0.k*. Where all hearers are known to be equally prone to misinterpret tellings, the (epistemic) probability that there has been a misinterpretation at some point in some testimonial chain increases as the number of testifiers in the chain increases. Thus, the (epistemic) probability that at least one of the hearers $S_2...S_n$ has misinterpreted a telling is greater than the (epistemic) probability that at least one of the hearers $S_2...S_{n-1}$ has misinterpreted a telling. Thus, it is possible that P(at least one of the hearers $S_2...S_{n-1}$ has misinterpreted a telling) < 0.k* < P(at least one of the hearers $S_2...S_n$ has misinterpreted a telling). So we cannot move from the claim that S_n ought to think that at least one of the hearers $S_2...S_n$ has misinterpreted a telling to the claim that she ought to think that at least one of the hearers $S_2...S_{n-1}$ has misinterpreted a telling. And so cases in which S_n does not have good undefeated evidence that no misinterpretation has occurred in the testimonial chain needn't also be cases in which she has no good undefeated evidence

that no misinterpretation occurred between S_2 and S_{n-1}. So it seems cases in which S_n ought to think a misinterpretation has occurred somewhere in the chain will come apart from cases in which S_n does not have the right sort of evidence for the reliability of S_{n-1}.

Now we can spell out more fully why this case makes trouble for the conjunction of KT and DEFEAT. Above, we noted that S_n's degree of confidence that *p* is inappropriately high given her evidence. If we like DEFEAT then S_n does not know that *p.* But if KT is right, then S_n *does* know that *p.* S_1 knew that p, and S_2 competently understood T_1, and had the right sort of evidence. So according to KT, S_2 knows that *p.* Similarly, S_3 competently understood T_2, and had the right sort of evidence, so S_3 knows that *p.* And so on until S_n receives a telling from S_{n-1}, who knows that p, whilst S_n has the right sort of evidence concerning her own competence and concerning the reliability of S_{n-1}. Either KT is false, or the inappropriateness of S_n's degree of confidence is no bar to knowledge, and DEFEAT is ill-conceived.

3.3 Costs.

An optimistic transmission theorist might wonder: what is so bad about having to reject DEFEAT? First of all, a lot of people -- though by no means all – think defeat a very significant epistemic phenomenon. As Lasonen-Aarnio notes: 'it has been an important

desideratum of epistemological theorising in the post-Gettier era that one's theory be able to accommodate the phenomenon of defeat' (2010).

There are other, less generic costs associated with DEFEAT's rejection. If one rejects DEFEAT, the transmission theorist may find herself unable to accommodate the sort of data that epistemologists of testimony typically want to be able to accommodate. For example, most epistemologists of testimony agree that only on a 'radically counter-intuitive picture' might hearers obtain or maintain testimonially based knowledge that *p* when they have very good evidence that a testifier is unreliable (Lackey, 2008). But rejecting DEFEAT would expose our theorist to just such a picture. To see this, suppose that A knows at t that S is reliable. But shortly after t, A begins to acquire a great deal of evidence that A is hopelessly unreliable (perhaps A is smeared by all the newspapers and badmouthed by normally trustworthy public figures). This evidence, though, is *misleading*: A remains reliable throughout these efforts to destroy his reputation. But suppose that A, after accumulating a great deal of this misleading evidence, continues to believe what S knowledgeably tells her. The transmission theorist who rejects DEFEAT is under pressure to allow these as cases of testimonially based knowledge: A knew at t that S was reliable and cannot be said to have lost this knowledge just because she has acquired misleading evidence.

It is worth pointing out too that even those epistemologists for whom rejecting DEFEAT is a matter of indifference tend to endorse an anti-luck condition on knowledge. Similarly, those epistemologists who do not mind rejecting ANTI-LUCK tend to be epistemologists who would be very unhappy about rejecting DEFEAT. Few epistemologists care about neither. Being sanguine about one of the tensions mapped by this paper may be a viable strategy. Being sanguine about both is not.

3.4 Generalising.

Suppose you're persuaded by my case against KT. Ought you to be wary of endorsing any transmission thesis? One might be tempted to think that rather than diagnosing a tension between transmission and DEFEAT, we ought to conclude that the transmission theorist needs to make more liberal use of the resources that DEFEAT offers, and go for something like this:

KT*: Necessarily, if

(i) S knows that *p,*

(ii) A comes to believe that *p* solely on the basis of competent understanding of S's testimony,

(iii) A has good, undefeated evidence (a) that *A's understanding was competent* and (b) that *S is a reliable testifier* and

(iv) A has no other relevant defeaters,

then A knows that *p.*

Whether or not we should think of such a proposed adjustment as a good one will not depend on whether KT* will turn out to mesh with DEFEAT. It looks very well placed to do *that*. The question is whether it does so at too high a cost.

Consider again the tension, diagnosed by Lasonen-Aarnio, between SPC and ANTI-LUCK. One might respond by reworking one's favoured closure principle so that it requires of deductions that they must not be too lucky, either via a distinct clause or by building such a condition into the notion of competent deduction. It seems that if these reworkings preserve closure principles, they do so at a cost, given that it looks like any grasp we have of what it is for beliefs to avoid lucky truth is insufficiently independent of our grip of what counts as knowledge for a principle reworked along these lines to count as non-trivial.[25] KT* confronts us with a similar puzzle. What we ought to think of it will depend on whether we can spell out clause (iv) -- or some functionally equivalent clause -- in such a way as to keep KT* from circularity.

One might try to spell out clause (b) as the requirement that A have no defeater for the claim that she is at the end of a suitably short testimonial chain. One obvious worry concerns how to characterise a 'suitably short' chain without reference to chains that have testimonially based knowledge as their output. *Maybe* there is some way of doing this, but if there is, it is

[25] But then, it may be that even once we strip away our grasp of what counts as knowledge, we can tell a story about what it is for beliefs to avoid lucky truth. If this is so, then a closure principle revised along these lines might retain the power and interest of simpler statements of single premise closure. Whether or not we should be optimistic about the prospects for such a revision depends on deep problems as to the proper way of thinking about anti-luck conditions on knowledge, which I am not in a position to address here.

not obvious. But even if we were to bracket this worry, this way of unpacking (b) would be insufficiently general. Suppose A knows that S knows some proposition, and beginning with this proposition, is about to perform a long sequence of *n* single premise deductions. S will tell A whichever proposition is the output of the *n*th deduction. If *n* is sufficiently large, surely A ought to think it likely *some* mistake -- *whether interpretive or inferential* -- has occurred. In fact S performs each of his *n* deductions competently and tells A, who competently understands his telling that *p.* If you are the sort of person who thinks that the set-up of TESTIMONIAL CHAIN is inimical to testimonially based knowledge, you should think that this too is a case in which A does not acquire testimonially based belief that *p.* But this is despite A *knowing* that she is not at the end of a long *testimonial* chain. Once again, cautious pessimism -- this time with respect to the prospects for finding a transmission thesis that meshes with DEFEAT -- seems to be in order.

Conclusion.

Let us take stock of our position. It seems that there are tensions between SPC and both ANTI-LUCK and DEFEAT. If these tensions should make us suspicious of SPC, they should make us suspicious too of at least one transmission thesis: KT. There may be ways of reworking KT so that it meshes with ANTI-LUCK and DEFEAT whilst retaining the

character of a transmission thesis (that is, without becoming trivial or circular). But the prospects for such a reworking do not look particularly bright.[26] After all, the project of the transmission theorist is linked to one of the canonical projects of analytic epistemology: that of *analysing* knowledge. Transmission theories look a lot like attempts to analyse a *particular species* of knowledge. If the moral of post-Gettier epistemology is that we should shy away from such analyses, then the cautious pessimism that this paper recommends should please epistemology's moralists -- or at least not surprise them.

I described my project as that of pointing out some unnoticed pressures that transmission theorists are vulnerable to, rather than tracing the consequences of this vulnerability in detail. But I want to end by sketching out what new shapes the dialectic might take once the pressures noted in this paper are taken seriously.

First, one might admit that *if* ANTI-LUCK considerations tell against single premise closure then they also tell against KT, but protest that ANTI-LUCK does not tell against single-premise closure. Here is how this line of thought might go. One might think that there is some way of cashing out ANTI-LUCK that will not generate the sorts of tensions charted by Lasonen-Aarnio and by this paper. Williamson, for example, has argued that ANTI-

[15] This may depend on some details of how we understand circularity that are not noted here.

LUCK can be glossed in ways that do not make trouble for closure principles, so long as we are willing to be suitably mercenary in our evaluations of which worlds are relevant to determining which beliefs count as lucky. In the best case scenario for the transmission theorist pursuing this argument, it turns out that there are natural and theoretically adequate ways of cashing out ANTI-LUCK that are hospitable to the transmission theorists' position. In the worst case scenario, there are no such alternatives.

Suppose that – relatively speaking – things go well for the transmission theorist. If they are willing to adopt enough of Williamson's epistemological package they may be able to cautiously endorse some species of anti-luck epistemology. But even if no such tensions emerge, before any endorsement of ANTI-LUCK could be made with a clear conscience, we would need to have to hand a more careful unpicking of the varieties of anti-luck epistemology, and some equally careful tracing of their interaction with closure principles.[27]

Second, I want to say a little about the upshot of this paper for the dispute between the reductionist and the anti-reductionist. To some extent, what one takes its implications to be will depend on how sympathetic one is to its recommendation of cautious pessimism. Suppose one is sympathetic. Then the paper has little or nothing to say to the dispute

[27] Of course, it might turn out that there are tensions not envisaged here between elements of the Williamsonian picture and the project of the transmission theorist.

between the reductionist and the anti-reductionist. It does not seem that reductionism is more wedded to transmission than anti-reductionism, or vice versa: if one can be a reductionist without being a transmission theorist, one can be an anti-reductionist without being a transmission theorist.[28] But suppose one is more unsympathetic. There are two different ways this lack of sympathy might manifest itself, with different implications for the dispute between the reductionist and the anti-reductionist. Suppose one finds oneself suspecting that there are ways to amend KT so as to avoid the tensions charted in this paper. Such adjustments are *likely* to pull KT in the direction of the reductionist, insofar as they seem likely to build up a more demanding picture of what tellings have to be like if they are to produce testimonially based knowledge. But suppose instead one is tempted to amend one's preferred reading of anti-luck epistemology – perhaps by buying into the Williamsonian epistemological package. This response looks apt to be friendlier to the anti-reductionist. After all, the Williamsonian reading of anti-luck touts itself as allowing us to hang on to a robust single premise closure principle. Perhaps it can allow us a similarly bold story when it comes to testimony.

Mapping the dialectical position that transmission theorists find themselves in is not, then, straightforward: their situation can be placed in better or worse lighting. But the interest of

[28] I thank an anonymous reviewer for *Philosophical Studies* for this point.

these structural problems need not lie in the obviousness of their upshot for problems in the epistemology of testimony. Rather, mapping these structural problems makes vivid quite how intimately problems in the epistemology of testimony are tied to more foundational problems in epistemology: how we should understand anti-luck conditions on knowledge; and how we should construe the ideology of defeat.[29]

Acknowledgements.

Thanks to Lizzie Fricker, John Hawthorne and Timothy Williamson for helpful discussion and comments. Particular thanks go to an anonymous reviewer for *Philosophical Studies* for detailed and insightful comments on earlier drafts.

[18] A historical note. In the 18th century Locke wrote: 'A credible man vouching his knowledge of it, is a good proof, but if another, equally credible, do witness it from his report, the testimony is weaker; and a third that attests from hearsay of an hearsay, is yet less considerable. So that in traditional truths, each remove weakens the force of the proof: and the more hands the tradition has successively passed through, the less strength and evidence does it receive from them (1997). One way of thinking about this paper is as spelling out this Lockean thought in the ideology of contemporary epistemology (although it is unclear whether what Locke had in mind is captured better by the ideology of anti-luck or by that of defeat).

References.

Bird, A., 1998. Dispositions and Antidotes. *The Philosophical Quarterly* 48 (191): 227-234.

Coady, C. A., 1992. *Testimony: A Philosophical Study*. Oxford: Oxford University Press.

Dummett, M., 1994. Testimony and Memory. In ed. Matilal, B. K. & Chakranarti, A., *Knowing from Words.* Dordrecht ; Boston: Kluwer, pp. 251-272.

Faulkner, P., 2011. *Knowledge on Trust.* Oxford: Oxford University Press.

Fricker, E., 2015. How To Make Invidious Distinctions Among Reliable Testifiers. *Episteme* 12 (2): 173 – 202.

Goldberg, S. C., 2008. Testimonial Knowledge in Early Childhood, Revisited. *Philosophy and Phenomenological Research* 76 (1):1–36.

Graham, P., 2006. Liberal Fundamentalism and Its Rivals. In eds. Lackey, J. & Sosa, E., *The Epistemology of Testimony*. Oxford: Clarendon Press, pp. 93—115.

Hawthorne, J., 2004. *Knowledge and Lotteries.* Oxford: Oxford University Press.

Hawthorne, J. & Lasonen-Aarnio, M., 2009. Knowledge and Objective Chance. In eds. Greenough, P.,& Pritchard, D., *Williamson on Knowledge*. Oxford: Oxford University Press, pp. 9-108)

Lackey, J., 2008. *Learning from Words*. Oxford: Oxford University Press.

– 2015. Reliability and Knowledge in the Epistemology of Testimony. *Episteme* 12 (02): 203 – 208

–

Lasonen-Aarnio, M., 2008. Single Premise Deduction and Risk. *Philosophical Studies* 141 (2):157 - 173.

- (2010). Unreasonable Knowledge. *Philosophical Perspectives* 24 (1):1-21.

Lewis, D., 1997. Finkish Dispositions. *The Philosophical Quarterly* 47 (187):143-158.

Locke, J., 1997. *An Essay Concerning Human Understanding.* London : Penguin Books.

Makinson, D. C., 1965. The Paradox of the Preface. *Analysis* 25 (6): 205-207.

McDowell, J., 1994. Knowledge by Hearsay. In eds. Matilal, B. K. & Chakrabarti, A., *Knowing from Words.* Boston; Dordrecht: Klumer, pp. 195--224.

Schechter, J., 2013. Rational Self Doubt and The Failure of Closure. *Philosophical Studies* 163 (2): 428-452.

Weiner, M., 2003. Accepting Testimony. *The Philosophical Quarterly* 53 (211): 256-264 .

Williamson, T., 2000. Knowledge and Its Limits. Oxford: Oxford University Press.

- 2011. Improbable Knowing. In ed. Dougherty, T., *Evidentialism and Its Discontents.* New York; Oxford: Oxford University Press, pp. 148-167.

Wright, S., 2015. In Defence of Transmission. *Episteme* 12 (1):13-28.

Abstract – Testimonial Pessimism.

Recent epistemological history has inclined towards 'testimonial optimism', keen to stress the division of epistemic labour and how ubiquitously we depend upon the words of others. Its counterpart, 'testimonial pessimism', marks out a cluster of gloomier views, which stress -- in different ways -- testimony's' epistemic shortcomings. My project in this paper is to establish a robust connection between pessimist readings of testimony, and two different commitments one might have in the philosophy of language: emotionist inferentialism, and what I call 'strong' readings of the *de re*. I do not aim to say, in this paper, what I think we ought to *do* with these connections; that is, I aim to remain agnostic on whether we should take the connections I sketch to give us a way of vindicating pessimism, or whether they are better read as part of an error theoretic project.

Testimonial Pessimism.

"But one who gives the learner not only the Truth
but also the condition for understanding it, is more than a
teacher"
- Johannes Climacus, *Philosophical Fragments.*

In Waugh's *Brideshead Revisited* the bluff Canadian, Rex, decides to convert to Catholicism in order to marry Julia Flyte. Julia's mother, Lady Marchmain, sends Rex to Father Mowbray. Despite Rex's determination to be converted, the instruction does not go well. After his third meeting with Rex, Father Mowbray has tea with Lady Marchmain:

> 'Well, how do you find my future son in law?'
> 'He's the most difficult convert I ever met.'
> 'Oh dear, I thought he was going to make it so easy.'
> 'That's exactly it. I can't seem to get anywhere near him. He doesn't seem to have the least natural piety or intellectual curiosity.
> 'The first day I wanted to find out what sort of religious life he had till now, so I asked him what he meant by prayer. He said '*I* don't mean anything. *You* tell *me*.' I tried to, in a few words, and he said: 'Right. So much for prayer. What's the next thing?' I gave him the catechism to take away. Yesterday I asked him whether Our Lord had more than one nature. He said: 'Just as many as you say, Father' (2000).

The problem is not that Rex is stupid (although we are not supposed to think him *clever* -- 'You must treat him like an idiot child, Father Mowbray', says Lady Marchmain.) Rather,

Rex lacks the *requisite sensibility* to have the sort of conversation Father Mowbray wants, and *needs* him to have. Rex's 'docility', as Father Mowbray calls it -- his indifferent deference to doctrine -- is no help at all. Rex easily -- too easily -- accepts everything Father Mowbray says. But something is missing. Father Mowbray's speech is somehow 'misfiring': he cannot bring about in Rex the state he seeks. Surely most of us have been in Father Mowbray's position: an interlocutor insists that they agree with us; we remain convinced that we have been misunderstood. Our speech is obstinate, indolent, and will not do the work we want it to. Probably most of us have been in Rex's position, too, perhaps dimly aware that we cannot quite grasp what is being said. Such misunderstandings are perhaps more naturally the terrain of the novelist than the philosopher: so much of why and how people fail to understand each other looks irreducibly particular, embedded in the peculiarities of biography and temperament.[30]

But some philosophers, undeterred, want to say something systematic about the ways in which testimony can fail. Recent epistemology has inclined towards 'testimonial optimism', keen to stress the division of epistemic labour and the ubiquity of our epistemic

[30] Kundera is a case in point. He tells the reader: 'If I were to make a record of all Sabina and Frank's conversations, I could compile a long lexicon of their misunderstandings' (1984). Instead, he compiles a short dictionary of misunderstood words: 'woman', 'fidelity', 'betrayal', charting the ways in which each's significance is different for Sabina than for Frank.

interdependence.[31] Its counterpart, 'testimonial pessimism', marks out a cluster of gloomier views, which stress, in different ways, testimony's shortcomings.[32] My project in this paper is to establish a robust connection between pessimist readings of testimony and two different commitments one might have in the philosophy of language: emotionism, and what I call 'strong' readings of the *de re*. I do not mean to say, in this paper, what I think we ought to *do* with these connections; that is, I aim to remain agnostic as to whether we should take the connections I sketch to give us a way of vindicating pessimism, or whether they are better read as part of an error theoretic project.

Typical pessimists start by identifying particular kinds of testimony, or particular kinds of testimonially based belief, that look (perhaps systematically) defective along some important axis of evaluation. For example, Meskin (2004) notes that even realists about aesthetic properties widely accept that 'aesthetic testimony cannot provide…justification or

[31] Lackey's remarks are typical: 'Testimony is an invaluable source of knowledge. We rely on the reports of those around us for everything from the ingredients in our food and medicine to the identity of our family members, from the history of our civilization to the limits and contents of our planet. If we refrained from accepting what others told us, our lives, both practically and intellectually, would be unrecognizable' (2008).

[32] This is a self-consciously rough statement: there is a good deal of heterogeneity in what the core commitments of a testimonial pessimist are taken to be: Mogensen, for example, takes pessimism about moral testimony to involve commitment to normative considerations 'specific to the moral domain' (2015). This picks out a much narrower cluster of views than Howell's characterisation of the view: 'moral deference is [in differing degrees] problematic' (2014). I take these disputes to be broadly terminological and do not want to adjudicate between them. Instead, I will be working with an expansive conception of the pessimist's commitments.

knowledge', Sliwa (2012) tells us that much of the literature assumes there is something defective about basing one's moral beliefs on testimony and Mogensen (2015) recounts that 'although we can happily rely on the word of others when it comes to the geography of India or the mating habits of penguins, it seems we can't be quite so willing to accept testimony on moral questions'. And in light of traditions that tie mathematical knowledge to the ability to articulate a *proof*, mathematical beliefs based only on testimony can seem to lack the epistemic gilding required for knowledge (Chisholm, 1966; Easwaran, 2009; Williams, 1972).

One reason to be wary of these ways of talking is doubt that the suggestive categories -- 'moral testimony', 'aesthetic propositions' -- can be made precise. It is easy enough to say that 'eating meat is wrong' is a moral proposition, but a little tricker to know what to say about 'eating meat is wrong or it is not', or 'eating meat is wrong *or no elephants are green*'; it is harder still to know what to make of testimony that exploits what look like moral predicates within the scope of intensional operators: 'Someone said that eating meat was wrong'; 'No one thinks it's OK to eat veal'.[33] A quick, vivifying example: suppose you

[33] It is likely that making good on these categories will require some sort of theory of 'aboutness'; see Yablo (2014) for a recent systematic treatment.

ask me whether we need to buy more bread and I tell you, knowing that you know perfectly well that we ought not to torture children, 'Either we are out of bread or we ought to torture some children', and you believe what I tell you (and work out that we need to buy more bread). We clearly ought *not* to count this as a case of moral testimony, but it's not easy to say *why* not. Theorising the pessimist's taxonomies is not a job for the optimist; it is the pessimists for whom things look bad if statements of their view turn out to lean heavily on rickety taxonomies.

An alternative approach to pessimism -- the approach I favour -- starts not with particular cases (or species) of what look like defective testimony, but begins by casting about for positions in the philosophies of mind or language which *predict* pessimist intuitions in a range of cases. Once one can show that testimonial pessimism looks to flow from a given commitment, one can endorse the commitment along with its pessimistic upshot, or use the connection to draft an error theory for pessimist intuitions: 'we are tempted by pessimism insofar as we are tempted by this other philosophical position, but the position is wrong, and we should reject it, along with its pessimistic implications.' Positions with some pessimistic upshot are not hard to track down. Suppose one thinks something like the following: 'I believe certain things -- certain propositions -- about myself using a *de se* mode of

presentation. This *de se* mode of presentation is essentially *private*; no one else can think of me in just this way that I do. Nevertheless, I can express and tell others these propositions that I believe about myself: typically, I use the word 'I' to do so.'[34] Someone who thinks along these lines, and who testifies using the word 'I', should take themselves to tell other people propositions that their audience can only approximate.[35] The private, *de se* way they have of thinking about themselves cannot be transmitted by testimony.

On standard models of successful communication, even more modest views of *de se* content are troublesome. The standard Stalnakerian 'package delivery' model -- taken for granted by many epistemologists of testimony – says that, in good cases 'when I communicate my beliefs to you, there is something that I believe, that I convey, and that you come to believe' (Moss, 2012).[36] This 'package delivery' model makes the following thought natural: A's belief that *p* is testimonially based only if someone has told A that *p.* If you are the sort of person who individuates propositions very coarsely -- you count *that Venus is Hesperus* and

[34] For classic statements of this sort of view, see Lewis (1979) and Perry (1979). For criticisms, see Magidor (2014) and Millikan (1990).

[35] This line of thought is compatible with thinking that we do not *always* use the word 'I' to ascribe *de se* attitudes. When we say sentences like 'I expected that I would win', the embedded first-personal pronoun admits of a third personal reading. (see Kaplan, 1989; Lakoff, 1972; Morgan, 1970; Moss, 2012).

[36] Goldberg writes 'the literature focuses exclusively on…cases in which the proposition attested to is identical with the proposition believed by the recipient' (2001). See also Fricker (2012), Goldberg (2006), Lackey (2008), Peet (2015, 2016). I borrow the 'package delivery' idiom from Sarah Moss (2012).

that Venus is Phosphorous as the same proposition -- you will probably want to supplement rather than discard this model. One natural supplementation exploits guises: A's belief that *p* is testimonially based only if someone has told A that *p* using sentence *s*, and A believes *p* under a linguistic guise appropriately related to *s*. The 'appropriate relation' cannot just be sameness of sentence-type. Suppose David Cameron tells the Queen, 'I will resign as Prime Minister'. If the Queen comes to have testimonially based belief that David Cameron will resign as Prime Minister, she does (should) not believe this proposition under the guise of the sentence 'I will resign as Prime Minister'. The 'package delivery' model sits less easily alongside views that individuate propositions more finely. Moss makes this point vividly (2012). Suppose we model propositions as sets of *centred* worlds -- world-time-individual triples -- so that when David Cameron believes that he will resign, his belief state should be represented as an attitude to a set of worlds which mark David Cameron as the centre. If David Cameron tells the Queen 'I will resign', we should not think his telling successful if the Queen ends up with a belief attitude towards a set of worlds which mark Cameron as the centre.[37]

[37] Something like these dynamics will be a feature of many views which individuate content finely. Suppose Mhairi is in London and Jane in Aberdeen, and Mhairi tells Jane, 'The sun is shining here'. Given roughly Fregean ways of individuating content, we ought *not* to think that, in good cases, the proposition Mhairi expresses is the same as the proposition that Jenny comes to believe when Jenny believes that the sun is shining in London. Egan's 'content proliferation view' presents a more distinctive challenge to the 'package delivery'

The rest of the paper comes in four parts. In the first section I present a reading of medieval mystics as 'proto-pessimists'. I show that we can read these authors' pessimism as flowing from a tacit commitment to particular (controversial) views in the philosophy of language: emotionism and 'strong' views of the *de re*. These readings suggest a general connection between pessimism and, respectively, emotionist commitments, and 'strong' readings of *de re* content. In slogan form: emotionists should be pessimists, and so should 'strong' theorists of the *de re.* In the second and third parts of the paper I show that these suggestions can do interesting work, by sketching the upshot of the proposed connections for moral and aesthetic testimony. I show too that the connections I sketch should be of interest regardless

---

model. According to content proliferation, when Tony Robbins says 'You can change your life' to a packed hall, he 'makes a single utterance which expresses different propositions to different members of his audience', each of whom assign a different value to 'you' (2009). This sort of view sits uneasily with the 'package delivery' model insofar as it interferes with our thinking of Tony Robbins as *believing each* of the propositions his audience members come to believe on his say-so: one may believe that everyone in the room can take control of their life without believing *of* each person in the room that they can change their life. Elsewhere, the ideology of 'transformative experience' -- much in vogue -- may suggest a view like the following: 'Some experiences one simply needs to 'have for oneself' if one is to know what they are like: seeing red, having a child, and taking part in a revolution. No testimony, however skilful or evocative, can teach me what it is like to do these things.' If this is right, there will be some propositions that revolutionary, red-seeing, child-bearers can grasp and communicate to each other, but which they will be unable to communicate to audiences who lack their experiences. For arguments along these lines, see Paul (2014). This sort of view has interesting moral and political consequences: Iris Young argues that in order for social institutions to be just, we need to grasp what life is like for those who are differently socially situated (2000). Testimony seems like the obvious way to acquire this grasp, unless occupying a particular social position is an experience that one must have 'for oneself'. Rachel MacKinnon has argued along these lines that 'gender transitions count as...transformative experiences' (2015).

of whether you want to vindicate pessimism or to explain away the intuitions in its favour. In the fourth section I respond to objections. One thing I do not mean to do here is rehearse arguments made elsewhere about how we ought to theorise testimonial pessimism.[38] My project here is exploratory and suggestive, gesturing towards some ways in which we might expand the theoretical territory available to the testimonial pessimist, rather than straightforwardly insistent as to where the testimonial pessimist must pitch her camp.

1.1 Testimonial Pessimism and Mysticism.

Medieval authors recording their mystical experience were anxious to regulate access to their testimony. The anonymous author of *The Cloud of Unknowing* gives detailed and vehement instructions to the reader:

[38] For accounts of testimonial pessimism which stress the epistemic significance of understanding see Hills (2009, 2013), Hopkins (2007), McGrath (2009), Nickel (2001). McGrath (2011) connects pessimism and non-cognitivism; Andow (2014), pessimism with contextualist and relativist semantics. Sliwa (2012) argues against taking testimonial pessimism seriously; Meskin (2004) that aesthetic testimony can transmit justification. I am open to the thought that what instincts we have that tell in favour of testimonial pessimism are overdetermined, or have a syncretistic character, drawing what force they have from multiple sources. As such, my suggestions might be read as complements rather than alternatives to views like Mogensen's (2015) 'authenticity' centred and Howell's (2014) 'aretaic' accounts of pessimism.

> In the name of the Father and of the Son and of the Holy Spirit, I command and beseech you, with as much strength and force as the bond of love is capable of bearing, whoever you may be that have this book in your possession, whether by carrying it as a messenger or borrowing it, that so far as you are able to you do not willingly and deliberately read it, copy it, speak of it, or allow it to be read, copied, or spoken of, by anyone or to anyone, except by or to a person who…has undertaken truly and without reservation to be a perfect follower of Christ…(‘Cloud’, 1924).

> …I do not desire that this book should be seen by worldly chatterers, public self-praisers or fault finders, newsmongers, gossips or scandal-mongers, or detractors of any kind, for it was never my intention to write of such matters for them… (‘Cloud’, 1924)

Instructions given in *The Mystical Theology of St Denis* are less exhaustive, but just as uncompromising:

> Take care that none of those unwise people still dwelling in their senses hear about these matters (‘Denis’, 2001).

One might regard these exhortations as philosophically uninteresting: most of us have kept secrets. Certainly, some mystical writers feared persecution, or -- like the Quakers and dissenters who worried that their accounts might titillate -- that widespread access to their texts would produce degeneracy (Atchenstein, 2002).[39] But the concerns of our medieval

[39] For example, reflecting on the poetry of Elizabeth Singer, Isaac Watts writes: 'Let it be observed, that it was much the Fashion, even among some Divines of Eminence in Former Years, to express the Fervours of devout Love to our Saviour in the Style of the Song of Solomon: And I must confess that several of my Composures

authors are distinctly epistemic: those who fail to 'prepare for the contemplative life by means of virtuous active living' are not just thought *unworthy* of religious knowledge but *incapable* of coming to have the *right sort* of attitude towards the propositions known and expressed by our mystical authors ('Cloud', 1924). Let us call the proto-pessimist position that 'unprepared' audiences are incapable of epistemically proper responses to religious testimony *epistemic calvinism.* The project of this paper is not to argue that we ought to endorse epistemic calvinism, but to show how we can read it as flowing from tacit commitments in the philosophy of language.

Here is a natural thought: which propositions one is able to get into cognitive contact with depends on one's conceptual apparatus. One may be *hermeneutically deprived*: unable to know certain propositions not because of a lack of evidence, but because one lacks the requisite conceptual apparatus (see Fricker, 2007). What is suggestive and interesting in our medieval authors is that they seem tacitly committed to the idea that 'reforming grace' acts

---

in Verse written in younger Life were led by those Examples unwarily into this Track. But if I may be permitted to speak the Sense of maturer Age, I can hardly think this the happiest Language in which Christians should generally discover their warm Sentiments of Religion, since the clearer and more spiritual Revelations of the New Testament. Yet still it must be own'd, there are some Souls favour'd with such beatifying Visits from Heaven, and raptur'd with such a Flame of divine Affection, as more powerfully engages all animal Nature in their Devotions, and constrains them to speak their purest and most spiritual Exercises in such pathetick and tender Expressions as may be perversely profaned by an unholy Construction.' Quoted in Atchenstein (2002).

to combat hermeneutic deprivation. The authors of *Denis* and *Cloud* are clear that the operations of grace have made 'perfect followers of Christ' capable of 'full knowledge of…the works of God' *by altering souls' capacities for love* ('Cloud', 1924). But how can a change in our *affective* dispositions be thought to overcome hermeneutic deprivation? I suggest that tacit commitment to an emotionist semantics for (some) religious vocabulary would explain why our authors anticipate changes in affective dispositions as combatting hermeneutic deprivation.

A second way of thinking about hermeneutic deprivation appeals to 'strong' views of singular content. One might think something like this: 'I believe *de re* as well as *de dicto* propositions. One can only believe a *de re* proposition about some object *O* if one is *acquainted,* that is, stands in a privileged *epistemic* or *causal* relation, with *O.* Some people think that one counts as acquainted with *O* just so long as one has heard the right sort of name for *O*; but they are wrong: acquaintance with *O* requires much more than this. If I know, and tell you, a *de re* proposition about Jones, but you are not *acquainted* with him, you cannot come to know the proposition I tell you.'

Think of Johann Galle looking through his telescope and seeing Neptune for the first time. If *seeing* Neptune is required for acquaintance, then if Galle tells Le Verrier 'Neptune is almost exactly where you predicted it would be', he is telling Le Verrier a *de re* proposition that Le Verrier, not having seen Neptune for himself, can neither entertain nor believe. This suggests a second way of reading our authors: if we think of mystical experience as acquainting an agent with God, a religious 'elect' will be able to entertain *de re* propositions that are cognitively inaccessible to other agents.

1.2 Emotionist Semantics, Sketched.

Let us get clear about the emotionist's commitments. Emotionist semantics are often understood as one aspect of an *inferentialist* programme in the philosophy of language. Williamson draws a self-consciously rough-and-ready distinction between referentialist and inferentialist tendencies. Referentialists lend explanatory priority to a referential semantics for a language and think the inference rules for a language are explained in terms of this semantics, where inferentialists take inference rules to be explanatorily prior to a referential semantics (2009). Consider these rules for the word 'fruit':

Fruit Introduction

X is an apple.

X is a fruit.

Fruit Elimination

X is a fruit.

X is healthy.

In the simplest inferentialist model available, iff one is disposed to perform *Fruit Introduction* and *Fruit Elimination* does one understand the word 'fruit' or have the concept *fruit* (Boghossian, 2003). So far our characterisation of the inferentialist position, even qua crude sketch, is too narrow. There are relevant rules other than inference rules: ''language-entry' rules that connect perceptual states to moves in the language game, and 'language-exit' rules that connect moves in the language game to non-linguistic actions' (Williamson, 2009). In the simplest available model of an emotionist semantics for some term 't'; only if one is disposed to perform *t Exit-Condition* does one understand term 't' or have the concept *t*:

t Exit-Condition

X is t.

Feel: emotion e towards X.[40]

Some details of the emotionist picture depend on how one likes to think of propositions. On fine-grained accounts an emotionist semantics will say: in order to have cognitive contact with some propositions, one's affective dispositions have to be structured in a certain way. An alteration of one's emotional capabilities might alter *which propositions one is capable of entertaining* (and believing, knowing, and so on). So long as a mystic has affective dispositions that a 'worldly chatterer' lacks, we can think the two capable of grasping distinct contents. This makes trouble for testimony, and the mystic unable to transmit her beliefs to the worldly.[41]

On more coarse-grained models, two people might have very different conceptual apparatuses, and yet be in cognitive contact with exactly the same propositions. For coarse-grainers, even once we tie affect to conceptual competence, mere cognitive contact comes too cheaply to let us map differences of affective disposition onto differences in cognitive reach. But, as noted above, it is natural for coarse-grainers to supplement the ideology of the

[40] Note the 'only if': accepting an emotionist semantics for some term 't' is consistent with taking 't' to be governed by inference-rules as well as exit-conditions and with taking 't' to make a semantic contribution to the sentences of which it is a constituent.

[41] You might think the relevant affective dispositions can only be generated by contact with God, but you might also think that certain important -- by the lights of the mystic -- affective dispositions can be cultivated by living a certain sort of life or by being a part of a religious community.

proposition with something like one of *guises*, and to say something like this: testimony succeeds only when an audience who is told that *p* using sentence *s* acquires belief that *p* under a linguistic guise that stands in some appropriate relation to *s*. We can tie the notion of guises to affective dispositions, and say that unless a hearer is equipped with the right sort of affective dispositions, they will be unable to entertain propositions under the *right sorts of guises.* Testimony between the mystic and the worldly will fail where the latter's affective dispositions preclude access to the relevant guises.

Together these pictures give us the *emotionist* model of transmission failure: cognitive access to content depends on affective resources; where these are lacking in an audience, they cannot have propositional attitudes towards what they are told.[42] Importantly, to endorse the emotionist model, you needn't be committed to an emotionist semantics. You just need to think that emotionist semantics codify some important intuitions we have, and grant that these same intuitions structure our pessimistic judgements. You can think all this

[42] 'Transmission' is used with a variety of resonances in the testimony literature (see, for example, Lackey, 1999, 2008; Wright, 2016). When I talk about transmission failure here, I aim to pick out cases in which an audience is somehow 'blocked' from believing what they are told.

whilst judging that the theoretical trade-offs involved in adopting emotionist semantics are too costly. [43]

1.3 Singular Thought.

So far I have been talking quite loosely about singular thought and *de re* propositions. Let's sharpen things up. Lots of people have observed that there is an intuitive contrast between the sense in which the sentence 'Saul Kripke is bearded' is about Saul Kripke, and the sense in which 'The author of *Naming and Necessity* is bearded' is about Saul Kripke. Sometimes this contrast is spelt out in the following way: the first sentence, and not the second, is *directly about* Kripke. We can distinguish between the following two claims:

(1) Sentences which are *directly about* some object *O* express *singular thoughts* about *O*, or *de re* propositions about *O*.

[43] To almost any way of thinking, the relationships we have with other persons certainly teach us a great deal. The medieval authors I have been considering took very seriously the capacity of a relationship with God to reconfigure one's affective life. But surely the sustained emotional entanglements most of us have with other human beings are capable of transforming us in similar ways. If proponents of emotionist semantics are right to tie together affective dispositions and conceptual apparatuses, then our interpersonal relationships become hermeneutic resources.

(2) In order to have cognitive contact with a *de re* proposition about *O*, one must be *acquainted* with *O*.[44]

'Acquaintance' has been construed in a variety of ways, sometimes with a causal, sometimes an epistemic flavour. Causal accounts often allow that mastery of a singular term with the right causal history provides acquaintance (see Sosa, 1970; Hawthorne and Manley, 2014); Evans proposed an epistemic constraint: 'the subject must have a capacity to distinguish the object of his judgment from all other things' (1982). But for our purposes, what matters is this: where there is some sufficiently strict notion of acquaintance on which hearing someone express a *de re* proposition about *O* is not enough to acquaint one with *O*, there is a theory of *de re* content that makes space for a distinctive species of transmission failure. Call a view of the *de re* a *strong* view if it says that to have *de re* thoughts about *O* one needs to have had contact with *O* that is not mediated by another speaker. Here is one strong view: only agents who have had mystical experience are acquainted with God. It follows that the 'unwise people still dwelling in their senses' will not be in a position to

[44] A third claim articulates the thesis that part of what is special about singular propositions is that they are 'object dependent'; that a singular thought (singular proposition) which is directly about *O* has *O* itself as a constituent (see, for example, Armstrong and Stanley, 2011). This claim will not be important in what follows.

entertain or believe the *de re* propositions believed and expressed by mystics. Call this the *de re* model of transmission failure.

We now have sketches of two novel models of transmission failure; both diagnose certain audiences as incapable of grasping the propositions they are told.[45] In what follows, I examine the prospects for extending these models to other species of testimonial pessimism.

2.1 Moral Testimony Pessimism.

Let us take stock. I have suggested that just as moral and aesthetic testimony can be given a pessimistic reading, so can mystical religious testimony, and have sketched both an emotionist and a *de re* model of transmission failure. In the next two sections of the paper I look at how these two models can be extended to other cases of testimonial pessimism. There are extant (if sketchy) emotionist semantics for moral vocabulary. First I show how to use these to extend the emotionist model to moral testimony. Second, I suggest that the *de re* strategy can be mapped onto the theoretical tradition of the 'aesthetic attitude' associated

[45] Note too that both the emotionist model and the *de re* model of transmission failure are *scalable*: neither posits domain specific norms.

with Shaftsbury, Kant and Schopenhauer, and that this mapping allows us to extend the *de re* model to aesthetic testimony.[46]

First let us a sketch an emotionist semantics for moral vocabulary. Prinz (2007) defends a version of emotionism for moral vocabulary. A simple emotionist semantics along the lines he suggests looks something like this:

(PRINZ-SEMANTICS): One understands the word 'wrong' and has the concept *wrong* iff one is disposed to perform *Wrong Exit-Condition.*

Wrong Exit-Condition

X is wrong

Feel: disapprobation towards X; be motivated not to X .[47]

[46] There are important structural disanalogies between epistemic calvinism and testimonial pessimism in other domains. Epistemic calvinism doesn't want to declare *all* audiences as incapable of forming non-defective belief on the basis of religious testimony. Standard statements of moral and aesthetic testimonial pessimism are more totalising, and often presented as a thesis typically about 'normal, mature' agents. These structural differences mean that an emotionist semantics interacts differently with different flavours of pessimism.

[47] I stress that this is a *toy* emotionist semantics: I am using it in order to get a sense of the dynamics one might end up with were one to adopt a semantics along these lines. There are all sorts of interesting questions that I cannot address here: for example, how might an emotionist semantics cope with boolean operators or probability operators? It is not obvious what kinds of affective dispositions a semantics should require for sentences like 'If Hitler did it, then it was wrong', 'Killing is probably wrong' or 'Either stealing is wrong or killing is'. Another puzzle involves generality: we can quantify over events by saying things like 'Most of what

If someone lacks the affective dispositions needed for moral concepts, then they will not be in a position to have moral testimonially based beliefs. Just as 'the worldly chatterers' will not understand the testimony of our mystical authors, those who fail to have the relevant dispositions will be incapable of the sort of uptake of testimony required for testimonially based belief.

But we need to say something more than this if we want to make space for testimonial perssimism. It seems that sometimes we morally testify to agents who *do* have the relevant affective dispositions. Such agents *are* in a position to believe what we tell them. Can an emotionist semantics explain our discomfort in these cases? I think so. Consider the following cases:

Hitler did was wrong'. But it is not obvious how to expand a simple emotionist semantics to such cases. Note too that PRINZ-SEMANTICS presents a very *thin* way of sketching the relevant exit-condition; one could enrich it in various ways. An exit-condition could appeal to 'clusters of beliefs, interests, norms, practices, values, affective dispositions, and ways of interpreting and interacting with the world' (Swanson, m.s.). When motivations are written into the exit-condition, this species of inferentialism has clear similarities to motivational judgement internalism; the thesis that at least in 'standard' agents, moral judgements themselves have motivational force (see Bjorklund et al., 2012; Drier, 1990; Smith, 1994).

RESTAURANT: You're in a restaurant with a friend and trying to work out how to split the bill. You both decide to do the arithmetic separately. At exactly the moment you work out for yourself that you each owe £18.67, your friend tells you 'We both owe £18.67'. You end up believing that you each owe £18.67.

CLASS: I'm teaching, and I ask my student to consider the sentence 'All bachelors are unmarried'. They consider it, and say 'I've never thought about that before, but now you mention it, that does seem right', and they end up with the belief that all bachelors are unmarried.

FLOWER: I am shopping with my friend, whom I mistakenly think colour blind, for a white carnation. In the florist's, he absentmindedly picks up a flower without looking at it and as he looks down at it, I tell him 'That flower is red'. He hears me just at the moment he sees the flower for himself. He ends up believing that the flower is red.

In RESTAURANT, you don't mistrust your friend, and you in fact come to believe exactly the proposition they tell you. Their testimony needn't even be epistemically irrelevant: that

they got the same answer as you should reassure you that your own calculations went well. But we don't think of your belief as *testimonially based.* You have, in some important sense, *judged for yourself* how much is owed. Similarly, in CLASS, I have prompted my student to entertain a novel (for them) proposition; but we ought not to think of my student as *taking my word* for it that all bachelors are unmarried. In FLOWER, again, we do not think of my friend as *taking my word for it* that the flower is red, although he trusts me and ends up believing the proposition I tell him.

We can sharpen these intuitions by thinking about the *I-thou* relations in play in these three cases. In standard cases, when I believe something on your say-so, I am entitled to *blame* you for my false belief if it turns out that you told me something false.[48] Characteristically, in acquiring a testimonially based belief that P, one acquires an entitlement to blame a speaker if it turns out that *p* is false. We can use the presence or absence of these entitlements to diagnose whether or not we should think of a belief as testimonially based. But in neither RESTAURANT, CLASS, nor FLOWER is the hearer entitled to blame the speaker if they end up with a false belief, (although they may be entitled to rebuke the speaker for having spoken falsely).

[48] Something along these lines has been defended by Goldberg (2006, 2015), Moran (2006), Ross (1986).

I suggest that a natural model of moral testimony maps its dynamics closely onto those in play in RESTAURANT, CLASS, and FLOWER, so that when an audience *believes* a moral proposition that I tell them, we judge that they fail to take my word for it, and so judge that the belief fails to be testimonially based. According to PRINZ-SEMANTICS, if Rex has the concept *wrong*, then he is disposed to feel disapprobation towards *a* when he believes the proposition that *a* is wrong. Without such a disposition, Rex either lacks the concept *wrong* or has the relevant concept but does not believe that *a* is wrong. You might think that the relevant affective dispositions only 'kick in' for Rex when he actually *believes* *a* is wrong. But it is rather more natural to associate these dispositions with a *broad range* of propositional attitudes (supposing, entertaining, wondering).[49] Once we make salient a background picture in which agents who have moral concepts are equipped with wide-ranging suites of affective dispositions, it stops looking apt to think of agents who believe the moral propositions they are told as having *testimonially based beliefs*, rather than as having beliefs which are (at least in part) the output of their own affective responses. In slogan form,

[49] This expansiveness can be motivated by reflecting on philosophical methodology. Moral philosophers who offer us thought experiments ask us to contemplate imaginary scenarios and, often, to have affectively laden responses to these imagined situations. Positing sharp contrasts between the affective responses associated with belief and more non-committal attitudes would make it difficult to account for the epistemic role played by thought experiments. See Gendler (2010) for discussion of the role of thought experiments in philosophy.

the view is this*: the dispositions that underpin Rex's ability to grasp a moral proposition are naturally read as part of the basis for his belief in that proposition.* After all, as Howell notes, we often expect an agent's moral beliefs to 'stem from…feelings and intuitions which are indicative of [their] moral character.'

More slowly: I propose that we have something like the following tacit model of what goes on in standard cases of moral testimony. If I tell you that *a* is wrong and you have the sorts of dispositions that allow you to grasp the communicated content, you have a suite of your own affective resources that 'kick in' when my testimony prompts you to entertain a novel moral proposition.[50] And so you do not present as having simply *taken my word* for it that *a* is wrong; we read your belief as, in Howell's idiom, 'stemming from' these affective responses (2012).[51] This model might, of course, get various things wrong. But what

[50] Exactly how one fills in the model will depend on how one likes to think of the emotions, and in particular, whether or not one is a cognitivist about emotions. If you favour cognitivism cases of moral testimony will map more neatly onto cases like CLASS than if you favour non-cognitivism, in which case, at least if you doubt that perceptual experience has cognitive content, what goes on in the moral case will look more like what goes on in FLOWER.

[51] It is important to see that the model I am suggesting here is neutral with respect to the *epistemic value* of our affective dispositions. In particular, I don't suggest that having a certain cluster of affective dispositions is required if one is to have *justified* moral beliefs. To think this would be to run together claims about a belief's etiology and its epistemic standing. But how we should relate these two things is unclear. For discussion, see Vavova (2015) White (2011).

matters, for my purposes, is that something *like* it structures our judgements about standard cases of moral testimony.

Let us regroup. The emotionist model makes two core suggestions: first, that agents who lack certain non-cognitive attitudes and abilities are incapable of acquiring testimonially based moral beliefs (as they cannot grasp the relevant contents), and second, that agents *with* these attitudes and abilities are typically blocked from becoming *epistemic dependents* in the way that testimonially based belief requires. We can, then, explain pessimism about moral testimony thus: moral testimony will *typically not give rise to testimonially based belief*. Agents without the resources to grasp the proposition they are told will not end up believing it. Agents with the resources to grasp the proposition they are told might end up believing it, but they will not, in general, believe it in the right sort of way to count as having *testimonially based* belief.

2.2 Varieties of Pessimism.

You might think these considerations suggest that the category of moral testimonially based belief is hopelessly confused -- that the way we think about what it is for a belief to be

testimonially based combines so badly with our tacit model of what goes on in cases of moral testimony that we cannot make sense of the category. But this is much too quick. We *can* make sense of genuine moral testimonially based beliefs. There are (at least) two different ways to do this. First, a disposition to α in circumstances *c* is not always *manifested* (Cohen and Handfield, 2007). A disposition is *finked* iff a disposition is triggered by its stimulus but some feature of the context acts to remove the disposition before it can be manifested (Cohen and Handfield, 2007). (A sorcerer who casts a spell to strengthen a vase as it falls towards the floor *finks* the vase's disposition to shatter when struck (Lewis, 1997).) A disposition is *masked* when the disposition's manifestation process is inhibited (Cohen and Handfield, 2007). (A pharmacist masks the disposition of a poison to kill when he ingests an antidote (Bird, 1998).) To model what look like genuine testimonially based moral beliefs, one thing we can do is imagine cases in which an audience's affective dispositions are *finked* or *masked*. A second thing we can do is imagine cases in which an agent has lots of reasons to think her dispositions *misleading* when in fact there is nothing wrong with them, but mistrust means that she really does *depend* on a testifier when he tells her that (say) she ought not to pay her friend back.

MASK: Jane has started taking some new medication. It is a common side-effect of this medication that one becomes almost totally affectless, but for just a few days -- after a week on the pills, one is back to normal. Normally, Jane has intense emotional responses to animal cruelty, but when, on her medication, she sees some children about to set fire to a cat, she has no emotional response. A friend is with her, and says 'Those children shouldn't be doing that!' She trusts her friend, and believes what she says.

DEFEAT: Fritz has spent a long time as an undercover spy in a foreign totalitarian regime. He suspects that his time there will have distorted his affective attitudes, and resolves not to trust his own moral judgements for his first few months back home. In fact, Fritz has been much less affected by his time abroad than he thinks. A friend begins to tell him about a court case in which the judge has excluded important and decisive evidence from consideration. Fritz's own intuitions, which he assiduously ignores, are that this is a terrible injustice, but he trusts his friend when he says 'This is a terrible injustice'.

What should someone drawn to pessimism say about such cases? Nothing in the emotionist model suggests that the testimonially based beliefs that Fritz or Jane end up with must fail to count as knowledge. But such a result might be compatible with a careful construal of pessimism. Howell, for example, is careful to say that accommodating pessimism about moral testimony should be compatible with allowing that sociopaths ought to defer; the moral testimony pessimist, he thinks, should defend a thesis about 'normal agents in normal circumstances' (2014). (Someone who takes the emotionist semantics seriously probably cannot insist that the sociopath defer, if they take the sociopath to be incapable of grasping moral contents. But they can, importantly, insist that the sociopath engage in what Howell calls 'active deference' -- acting *as though* a claim that *p* is true, without having a belief that *p*.[52]) Howell's own account does not predict that we ought to have pessimist intuitions about MASK or DEFEAT style cases. He argues that when agents rely on testimony for their moral beliefs, those beliefs will typically fail to be 'subjectively integrated': 'motivations and feelings are likely to be disconnected from beliefs sustained by deference, but so are moral intuitions and related non-cognitive attitudes' (2014). But in FINK and DEFEAT, the testimonially based beliefs are 'integrated' in Howell's sense -- they just fail to be

[52] 'Active deference' in Howell's sense seems closely related to Moore's account of 'disengaged grasp' of thick concepts, where '[t]o grasp a thick ethical concept in the disengaged way is to be able to recognize when the concept would (correctly) be applied' but without sharing in the 'beliefs, concerns, and values [that] give application of the concept its point' (2006). Thanks to Nakul Krishna for this point.

etiologically connected to those moral intuitions and related non-cognitive attitudes. Howell's verdict might be the right one: FINK and DEFEAT certainly do not look much like the paradigm cases which motivate pessimism about moral testimony.[53] Optimistic readings of FINK and DEFEAT need not vex a cautious pessimist.

More pessimistic readings of FINK and DEFEAT are available. Whilst they show that we can make sense of the idea of testimonially based moral beliefs, they also show us cases of genuine moral testimonially based beliefs that certainly involve some kind of *flaw*. Jane's medication has cut her off from some important epistemic resources. For a few days, she has become *epistemically dependent,* and is correspondingly *epistemically vulnerable* in ways that she normally is not. Fritz has good reason to think he is an unreliable judge, when he is in fact reliable: he is responding to misleading evidence about his own abilities, and this looks like a bad situation for Fritz to be in. We should not be surprised if we have

[53] A paradigm case of the sort of testimony used to motivate pessimism is Hills' VEGETARIAN:

> Eleanor has always enjoyed eating meat but has recently realized that it raises some moral issues. Rather than thinking further about these, however, she talks to a friend, who tells her that eating meat is wrong. Eleanor knows that her friend is normally trustworthy and reliable, so she believes her and accepts that eating meat is wrong. (2009).

pessimistic intuitions about these cases -- after all, they are cases in which *something* is definitely going wrong.[54]

The emotionist model, then, bifurcates, diverging in what it says about cases like FINK and DEFEAT. One version of the thesis takes cases like FINK and DEFEAT to be troubling; another is happy to grant that these are cases of deference about which we should be sanguine.

My project here has been to establish a theoretically significant connection between pessimism about moral testimony and emotionist inferentialism, not to endorse either. When introducing the emotionist model of transmission failure, I stressed that one might endorse the emotionist model without endorsing its inferentialist commitments, so long as one is happy to allow that the model's inferentialist mechanics codify some potent intuitions. But one might go further than this, and suggest that the emotionist model could be stripped of its distinctively inferentialist commitments, and run on non-inferentialist machinery. As it stands, the emotionist model tells us that having certain affective

[54] This diagnosis of what goes wrong in FINK and DEFEAT shares little with the standard emotionist diagnosis. We ought not regard this discontinuity as a theoretical cost, given the significant structural differences between FINK and MASK and more standard cases like VEGETARIAN.

dispositions is *constitutive* of competence with moral concepts. Someone unhappy with inferentialism will probably be unhappy with this thought, but they might like something considerably weaker, such as the claim that *typical* conceptually competent agents exhibit the sorts of dispositions codified in *Wrong Exit-Condition*, but that a failure to exhibit these dispositions is *compatible* with conceptual competence.[55] An emotionist model constructed around this commitment would lose its distinctively inferentialist flavour but would retain much of its explanatory and predictive power.

3.1 Aesthetic Testimony Pessimism.

In this third section I show that the *de re* strategy -- or something like it -- might be expanded to the aesthetic case. Above, I suggested that some views of *de re* content will predict that audiences lacking 'acquaintance' will be incapable of coming to believe *de re* propositions expressed by mystics. If one thinks that there are *de re* aesthetic propositions -- propositions that are 'directly about' particular novels, or pieces of music, or paintings -- then a theory of *de re* content which exploits a relevantly strict notion of acquaintance will

[55] See Williamson (2007) for general arguments against thinking conceptual competence (even partly) constituted by dispositions to assent or infer.

naturally generate an account on which some aesthetic testimony fails to transmit. And there are influential traditions in aesthetics that are suggestive of something like a relevantly strict notion of acquaintance. For example, the tradition of 'the aesthetic attitude', usually traced from Shaftesbury and Hutcheson through Kant and Schopenhauer, seeks to posit a *distinctively aesthetic* cognitive relation, or mode of perception, between agent and aesthetic object (see Eagleton, 1990; Gaston, 2006; Mitias 1982; Rind, 2002b; Stolnitz, 1961).[56] For Schopenhauer, where ordinary cognition is bound up with the will:

> The momentary silencing of all willing...begins as soon as we give ourselves up to aesthetic contemplation as the pure will-less subject of cognition. (2010).

For Schopenhauer, adopting the aesthetic attitude is not an inevitable consequence of an encounter with a work of art: only the 'genius' can relinquish the will and become the 'pure cognitive subject' (2010). Other strains of the 'aesthetic attitude' tradition do not exploit Schopenhauer's peculiar metaphysics but articulate, using different ideologies, the distinctive cognitive relation we can bear to aesthetic objects. The ideology of *disinterest* -- central to both Shaftesbury's and Kant's accounts -- recurs. Kant, for example, writes that aesthetic

[56] Rind argues *against* the 'widely held' view that the idea of a 'distinctively aesthetic mode of perception', closely associated with 'disinterest', can be traced to Shaftesbury.

'[t]aste is the faculty for judging an object or a kind of representation through a satisfaction or dissatisfaction without any interest' (1987). Stolnitz, characterises the aesthetic attitude as 'disinterested and sympathetic attention to and contemplation of any object of awareness, for its own sake alone'; for Bullough the idea of psychical distance is key: we put an object 'out of gear' with our practical 'needs and ends' (1957).[57] If we think of these aesthetic attitudes as *distinctive modes of cognition or perception*, it becomes easy enough to see how to slot the ideology of the aesthetic attitude into a theory of *de re* content: to have singular thoughts about an art-object *O*, one must have had 'the aesthetic attitude' towards *O*.[58] Someone who has not read and contemplated *Middlemarch*, then, just cannot get into cognitive contact with the *de re* proposition expressed by Virginia Woolf when she says that Middlemarch is one of the few English novels written for grown up people.[59]

[57] Classic criticisms of the idea of the aesthetic attitude are found in Dickie (1964) and Bourdieu, who accuses aesthetic attitude theorists of mistaking the 'experience of a cultured person from a certain social mileu' for an 'ahistorical essence' (1987).

[58] When conjoined with accounts that tie the *de re* tightly to names, this story predicts stronger pessimist intuitions for testimony of the form 'Middlemarch is terribly good' than for testimony of the form 'The first novel of the 1850s written by a woman is terribly good'. But it is unclear that there is any sharp 'semantic rift' between names (and demonstratives) and definite (and indefinite) descriptions (see Hawthorne and Manley, 2012).

[59] Woolf, in the *Times Literary Supplement,* called Middlemarch a 'magnificent book which with all its imperfections is one of the few English novels written for grown-up people' (1919).

On the other hand, it might be better to think of what is going on in the aesthetic case not as an instance of the *de re* but as a related phenomenon with some structurally similar features. We are familiar with the claim that certain sorts of thoughts about objects require a thinker to stand in a special relation to those objects. Theories of the *de re* are special cases of this claim, and so are aesthetic attitude theories (call them *de artem* theories). But the latter need not be thought of as special cases of the former, and we need not think of the peculiar form of contact one stands in with a work of art, in virtue of having contemplated it, as a species of acquaintance in the sense that interests the theorist of the *de re.* Instead, we might think of there being distinctive 'contemplative' cognitive relations in which one can stand to works of art, and specify that an agent can have *de artem* thoughts about *O* only if they have contemplated *O*. Indeed, thinking of things in this second way, in which some of the *structure* of the de re framework is borrowed, but in which aesthetic attitudes directed towards *O* are not classified a species of the *de re*, comes with some advantages. Perhaps you think it implausible that an agent who has bought a copy of *Middlemarch*, and who has read its blurb, and the editor's introduction, but who has not 'contemplated' the novel, cannot have *de re* thoughts about Middlemarch. One obvious response is just to weaken the requirements for an agent to count as having 'contemplated' a work. But one might want to say something else: that such an agent can have singular thoughts about Middlemarch, but

not *de artem* thoughts. If a piece of testimony *t* expresses a *de artem* claim, such an agent will not be able to grasp the *de artem* claim expressed by *t*.[60]

An important aspect of the picture we got when expanding the emotionist model to moral testimony was the machinery explaining why testimonially based moral belief in agents with the 'right' affective dispositions nevertheless look odd. Putting together aesthetic attitude theory with the *de re* (or the *de artem*) model generates similar mechanics: for many proponents of the aesthetic attitude, once one adopts the aesthetic attitude, one is *thereby in a position to make authoritative aesthetic evaluations for oneself* (Atalay, 2007; Fenner, 2003; Rind, 2002a). For example, for Kant, once one considers an object 'disinterestedly', the cognitive powers one brings to bear on an object are the same as those of any other disinterested observer:

> In all human beings, the subjective conditions of this [aesthetic power of judgment], as far as the relation of the cognitive powers therein set into action to a cognition in general is concerned, are the same (1987).

Once one is in a position to entertain *de re* thoughts about an art object, one is also in a position to judge the truth of that *de re* proposition for oneself, and need not rely on

[60] Thanks to Ralph Wedgwood and Timothy Williamson for comments and discussion on these points.

testimony, in much the way that the emotionist model suggests that agents with the right sort of affective attitudes and dispositions need not depend on testimony.

Aesthetic attitude theorists are not in vogue in the academy; nevertheless, I think the thought that there is a distinctive, special attitude -- something like 'disinterested contemplation' -- that one can and ought to take towards works of art is deeply embedded in contemporary Anglo-American culture. The cluster of views that comprise the aesthetic attitude tradition makes it natural for us to think that someone who has failed to have this attitude towards a given work of art is not able to think about it and make judgements about it in the same way as someone who has.[61] The ideology of acquaintance and *de re* propositions offers just one more way of making these claims precise (if not more plausible). Note too that one needn't endorse either the claim that there are any genuine *de re* propositions or the claim that acquaintance hooks up in some theoretically interesting way with singular thought to find the story I am telling plausible. All you need to think is that the ideologies of the *de re* and of acquaintance codify some powerful intuitions about content. Even if you think these intuitions are ultimately best explained some other way,

[61] You might think something like this plays a part in explaining one widely noted feature of predicates of personal taste: that to assert a simple sentence using a predicate like tasty or beautiful typically suggests that the speaker has first-hand knowledge of the evaluated object (Ninan, 2014).

you can still suppose that these intuitions about content structure our intuitions about testimonial exchanges.

To recap, I have developed two novel models of transmission failure that diagnose pessimism as flowing, respectively, from emotionism and from strong views of the *de re*. Importantly, endorsing these models is compatible with rejecting as well as with endorsing the views they identify as the source of pessimist inclinations. I started the paper by saying that I did not mean to compare or demonstrate the superiority of the views sketched here to their alternatives -- my project here is exploratory and suggestive. In the last section of the paper I will consider and respond to an objection to the emotionist model, before moving on to make concluding remarks.

4.1 Gurus, Drugs and Martin Luther King.

Mogensen has argued against accounts of testimonial pessimism which -- like my emotionist model -- propose that pessimist intuitions track testimony's failure to guarantee appropriate affective dispositions in hearers.[62] He asks us to consider the following case:

> GURU: Imagine a guru whose words have hypnotic powers, altering not only a person's moral beliefs but also her character, including her affective dispositions. These hypnotic powers ensure that the beliefs the guru communicates carry with them the appropriate non-cognitive response (Mogensen, 2015).

Mogensen argues that if pessimist intuitions track testimony's failure to guarantee appropriate affective dispositions in hearers, we should expect deference to the guru to look

[62] Mogensen is responding to Howell's 'aretaic model'. Howell argues that we may be suspicious of testimonially based moral beliefs insofar as an agent who relies on testimony for her moral beliefs looks to lack the ability to acquire these beliefs by other means -- for example, by exercising the virtues – and whose beliefs fail to be 'subjectively integrated' or find 'rich support within her [the believer's] own character' (Howell, 2014).

unproblematic. Instead, deference to the guru seems 'more disconcerting' than standard moral deference (Mogensen, 2015).

I agree that the prospect of encountering Mogensen's guru is unsettling. But I do not think he draws quite the right moral from our unsettlement -- it is not true that if pessimist intuitions track testimony's failure to guarantee appropriate affective dispositions in a hearer that deference to the guru should look unproblematic.

To see this, compare GURU first with PERSI:

> Persi is 'preternaturally eloquent'. When you're not around Persi, you're inclined to think that Persi is not particularly trustworthy, and for most p, your conditional credence in *p* conditional on Persi telling you that *p* is around 0.55. But you also know that when you're around Persi, you find him *so incredibly compelling* -- such is his charisma -- for all p, Cr(p/Persi has told me that p) = 1 (Briggs, 2009).[63]

[63] Briggs borrows the case from Maher (1992).

It seems that you should want to avoid Persi. By your present lights, *you will be irrational around him,* for around Persi you respond to evidence in ways that you do not presently endorse. I propose that we can give a common explanation of why we want to avoid both Persi and the Guru.

I want to avoid Persi because around him I will respond to evidence in ways that I do not presently endorse. But suppose one's epistemic dispositions in the 'thin' sense tracked by Briggs' PERSI interact with one's affective dispositions. Empirical work shows that 'fear… can affect low-level visual processes, sad moods can alter susceptibility to visual illusions, and goal-directed desires can change the apparent size of goal-relevant objects' and that mood affects one's reasoning abilities in various different ways: for example, under certain conditions negative emotions appear to 'attenuate' biases during logical reasoning (Zadra and Clore, 2011; see also Goel and Vartanian, 2011; Jung, Wranke, Hamburger, and Knauff, 2015).[64] Once one moves away from a sparse picture of epistemic dispositions on which they are 'sealed off' from their affective counterparts, GURU and PERSI start to look more similar. The details depend on how we capture the interactions between the epistemic and

[64] See also Blanchette (2006, 2014), Bless et al. (1996), Chang and Wilson (2004), Channon and Baker (1994), Clore and Palmer (2009), Dalgleish and Power (1999), Schwarz and Skurnik (2003) & Vroling and de Jong (2009).

the affective. One thing you might say is this: shifts in an agent's affective dipositions will typically be accompanied by shifts in her conditional credences. If we capture the dependence this way, then there is an important affinity between GURU and PERSI. In both cases, the encounter promises to do more than change what our unconditional degrees of confidence are. Instead, an encounter with Mogensen's guru, like an encounter with Persi, will alter our credence functions, and make us respond to evidence in ways that we do not presently endorse.

But you might prefer to think of emotions as constituting, or as making us more likely to conditionalize on non-veridical evidence. If you think this way, GURU looks less like PERSI and more like a third case:

> DRUG: LSQ makes people who have taken it certain of *false* propositions. Conditionalizing on these propositions, they become highly confident that they can fly by flapping their arms (Christensen, 1991).

Right now, you should think that taking LSQ will make you irrational: you think taking LSQ will make you treat as evidence some propositions that right now, you do *not* think

should be treated as evidence (1991). If meeting the guru, like taking LSQ, will incline you to treat as evidence propositions which you presently judge ought *not* to be treated as evidence, we can explain a desire to avoid the guru in much the same way as we explain the desire not to take LSQ.[65]

In light of these comparisons, GURU does not seem to tell us anything distinctive about deference and moral testimony. Accounts which stress the centrality of non-cognitive dispositions to moral judgement do not predict that we ought not be disturbed at the prospect of meeting the guru any more than accounts which stress the centrality of evidence to rationality predict that we should not be disturbed at meeting Persi. Once you grant that non-cognitive and affective dispositions have epistemic import, then someone who cares about being rational will care about what their non-cognitive and affective dispositions are like. And if they think that they have 'good' dispositions right now, then they should want to maintain those dispositions, and not have them displaced by others. So GURU does not show that pessimist intuitions fail to track testimony's failure to guarantee appropriate affective responses: the sort of trepidation we feel at the prospect of encountering

[65] There might be analogies between epistemic self-trust and emotional self-trust that my sketches do not capture. Zagzebski, for example, has argued that arguments for emotional self-trust parallel those for epistemic self-trust (2012).

Mogensen's guru is not *sui generis*, but continuous with the wariness provoked by PERSI and DRUG.

It matters, of course, for how you feels about meeting Persi, that you approve of your current conditional credences. One would -- and *should* -- feel quite differently about Persi if one had a compelling reason to think one's present credence function rationally defective, and to think that spending time with Persi would improve it. (Maybe around Persi you start to find bad arguments less compelling, and good arguments more persuasive.) But it seems that just as one might recognise deficiencies in one's credence function, and seek to change it, one might recognise oneself as emotionally wanting, and seek to cultivate a different set of affective dispositions. (Perhaps around Persi one is moved more acutely by suffering.) Once we grant a connection between affective dispositions, credence functions, and evidence, we should think that agents concerned for their epistemic rationality might *seek out,* as well as fear, novel emotional responses. We often take it to be a mark of *good* political rhetoric and moral discourse that it engages and directs the emotions. In her defence of the role of emotions in liberal politics, Nussbaum writes that Lincoln, Martin Luther King, Gandhi and Nehru were:

> '…great political leaders for their liberal societies because they understood 'the need to touch citizens' hearts and to inspire, deliberately, strong emotions directed at the common work before them' (2013).

Successful moral and political speech channels our sympathies in new ways, towards the oppressed and overlooked. But the prospect of encountering a Martin Luther King ought not to fill us with the same trepidation as the prospect of encountering the guru. It would be a mistake too, to grant as a data point that we shy away from speech which we anticipate will stir our affective capacities in new ways. We can feel that shifts in our emotional temperament might make us better or worse, much as we can feel that shifts in how we are inclined to respond and evaluate evidence might make us more or less rational.

Conclusion.

I began by telling you a story of a failed conversion, and with an account of testimonial pessimism as marking out a cluster of views which stress testimony's epistemic shortcomings. Typical pessimists, I said, offer us pessimisms that come in domain specific flavours. I have suggested a move away from this way of packaging testimonial pessimism: instead of starting with particular cases of apparently defective testimony, one can begin by casting about for

positions in the philosophy of language which *predict* pessimist intuitions in a range of cases. My project has been to establish just such a connection between pessimist readings of testimony and two different commitments one might have in the philosophy of language: emotionism, and strong accounts of *de re* content. I have remained agnostic -- as I said I would -- on the question of whether we ought to use the connection to draft an error theory, or endorse the positions from which pessimism flows.

Contemporary testimonial pessimism tends to fix its gaze on what it treats as relatively discrete domains of discourse; 'the moral' and 'the aesthetic'. But much of our talk is normatively or politically loaded without being the sort of thing we think of as 'moral testimony'. One clear example of this is testimony that exploits slurs. To my ear, slurring testimony can generate something like pessimist intuitions. I don't want to use a 'real' slur; so I borrow a slur from J. K. Rowling: 'mudblood' -- a slur used to denigrate wizards whose parents are non-magical (1999). Malfoy hates mudbloods, but Harry is a committed egalitarian, and, having grown up in isolation from the wizarding world, is relatively untouched by its anti-mudblood ideology. Suppose that Malfoy tells Harry:

(M) 'The mudbloods are in the forest'.[66]

Harry finds Malfoy politically repulsive, but he trusts him to be right about what is going on in the forest. Nevertheless, it seems that Harry (although he may infer and come to know various things about *where* the wizards of non-magical parentage are from the fact that Malfoy has said (M)) *cannot believe what Malfoy says*. I do not want to pretend that it is easy to see what we should make of slurs and their semantics. But suppose you accept something like the following view of slurs, recently argued for by Swanson: slurs 'cue' or 'invoke' ideologies, where ideologies are understood as 'clusters of mutually supporting beliefs, interests, norms, practices, values, affective dispositions, and ways of interpreting and interacting with the world' (m.s.). It is tempting to diagnose the transmission failure between Malfoy and Harry as closely related to their differing ideologies. Harry lacks access to the ideology that Malfoy is embedded in and exploits in his testimony. One way of putting this would be to think of sentences with 'mudblood' as a constituent as 'ideologically infused' guises -- guises an agent only has access to if they have enough of the 'mutually supporting beliefs, interests' and so on, that together make up anti-mudblood ideology.

[66] Some people have argued that slurs simply 'fail to refer' and thus that we should think of Malfoy as having failed to express a proposition. I assume this view is false; for arguments against it, see Williamson (2009).

Doubtless there are other things we could say about what is going on in the slur case: that Malfoy's testimony presupposes or implicates something that Harry does not accept. But these are a new kind of story, not a natural extension of one we already have to hand. It should be considered a strength of the emotionist model of transmission failure that it can be extended to give a common treatment of explicitly normative testimony and more subtly normatively charged speech. Even if you do not share my sense that slurring testimony ought to bring out the pessimist in us, the case should sensitise you to the ways in which testimonially pessimistic instincts have been moulded by the literature into highly organised yet oddly truncated forms.

My approach to testimonial pessimism pictures pessimist intuitions as tracking a kind of *perlocutionary* failure: a failure that arises in cases where our tellings fail to have a certain perlocutionary effect -- belief -- in virtue of an audience's hermeneutic deprivation.[67] Thinking of testimonial pessimism in this way suggests a continuity between the kinds of failure on show when Father Mowbray speaks with Rex, and the more heavily theorised forms of failure we call 'silencing', in which an agent cannot use her words to perform the

[67] See Langton (1993) for the classic statement connecting Austinian speech act theory to feminist concerns about silencing.

action she intends. If I am right to diagnose testimonial pessimism as bound to perlocutionary failure, there is an unnoticed commonality between the sorts of concerns that get tracked, bundled and theorised in the literature on silencing and those that are processed in the literature on testimonial pessimism. Father Mowbray may end up with some unlikely allies.

Acknowledgements.

I am grateful to Max Baker-Hytch, Sam Carter, Lizzie Fricker, John Hawthorne, Alex Kaiserman, Nakul Krishna, Andreas Mogensen, Andrew Peet, Robert Simpson, and three anonymous reviewers for reading and commenting on earlier versions of this paper. An early version of this paper was presented at a workshop on testimony and religious epistemology at the University of Oxford and material from it was presented at the University of St Andrews' Testimony and Context' workshop. I thank my audience on each occasion.

References.

Achinstein, S., 2002. Romance of the Spirit: Female Sexuality and Religious Desire in Early Modern England. *English Literary History* 69 (2): 413–38.

Andow, J., 2014. A Semantic Solution to the Problem with Aesthetic Testimony. *Acta Anal* 30: 211–218.

Armstrong, J. & Stanley, J., 2011. Singular Thoughts and Singular Propositions. *Philosophical Studies* 154 (2): 205–222.

Atalay, M., 2007. Kant's Aesthetic Theory: Subjectivity vs Universal Validity. *Percepi* 1: 44–52.

Beardsley, M. C., 1970. The Aesthetic Point of View. *Metaphilosophy* 1: 39–58.

Bird, A., 1998. Dispositions and antidotes. *The Philosophical Quarterly* 48 (191): 227–234.

Bjorklund, F., Bjornsson, G., Erikson, J., Olinder, R., and Strandberg, C., 2012. Recent Work on Motivational Internalism. *Analysis* 72 (1): 124–137.

Blanchette, I., 2006. The effect of emotion on interpretation and logic in a conditional reasoning task. *Mem. Cogn. 34:* 1112–1125.

- ed., 2014. *Emotion and Reasoning.* New York: Psychology Press.

Bless, H., Clore G. L., Schwarz N., Golisano V., Rabe C. & Wölk M., 1996. Mood and the use of scripts: does a happy mood really lead to mindlessness? *J. Pers. Soc. Psychol.* 71: 665–679. 10.1037/0022-3514.71.4.665.

Boghossian, P., 2003. Blind Reasoning. *The Aristotelian Society sup.* vol 77: 225–48.

Bourdieu, P., 1987. The Historical Genesis of a Pure Aesthetic. *Journal of Aesthetics and Art Criticism* 46: 201–210.

Briggs, R., 2009. Distorted Reflection. *The Philosophical Review* 118(1): 59–85.

Chang, A., & Wilson, M., 2004. Recalling emotional experiences affects performance on reasoning problems. *Evol. Hum. Behav.* 25: 267–276.

Channon, S. & Baker, J., 1994. Reasoning strategies in depression: effects of depressed mood on a syllogism task. *Pers. Indiv. Differ.* 17: 707–711.

Chisholm, R., 1966. *Theory of Knowledge.* New Jersey, Englewood Cliffs: N.J.,Prentice-Hall.

Christensen, D., 1991. Clever Bookies and Coherent Beliefs. *Philosophical Review* 100: 229–47.

Clore, G. L. & Palmer, J. E., 2009. Affective Guidance of Intelligent Agents: How Emotion Controls Cognition. *Cognitive systems research* 10(1): 21-30. doi:10.1016/j.cogsys.2008.03.002.

Cohen, D. & Handfield, T., 2007. Finking Frankfurt. *Philosophical Studies* 135(3): 363–374.

Dalgleish T., & Power, M., eds., 1999. *Handbook of Cognition and Emotion.* New York: John Wiley and Sons Ltd.

Dickie, G., 1964. The Myth of the Aesthetic Attitude. *American Philosophical Quarterly* 1(1): 56-65.

Drier, J., 1990. Internalism and Speaker Relativism. *Ethics* 101: 6–26.

Eagleton, T., 1990. *The Ideology of the Aesthetic.* Oxford: Blackwell.

Easwaran, K., 2009. Probabilistic Proofs and Transferability. *Philosophia Mathematica* 17 (3): 341-362.

Egan, A., 2009. Billboards, Bombs, And Shotgun Weddings. *Synthese* 166(2): 251 - 279.

Evans, G., 1982. *The Varieties of Reference.* Oxford: Clarendon Press.

Fricker, E., 2012. Stating and Insinuating. *The Aristotelian Society supp.* 86 (1): 61-94.

- m.s. On Not Being Able to Believe What One is Told.

Fricker, M., 2007. *Epistemic Injustice.* Oxford: Oxford University Press.

Gaston, S., 2005. *Derrida and Disinterest.* London; New York: Continuum.

Gendler, T., 2010. *Imagination, Intuition, and Philosophical Methodology*. Oxford: Oxford University Press.

Goldberg, S. C., 2001. Testimonially Based Knowledge From False Testimony. *Philosophical Quarterly* 51(205): 512–526.

- 2006. Reductionism and The Distinctiveness of Testimonial Knowledge. In eds., Lackey, J. and Sosa, E., *The Epistemology of Testimony.* Oxford: Oxford University Press.
- 2015. *Assertion*. Oxford: Oxford University Press.

Goel, V. & Vartanian, O., 2011. Negative emotions can attenuate the influence of beliefs on logical reasoning. *Cogn Emot.* 25(1): 121–31. doi: 10.1080/02699931003593942.

Hills, A., 2009. Moral Testimony and Moral Epistemology. *Ethics*, 120(1): 94–127.

- 2013. Moral Testimony. *Philosophy Compass* 8(6): 552–559.

Hopkins, R., 2007. What is wrong with moral testimony? *Philosophy and Phenomenological Research* 74 (3): 611–634.

Howell, R. J., 2014. Google Morals, Virtue, and the Asymmetry of Deference. *Noûs* 48 (3): 389–415.

Jung N., Wranke C., Hamburger K. & Knauff M., 2014. How emotions affect logical reasoning: evidence from experiments with mood-manipulated participants, spider phobics, and people with exam anxiety. *Frontiers in Psychology* 5: 570. doi:10.3389/fpsyg.2014.00570.

Kaplan, D., 1989. Afterthoughts. In eds. Almog,J.. Perry, J. & Wettstein, H., *Themes from Kaplan*. Oxford: Oxford University Press, pp. 565–614.

Kant, I., 1987. *Critique of the Power of the Judgement* , trans. Meredith, J. C.. Oxford: Clarendon Press.

Kierkegaard, S., 1985. *Philosophical Fragments; Johannes Climacus*. Trans. Hong, E. H. and Hong, H. V. Princeton: Princeton University Press.

Kundera, M., 1984. *The Unbearable Lightness of Being*. Trans. Heim, M. H.. London: Faber and Faber.

Lackey, J., 1999. Testimonial Knowledge and Transmission. The Philosophical Quarterly 49(197): 471–490.

- 2008. *Learning From Words.* Oxford: Oxford University Press.

Lakoff, G., 1972. Linguistics and Natural Logic. In eds. Davidson, D. & Harman, G., *Semantics of Natural Language.* Reidel: Dordrecht, pp. 545–665.

Langton, R., 1993. Speech Acts and Unspeakable Acts. *Philosophy and Public Affairs* 22 (4): 293–330.

Lewis D.,1979. Attitudes *de dicto* and *de se*. *Philosophical Review* 88: 508–543.

- 1997. Finkish dispositions. *The Philosophical Quarterly* 47(187): 143–158.

Maher, P., 1992. Diachronic Rationality. *Philosophy of Science* 59: 120–41.

McCann, D. J., ed., 1924. *The Cloud of Unknowing and Other Treatises*. London : Burns, Oates and Washbourne.

McGrath, S., 2009. The Puzzle of Pure Moral Deference. *Philosophical Perspectives* 23(1): 321–344.

- 2011. Skepticism About Moral Expertise as a Puzzle for Moral Realism. *Journal of Philosophy* 108 (3): 111–137.

MacKinnon, R., 2015. Trans*formative Experience *Res Philosophica* 92 (2): 419–440.

Magidor, O., 2015. The Myth of the De Se. *Philosophical Perspectives* 29: 249–283.

Meskin, A., 2004. Aesthetic Testimony: What Can We Learn from Others about Beauty and Art? *Philosophy and Phenomenological Research* 69: 65–91.

Millikan, R., 1990. The myth of the essential indexical. *Nous* 24: 723–34.

Mitias, M. H., 1982. What Makes an Experience Aesthetic? *Journal of Aesthetics and Art Criticism* 41 (2): 157–169.

Moore, A. W., 2006. Maxims and Thick Ethical Concepts. *Ratio* 19 (2): 129–147.

Moran, R., 2006. Getting Told and Being Believed. In Lackey, J. & Sosa, E., eds., *The Epistemology of Testimony.* Oxford: Oxford University Press.

Morgan, J., 1970. On the Criterion of Identity for Noun Phrase Deletion. *Proceedings From the Annual Meeting of the Chicago Linguistics Society ,* vol. 6.

Mogensen, A., 2015. Moral Testimony Pessimism and The Uncertain Value of Authenticity. *Philosophy and Phenomenological Research* 92(2). http://onlinelibrary.wiley.com/doi/10.1111/j.1933-1592.2011.00532.x/abstract.

Moss, S., 2012. Updating as Communication. *Philosophy and Phenomenological Research* 85 (2): 225-248.

Nickel, P., 2001. Moral Testimony and Its Authority. *Ethical Theory and Moral Practice* 4 (3): 253–266.

Nussbaum, M., 2013. *Political Emotions: Why Love Matters for Justice.* Cambridge: Harvard University Press.

Ninan. D., 2014. Taste Predicates and the Acquaintance Inference. *Proceedings of Semantics and Linguistic Theory* 24: 290-309.

Paul, L. A., 2014. *Transformative Experience*. Oxford: Oxford University Press.

Peet, A., 2015. Testimonial Knowledge Without Knowledge of What Is Said. *Pacific Philosophical Quarterly*. doi: 10.1111/papq.12137.

- 2016 Testimony and the Epistemic Uncertainty of Interpretation. *Philosophical Studies* 173 (2): 395-416.

Perry , J., 1979. The problem of the essential indexical. *Nous* 13: 3 -21.

Prinz, J.J., 2007. *The emotional construction of morals*. Oxford: Oxford University Press.

Rind, M., 2002a. Can Kant's Deduction of Taste Judgements Be Saved? *Archiv für Geschichte der Philosophie* 84 (1): 20-45.

- 2002b. The concept of disinterestedness in Eighteenth Century British Aesthetics. *Journal of The History of Philosophy* 40 (1): 67 –87.

Rowe, E. S., 1795. *Devout exercises of the heart.* New-York: Printed by Tiebout & O'Brien for E. Mitchell, book-seller no. 9 Maiden-Lane. http://ezproxy-prd.bodleian.ox.ac.uk:2300/iw-

search/we/Evans?p_action=doc&p_theme=eai&p_topdoc=1&p_docnum=1&p_sort=YMD_date:D&p_product=EVAN&p_text_direct-0=u433=%28%2047583%20%29|u433ad=%28%2047583%20%29&p_nbid=M5DB55NQMTQ2MDQ4MDY4OS42MDA5MTk6MToxMzoxMjkuNjcuMjQ2LjU3&p_docref=.

Rowling, J. K., 1999. *Harry Potter and The Prisoner of Azkaban.* London: Bloomsbury.

Ross, A., 1986. Why Do We Believe What We Are Told? *Ratio* 1: 69-88.

Schwarz N.& Skurnik I., 2003. Feeling and thinking: implications for problem solving. In eds. Davidson J. E. & Sternberg R., *The Psychology of Problem Solving.* Cambridge: Cambridge University Press, pp. 263–292.

Schopenhauer, A., 2010. *The World as Will and Representation.* Ed. and trans. Norman, J., Welchman, A. & Janaway, C.. Cambridge: Cambridge University Press.

Sliwa, P., 2012. In defense of moral testimony. *Philosophical Studies*, 158(2): 175–195.

Smith, M., 1994. *The Moral Problem.* Oxford: Basil Blackwell.

Sosa, E., 1970. Propositional Attitudes *De Dicto* and *De Re*. *Journal of Philosophy* 67 (21) :883-896.

Spearing, A.C., ed. 2001. *The Cloud of Unknowing and Other Works*, London: Penguin Books.

Stolnitz, J., 1961. On the origins of aesthetic disinterestedness. *Journal of Aesthetics and Art Crticism* 20: 131-144.

Swanson, E., m.s. Slurs and Ideology. http://www-personal.umich.edu/~ericsw/research/Swanson,%20Slurs%20and%20Ideologies%20November%202015.pdf.

Vavova, K., 2016. Irrelevant Influences. *Philosophy and Phenomenological Research.* doi: 10.1111/phpr.12297.

Vroling, M. S., & de Jong P. J., 2009. Deductive reasoning and social anxiety: evidence for a fear-confirming belief bias. *Cogn. Ther. Res.* 33: 633–644.

Waugh, E., 2000. *Brideshead Revisited.* London: Penguin Classics.

White, R., 2011. You just believe that because…. *Philosophical Perspectives* 24 (1): 573–615.

Williams, B., 1972. Knowledge and Reasons. In ed. von Wright, G. H., *Problems in the Theory of Knowledge.* The Hague pp.1-11.

Williamson, T., 2009. Reference, Inference and the Semantics of Pejoratives. In eds. Almog, J. & Leonardi, P., *The Philosophy of David Kaplan*. Oxford: Oxford University Press, pp. 137–158.

- 2007. *The Philosophy of Philosophy*. Oxford: Oxford University Press.

Woolf, V., 1919. George Eliot. *The Times Literary Supplement,* 20th November.

Wright, S.,2016. The Transmission of Knowledge and Justification. *Synthese* 193(1): 293-311.

Yablo, S., 2014. *Aboutness*. Princeton: Princeton University Press.

Young, I. M., 2000. *Inclusion and Democracy.* Oxford: Oxford University Press.

Zadra, J. R. & Clore, G. L., 2011. Emotion and Perception: the role of affective information. *WIREs Cogn Sci* 2: 676-685. doi: 10.1002/wcs.147.

Zagzebski, L., 2012. *Epistemic Authority.* Oxford: Oxford University Press.

Abstract – Epistemic Ambivalence

I argue that we ought to do away with the ideology of higher-order defeat.

Here are two observations. First, some epistemic attitudes count as good, and others as bad. Second, not all of our epistemic attitudes are *first order* attitudes; sometimes we have epistemic attitudes towards our own epistemic attitudes.[68] There are different pictures of how these two observations fit together. For some, first-order attitudes tightly constrain which higher-order attitudes we can permissibly hold, or vice versa.[69] For others, the two operate more discretely. Call these, respectively, *integrationist* and *ambivalent* epistemologies.

The ideology of higher-order defeat plays a key role in many contemporary integrationist epistemologies. In this paper, I argue that the ideology of higher-order defeat is paradox prone, and as such the ideology should be rejected, or substantially revised. In the first

[68] Or towards propositions which are *about* our own epistemic attitudes, if you prefer.

[69] To say that which first order attitudes I have constrains which higher-order attitudes I can permissibly hold, is to say that there is a relatively narrow range of second-order attitudes which can be paired with a given first-order attitude without the overall state which contains those attitudes being an impermissible one. To say that which second order attitudes I have constrains which first-order attitudes I can permissibly hold, is to say that there is a relatively narrow range of first-order attitudes which can be paired with a given second-order attitude without the overall state which contains those attitudes being an impermissible one. In other words, the constraints in question generate wide, rather than narrow, scope requirements.

section of the paper, I motivate the ideology of higher-order defeat, and specify the theoretical roles it is supposed to play. In the second section, I show that we can derive contradictions from apparently innocuous claims specifying the theoretical roles played by higher-order defeat. In the third section, I demonstrate the generality of the tensions sketched. Last, I sketch and respond to some objections to my view.

1.1 Motivating Higher-Order Defeat

Suppose you like – as many do – a principle along the following lines:

Single Premise Closure (SPC):

Necessarily, if S

(i) knows that *p,*

(ii) competently deduces *q* from *p,*

(iii) comes to believe *q* solely on the basis of this deduction, and

(iv) has throughout no relevant rebutting or undercutting defeater,

then S knows *q*.[70]

Sometimes, we ought to doubt that we are capable of making competent deductions. Classic sharpenings of this worry invoke the condition of hypoxia: 'a condition brought upon by high altitudes. It makes sufferers prone to errors in their reasoning…In mild cases, it is introspectively undetectable' (Schechter, 2013). Once you are in a situation in which you realise you are likely to be suffering from hypoxia, the thought goes, you ought not to have much confidence in the output of your deductions:

> HYPOXIA: Sam is climbing a very large mountain, alongside Norgay, a much more experienced mountaineer. Sam is not primarily responsible for navigation, but he knows that they need to keep heading in some direction between 44.6 degrees and 46.3 degrees. Once they reach a large cairn of white stones, Sam's more experienced companion knows that they are at an altitude where Sam -- inexperienced at high altitudes -- is very likely to be suffering from hypoxia. He tells Sam this, and Sam comes to know that he very likely has hypoxia. But Norgay wants to give Sam some

[70] To my way of thinking, *d* is a *rebutting* defeater for a belief that *p* just if Cr(*p*/*d*) but not Cr(*p*) is below some relevant threshold. One acquires an *undercutting* defeater *d** for a belief that *p* where one believes *p* on the basis of an inference from *q*, and Cr(*q*/*d*) but not Cr(*q*) is below some relevant threshold.

> practice at navigating, and asks Sam whether they ought to be heading in some direction between 43.9 and 46.3 degrees. Sam competently deduces that they should be heading in such a direction, and comes to believe, solely on the basis of his deduction, that they ought to be heading in some direction within this margin.[71]

Does Sam get to know the output of his competent deduction? The standard judgement seems to be that he does *not* – that his competent deduction from a known premise fails to be knowledge producing.[72] He *ought,* the thought goes, to think that he is unlikely to be making competent deductions, and this prevents him from knowing what he truly believes – that they ought to be heading in some direction between 43.9 and 46.3 degrees.[73] But if this is right, we look to have a counterexample to SPC, for on standard ways of taxonomising defeat, Sam acquires neither a rebutting nor an undercutting defeater.[74]

---

[71] Many others rely on similar cases, including Christensen (2010a), Elga (m.s.) and Schechter (2013).

[72] See, for example, Schechter (2013).

[73] There are tricky questions about what we ought to say in cases in which an agent A (i) believes *p* on the basis of some method M, (ii) ought to think *m* is a bad method and (iii) ought to think they believe *p* on the basis of some other, good, method. I bracket complications of this sort for the rest of the paper.

[74] Taxonomies of defeat which exploit different bits of background theory will not always produce perfectly overlapping categories. But an intuitive way of distinguishing higher-order defeat from its fellows begins with a contrast between two different ways of being fallible. On the one hand, the way the world presents itself to you may be misleading. But on the other hand, you might fail to respond to the world as you ought to, even given how it is presented to you (Christensen, 2010a). If rebutting and undercutting defeat map onto the first way of being fallible, higher-order defeat maps onto the second. And if we have in place the ideology of deductive inference as a *belief forming method*, we can draw the distinction in these terms. One thing you

There are (at least) two ways to respond. One can stick to SPC and tell some story about why we ought not to care too much about the intuitions HYPOXIA generates. This is the strategy I associate with ambivalent epistemologies.[75] But if we want to vindicate the standard intuition that HYPOXIA elicits – that Sam's deduction is not knowledge producing – then we need to revise SPC, which, as it stands, predicts that Sam's deduction *is* knowledge producing. Call this the integrationist strategy.

A natural way of tweaking SPC will leave you with something like this:

SPC*: Necessarily, if S

(i) knows that *p,*

(ii) competently* deduces *q* from *p,*

(iii) thereby comes to believe that *q* solely on the basis of this deduction,

---

might think is that the input to a deduction, or the output of a deduction, was somehow defective. We can tie undercutting and rebutting defeat to this sort of thought. Another thing you might think is that the deduction itself, rather than its input or output, was faulty. We can tie higher-order defeat to this sort of thought. Given either of these ways of taxonomising defeat, Sam should not be thought of as having a rebutting or an undercutting defeater. But note that these glosses need not be taken too seriously. So long as we have some characterisation of the varieties of defeat according to which an agent *can* have a higher-order defeater without having any other kind of defeater, there will be a problem for the proponent of higher-order defeat. I return to this point later in the paper.

[75] For examples of this sort of strategy, see Lasonen-Aarnio (2010) and Williamson (2009).

(iv) has no higher-order defeater for the claim that his deduction was competent,

then S knows that *q*.

For the sake of brevity, we can pick out a particular operation – competent* deduction – where S competently* deduces *q* iff S competently deduces *q* from a known premise, comes to believe *q* solely on this basis, and S throughout has no relevant rebutting or undercutting defeater. I will use this terminology for much of the paper.

SPC* is consistent with thinking that in *some* cases, S comes to know that *q* on the basis of competent* deduction even whilst having an undefeated higher-order defeater for the claim that his deduction was competent. But it would be odd to have this combination of views: much of the literature on defeat, tacitly or otherwise, understands the theoretical function of (undefeated) defeat as follows: (undefeated) defeat is something that stops one from having knowledge:

> "[D]efeatists" tend to agree that there is an additional condition, beyond a belief's being non-Gettiered, justified, and true (or perhaps built in as a condition on

justification), which must be fulfilled for that belief to count as knowledge." (Baker-Hytch & Benton, 2015.)[76]

As a first pass we might codify this thought by saying *necessarily, if S knows that q, then S does not have an (undefeated) higher-order defeater for q.* But this may well not do, given the hybrid etiology of certain beliefs:

QUIZ: You are taking part in a quiz. Some of the questions require you to do mental arithmetic. You need to do the sum 'What is 989 + 72?' in your head, and you try to do it more quickly than any of your competitors. But at exactly the moment that you work out that 989 + 72 = 1061, a competitor shouts out '1061!'.

It is tempting to think that the belief you end up with, that 989 + 72 = 1061, is not based *just* on your deduction, or *just* on your competitor's answer. And maybe you ought to think your *own* deductive capacities intact, but that your competitor's answers are wholly unreliable. It's not clear how we should theorise such cases. But we can screen them out by

[76] From hereon, whenever I say 'defeater' I mean 'undefeated defeater' unless otherwise stated.

restricting our attention to cases involving beliefs whose etiologies lack hybridity. For the rest of the paper I will be speaking only about beliefs with non-hybrid etiologies.

> DEFEATISM: Necessarily, if S knows that *q*, and believes *q* solely on the basis of some method M, then S does not have an (undefeated) higher-order defeater for *q*.

SPC* is also consistent with thinking that an agent has a higher-order defeater whenever the degree of confidence it is rational for one to have that one's deductions are competent is less than 1. But if this is how the notion of higher-order defeat gets filled in, SPC* will look to rule out the possibility of creatures like us acquiring knowledge via deduction. For at least given standard ways of construing what it is to have a given degree of confidence, none of *us* ought to have a degree of confidence of 1 that our deductions are competent. It follows that we ought to think that an agent having some non-zero rational degrees of confidence in their deductions being incompetent is compatible with said agent being higher-order defeat free.

In the next section of this paper I show that there is a deep-running tension between DEFEATISM and SPC*. Each of these theses can be thought of as encoding one role that

higher-order defeat is expected to play. What I show is that it *cannot play both* of these roles. I take these tensions to suggest that there are significant structural problems embedded in talk of higher-order defeat.

2.1 Tensions

Let us consider a particular case, ORACLE:

> ORACLE: Suppose that an oracle tells Watson 'You are bad at deduction'. Call the proposition she tells Watson *'SHERLOCK'*. Watson believes the oracle, and comes at t_1 to have testimonially based knowledge that *SHERLOCK*. Later, at t_2, Watson forgets his exchange with the oracle, but continues to believe, and know, the proposition she told him. At t_3, Watson competently* deduces from *SHERLOCK* that *SHERLOCK*, and comes to believe that *SHERLOCK* solely on the basis of this competent* deduction. Other than his belief that *SHERLOCK,* Watson has no evidence concerning his cognitive capacities.[77]

[77] You might think there is something suspicious about the idea of deducing a proposition from itself. I address this worry in the final section of the paper.

For the mechanics of the case to work it is important to remember that being bad at deduction does not mean that *all* of one's deductions are incompetent. Suppose you want to resist this assumption: you think that if someone is bad at deduction, then *none* of their deductions are competent. If you think this, then you should think it impossible that claims like 'Exactly 95% of Jones' deductions are competent' are true. But these seem like claims that could easily be true. But if 95% of your deductions are incompetent, surely you are bad at deduction! So it may be the case that Jones is bad at deduction but, nevertheless, that 5% of his deductions are competent. Consider too that someone might count as bad at making deductions if one third of their attempts at something very simple (like conjunction elimination) go wrong. But surely one should not count as incapable of performing competent deductions just because one third of one's deductions go badly. Deduction, is then, a *gappy* operation: being bad at it does not mean one can never competently perform it.

We can represent the space of possibilities for Watson in ORACLE at t_3 as looking something like this:

a. Watson knows that *SHERLOCK* and has a higher-order defeater for *SHERLOCK*.	c. Watson does not know that *SHERLOCK* and has a higher-order defeater for *SHERLOCK*.
b. Watson knows that *SHERLOCK* and doesn't have a higher-order defeater for *SHERLOCK*.	d. Watson does not know that *SHERLOCK* and does not have a higher-order defeater for *SHERLOCK*.

2.2. Any theory of higher-order defeat that commits us to being in the space represented by (a) or (d) is no good. To see this, consider the following pair of arguments.

Argument 1

Assume that Watson knows that *SHERLOCK* and has a higher-order defeater for *SHERLOCK*. Given that S has a higher-order defeater for *SHERLOCK*, it follows from DEFEATISM that either *SHERLOCK* has a hybrid etiology, or S does not know that

SHERLOCK. But from the description of the case, *SHERLOCK* does not have a hybrid etiology: S believes *SHERLOCK* solely on the basis of deduction from a single premise. So S does not know *SHERLOCK*. We have a contradiction: we must reject the assumption that S knows that *SHERLOCK* and has a higher-order defeater for *SHERLOCK*.

Argument 2

Assume that Watson does not know *SHERLOCK* and does not have a higher-order defeater for *SHERLOCK*. From the description of the case, Watson has formed belief in *SHERLOCK* solely on the basis of competent* deduction from a known single premise. So by SPC* either Watson knows *SHERLOCK* or has a higher-order defeater for *SHERLOCK*. But then it is false that Watson does not know *SHERLOCK* and does not have a higher-order defeater for *SHERLOCK*, and we must reject our starting assumption.

Locating ourselves in the logical space represented by either (a) or (d) generates a contradiction. As such, if we want to endorse both SPC* and DEFEATISM, we must locate ourselves elsewhere in logical space.

It follows that if we want to defend the ideology of higher-order defeat, we need to locate ourselves in the space represented by (b) or (c).[78] Call these positions *weak integrationism* and *strong integrationism* respectively. I will show now that neither weak nor strong integrationism is viable. Let us take them in turn.

2.3 Weak Integrationism

Weak integrationism allows for S to know that *p* without having a higher-order defeater for *p*.

The trouble with this position is as follows. If one can *know* that one is very likely to have hypoxia and yet fail to have a higher-order defeater for deduced beliefs, it is difficult to see that higher-order defeat can do the theoretical work that its proponents want it to do, and which motivates the ideology in the first place. This is easiest to see if we look at a so-called (by Horowitz) 'paradigm' case of higher-order defeat:

[78] Note that the space represented by (b) is open to you only if you reject B, and the space represented by (c) is open to you only if you reject C.

SLEEPY DETECTIVE: Dan is a police detective, working to identify a jewel thief. He knows he has good evidence—out of the many suspects, it will strongly support one of them. Late one night, after hours of cracking codes and scrutinizing photographs and letters, he finally comes to the conclusion that the thief was Lucy. Dan is quite confident that his evidence points to Lucy's guilt, and he is quite confident that Lucy committed the crime. In fact, he has accommodated his evidence correctly, and his beliefs are justified. He calls his partner, Alex. "I've gone through all the evidence," Dan says, "and it all points to one person! I've found the thief!" But Alex is unimpressed. She replies: "I can tell you've been up all night working on this. Nine times out of the last ten, your late-night reasoning has been quite sloppy. You're always very confident that you've found the culprit, but you're almost always wrong about what the evidence supports. So your evidence probably doesn't support Lucy in this case." Though Dan hadn't attended to his track record before, he rationally trusts Alex and believes that she is right—that he is usually wrong about what the evidence supports on occasions similar to this one (Horowitz, 2013).

If this is a case in which higher-order defeat can explain why Dan fails to know that Lucy is guilty, then it must be the case that Dan is in a *good enough* epistemic position with respect to the proposition Alex asserts -- call it *FATIGUE* -- for it to act as a higher-order defeater. But if we have a story about defeat that tells us that Dan isn't in a good enough epistemic position with respect to *FATIGUE* for it to have defeating force even when he knows *FATIGUE*, it looks like we will have trouble maintaining a story that tells us that Dan is in the right sort of position to have a higher-order defeater for the claim that Lucy is guilty. But if we are so strict about what counts as having a (higher-order) defeater that S does not have one in ORACLE even though S knows they are bad at making deductions, it seems hard to maintain that Dan is in a good enough epistemic position with respect to *FATIGUE* for *FATIGUE* to act as a higher-order defeater. But if *FATIGUE* does not function as a higher-order defeater for Dan, then higher-order defeat won't be able to play the theoretical role its proponents want it to. It won't be a way of explaining or theorising our intuitions in *precisely those cases* that are supposed to motivate the integrationist picture.

2.4 Strong Integrationism

Suppose instead we try occupying the space given to us by (C). In Dan's case, occupying position (C) amounts to saying that Dan is in a good enough epistemic position with respect to *FATIGUE* for it to have defeating force, but that Dan does not know *FATIGUE*. On the face of it, this looks like an appealing story. There does not seem to be very much that is troubling about thinking that some proposition can act as a defeater for an agent even when they are not in a strong enough epistemic position to *know* that proposition. But if we think through the implications of this view more carefully, we run into trouble. There are a few different arguments for this thought. Let us take them in turn.

2.5. Against Strong Integrationism

I take it that most people's motivation for buying into the ideology of defeat has to do with a desire to *avoid dogmatism* – to accommodate the thought that our epistemic states are *unstable.* A core anti-dogmatic commitment is this: a belief that is knowledge constituting today may *not* be knowledge constituting tomorrow. More generally, we often shift from good to bad and from bad to good epistemic positions with respect to particular claims (or so it seems). Part of believing *as one ought* is being responsive to these shifts.

It is natural, then, that people who like to talk about defeaters generally take it that defeaters, like knowledge, can be lost: defeaters themselves can be defeated.[79] Here is one way of making this thought precise:

> DEFEAT-DEFEAT: If A believes some proposition *p* at t, any *d* such that *d* acts a defeater for this belief that *p*, is not itself defeated.

DEFEAT-DEFEAT codifies the idea that learning new information can *ameliorate* the effect of a defeater on a subject's epistemic position. Rejecting DEFEAT-DEFEAT means a strange dogmatism, where defeaters, once acquired, continue to exert their defeating force regardless of further shifts in an agent's epistemic position. The point is not just that DEFEAT-DEFEAT sounds like a compelling principle. Rather, the ideology of defeat is only well placed to do the work its proponents want it to do – avoid dogmatism – if it comes as part of a package with DEFEAT-DEFEAT. Otherwise defeaters – and remember, defeaters stop you from knowing – are something one is *stuck* with once acquired. Without

[79] See Bergmann (2006), Bernecker (2011), Cohen (1987), de Ridder (2014), Fumerton (1995), Goldman (1986), Greco (2010), Lackey (2005, 2006), Plantinga (1993).

DEFEAT-DEFEAT, defeaters become a way of *generating* dogmatism, rather than avoiding, dogmatism.

Let us think through the implications of DEFEAT-DEFEAT for strong integrationism. The strong integrationist tells the following story about Watson's situation: Watson (i) does not know *SHERLOCK* but (ii) is in a sufficiently good position, with respect to *SHERLOCK*, for *SHERLOCK* to act as a defeater.

Suppose we combine this story with DEFEAT-DEFEAT. It follows that no *d* acts as a defeater-defeater for *SHERLOCK*. But if Watson is in a sufficiently good epistemic position with respect to *SHERLOCK* for it to act as a defeater for Watson, then presumably *SHERLOCK* will have defeating force with respect to any proposition that Watson believes on the basis of deduction. But Watson believes *SHERLOCK* on the basis of deduction, so *SHERLOCK* will have defeating force for *SHERLOCK*. So there is some *d* such that *d* is acting as a defeater-defeater for *SHERLOCK*. We have a contradiction: it follows from the conjunction of DEFEAT-DEFEAT and the strong integrationist reading of Watson's situation that there is no proposition *q* such that *q* is acting as a defeater for *SHERLOCK*.

But this is too fast. For a defender of higher-order defeat might argue as follows. Suppose we grant DEFEAT-DEFEAT. The above argument, the defender might argue, assumes that if Watson is in the right sort of epistemic position for some *d* to act as a defeater for Watson's belief that *SHERLOCK*, then he is in the right sort of epistemic position for *d* to act as a *defeater-defeater* for *SHERLOCK*. But DEFEAT-DEFEAT just tells us that if *d* is to act as a defeater for a given belief, *d* is not itself defeated. It does not tell us *what it would take* for *d* itself to be defeated.

So long as one notes this, one is free to say the following: after performing the relevant deduction, Watson is in the right sort of position with respect to *SHERLOCK* for it to act as a defeater for his beliefs formed on the basis of competent deduction, but not in the right sort of position with respect to *SHERLOCK* for it to act as a defeater-defeater for such beliefs. For example, suppose that Watson, on the basis of deduction, ends up with a degree of confidence of 0.7 in *SHERLOCK*, and that having this degree of confidence in *SHERLOCK* is enough for Watson to have a defeater for *SHERLOCK*. So long as having a degree of confidence of 0.7 that he is bad at deduction is not enough to prevent *SHERLOCK* – which is, recall, believed on the basis of deduction – from acting as a

defeater, there will not be a problem.[80] The higher-order defeatist has a decent line of defence. Let us, then, look at a different argument.

2.6 Rational Degrees of Belief

If you are the sort of person who thinks that the notion of rational degrees of belief is an important one in epistemology, then the following two views should be very natural:

1. Just when the degree of confidence that it's rational for me to have in q is greater than or equal to n (where $0 \leq n \leq 1$) can q act for me as a higher-order defeater.

2. There is some degree of confidence k such that I know that p only if my rational degree of confidence in p is greater than or equal to k (where $0 \leq k \leq 1$).

If you like both of these and single premise closure; you should be inclined to endorse:

RDB: Necessarily, if,

(i) S's rational degree of confidence in p is n and

[80] Thanks to Ralph Wedgwood and Timothy Williamson for discussion and written comments on this point,

(ii) S competently deduces *q* from *p* and

(iii) S does not acquire a relevant higher-order, rebutting or undercutting defeater, then S's rational degree of confidence in *q* is equal to or greater than n.[81]

To see this, consider the following argument. If you don't endorse RDB, it looks like you will end up with counterexamples to single premise closure (both SPC and SPC*). Suppose my rational degree of confidence in the proposition that some cats are striped is *k* (recall: I know some proposition only if my rational degree of confidence in that proposition is greater than or equal to k). Suppose I competently deduce that some animals are striped. Suppose too that I have no relevant defeaters. Unless my rational degree of confidence in the deduced proposition is greater than or equal to *k* I will fail to know the deduced proposition, and we will have a counterexample to single premise closure (again, both SPC and SPC*).

With this backdrop in place, let us turn to a variation of HYPOXIA.

[81] You might think that RDB holds only in cases where $n > k$. This would be a very strange view, at least for anyone who wants to understand rational degrees of belief in terms of knowledge, rather than vice versa.

HYPOXIA*: One begins such that one's rational degree of confidence in the proposition that one is very likely to have hypoxia (*'POXY'*) is equal to *n*. That is, let it be the case that $Cr_1(POXY) = n$. One competently deduces *POXY* from itself and comes to believe *POXY* solely on the basis of this deduction.

Either $Cr_2(POXY) \geq n$ or $Cr_2(POXY) < n$. Suppose the former is the case. Then one has a defeater for *POXY*. But then we have trouble: surely one's rational degree of confidence should be lower after one acquires a defeater. So $Cr_2(POXY) < n$, and we have a contradiction. So suppose the latter is the case. If my rational degree of confidence in *POXY* at t_2 is below *n*, then I lack a higher-order defeater for *POXY*. But if this is the case, then it seems that we have a counterexample to RDB – and we already have a sense of the costs involved in rejecting *that*. In other words, so long as there is some minimum rational degree of confidence that an agent must have in a proposition for it to act as a higher-order defeater, and we grant that competent* deductions preserve these degrees of confidence in the absence of higher-order defeaters, the defender of higher-order defeat will be in

trouble.[82] The upshot of these considerations is that there is no position in logical space that the integrationist can occupy that makes their joint endorsement of SPC* and DEFEAT look feasible.

3. Objections and Replies

The paper so far has sketched the prospects for endorsing both SPC* and DEFEATISM – two theses which each codify one role that higher-order defeat is supposed to play. In the next section of the paper, I will sketch and respond to two objections.

3.1 Generality

In the face of the arguments so far, one might be inclined to argue as follows: 'you have introduced higher-order defeat by way of closure (SPC and SPC*) principles, and characterised its theoretical role in terms of its interaction with closure principles. If

[82] It is worth noting that this argument may be run whilst treating 'rational degrees of belief' as a placeholder for any doxastic attitude or property that comes in degrees and which has the requisite connections to defeat and competent deduction.

someone likes the ideology of higher-order defeat but was never tempted by closure principles, they should remain sanguine about the prospects for higher-order defeat'.

But things are not quite so rosy for the higher-order defeatist: they cannot avoid the tensions charted in this paper by jettisoning SPC*. The dynamics that emerge when we trace the interaction of DEFEATISM and SPC* replicate elsewhere in epistemology and should trouble even those epistemologists who are perfectly cavalier about abandoning closure principles.

Consider: if higher-order defeat is to do any important explanatory work it should be the case that there are some ways of acquiring beliefs, or assigning degrees of belief that come with some epistemically significant property, (such as knowledge) just so long as an agent lacks a higher-order defeater, but are not knowledge producing otherwise. In other words, there is pressure for the fan of higher-order defeat to endorse *some* theory with something like following general form:

> INTEGRATIONISM: Necessarily, if S believes that (assign a degree of confidence of *n* to) *p* solely on the basis of competently using method M, then just so long as S

has no relevant higher-order defeater for *p,* S knows that (has a rational degree of confidence in) *p*.[83]

If no theory with the general form of INTEGRATIONISM is true, it is unclear what theoretical or explanatory work is actually being done by the ideology of higher-order defeat.

The trouble is, wherever *m* is gappy – that is, such that an agent may come to truly believe, *on the basis of competently M-ing,* that they ought not to trust method M's outputs – there will be the potential for problems that are structurally isomorphic to those that arise when one tries to endorse SPC* and DEFEATISM. To see this more clearly, let us look in a little more detail at two different cases. First up is the case of memory.

3.2 Memory.

[83] Two points. (i) Obviously the integrationist is not under pressure to accept *all* theories with this general form. (ii) Earlier in the paper I claimed that exactly how one likes to distinguish between higher-order defeat and other species of defeat would not, as far as the arguments of this paper go, matter that much. We are now in a position to see why. So long as you think that higher-order defeaters can come apart from other sorts of defeaters, there will be cases in which INTEGRATIONISM style theses predict that an agent knows some proposition just if they lack a relevant higher-order defeater.

Epistemologists who think about memory are often inclined to endorse their own 'transmission' principles. Consider the following appealing story about memory:

MEMORY: Necessarily, if

(i) S knows that *p* at t_1,

(ii) *p* is competently stored in S's preservative memory until t_2 and

(iii) S believes that *p* at t_2 solely on the basis of the output of his preservative memory that *p,* and

(iv) at t_2 S has no relevant rebutting, undercutting or higher-order defeater for *p,*

then S knows that *p*.[84]

It should be easy to see that this has the same general structure as INTEGRATIONISM.

We noted above that deduction is *gappy*: one can be generally bad at deduction and nevertheless *sometimes* manage a competent deduction. This gappy structure is not peculiar to deduction. Memory is gappy too: the arguments given for thinking that an agent may, on

[84] Principles along these lines are defended by Blaauw (2008), Bernecker (2008), and Burge (1997) among others.

a given occasion, perform a competent deduction, even if they are generally bad at deductions apply, *mutatis mutandis*, to memory. One might be generally bad at encoding information in preservative memory, but properly encode in preservative memory the information that *you are generally bad at encoding information in preservative memory.*[85] Consider the following case:

> GURU: On Monday morning, a guru, who knows that Hercule's preservative memory is terrible, tells Hercule that his preservative memory is terrible. Hercule comes to know this on Monday morning, and competently encodes the information. On Tuesday evening, he comes to believe, on the basis of his preservative memory, that his preservative memory is terrible. Hercule has acquired no rebutting or undercutting defeaters for this claim since Monday.

If we like MEMORY, it follows that on Tuesday evening either Hercule knows that his memory is terrible and lacks a relevant higher-order defeater for this claim, or has a higher-

[85] Alternatively, consider the following. One might know, in the morning, that there is a 95% chance that by the evening, you will have taken a memory altering drug, and a 5% chance you will have taken the placebo. It seems that you can know, in the morning, that in the evening, you should probably think your memories of the day inaccurate. You take the placebo and your memory is functioning perfectly. In the evening, you remember what you knew that morning: that you should probably think your memories of the day inaccurate.

order defeater for said claim, and fails to know that his preservative memory is terrible. But against each of these options, we can run an argument structurally analogous to the arguments run against their counterparts above – weak and strong integrationism, respectively. Let us take them in turn.

The weak integrationist will say that Hercule *knows* that his memory is terrible, but lacks a higher-order defeater for claims he believes on the basis of preservative memory.[86] But this looks bad for the same reasons, explored above, that it looks bad to think that, in SLEEPY DETECTIVE, Dan's knowing that his laye-night reasoning tends to be very sloppy fails to give him a higher-order defeater.

In the case of deduction, we saw that adopting the strong integrationist strategy threatened to undermine plausible stories about the connections between competent* deduction and rational degrees of belief. We will encounter a similar problem here. For if you are attracted to MEMORY, you should like the claim that if at t, Hercule has a rational degree of confidence of *n* in some claim *p*, and he encodes this information in preservative memory, if

[86] Recall that are abstracting away from complications that arise given an agent's uncertainty about the source of their beliefs.

Hercule, at t_2, comes to have a degree of belief in *p* on the basis of said memory, then, absent relevant defeaters, his rational degree of confidence in *p* will be no less than *n.* We can then fill in the argument as follows. Suppose that on Monday morning, Hercule has a rational degree of belief of *m* in the claim that his preservative memory is terrible, where *m* is such that if one's degree of belief in the claim that one's memory is terrible is any lower than *m*, one lacks a relevant higher-order defeater for claims one endorses on the basis of memory.[87] Hercule competently encodes this information in preservative memory. Then, on Tuesday afternoon, Hercule assigns a degree of beliefto the claim that his preservative memory is terrible on the basis of his preservative memory. He has acquired no rebutting or undercutting defeaters for the claim since Monday morning. By familiar reasoning, there will be trouble regardless of whether Hercule ends up with a degree of belief of at least *m*, or ends up with a degree of belief of less than *m.*

3.3 The Tychometer

Let us take stock. We have argued that if higher-order defeat is to do the theoretical work claimed for it, something along the following lines must be true: there must be some

[87] The term 'endorse' here ought not to be taken as imposing any lower limit on *m.*

procedure for assigning degrees of confidence to propositions such that the procedure generates, say, rational degrees of confidence in just those cases in which an agent lacks a higher-order defeater.

With this in mind, suppose that one has access to a *tychometer*. A *tychometer* is a machine which, for any p, infallibly tells you which degree of confidence it is rational for you to have, given one's first-order evidence. Just so long as you have no relevant higher-order defeaters, the tychometer will tell you what your rational degree of confidence in *p* is. The tychometer has a microphone and a screen. If you speak a sentence *s* into the microphone, the tychometer will display a number between 0 and 1. After you say a sentence s to the tychometer, whichever *n* the tychometer displays, that is the degree of confidence you should have in the claim articulated by s, just so long as you have no relevant higher-order defeater.

What would count as a higher-order defeater in such a case? Well, one might be highly confident that one is misreading the numbers displayed on the tychometer. If it seems to you that, for proposition p, the tychometer is displaying a <0.7> but you are pretty sure that you are misreading the tychometer, assigning a degree of confidence of 0.7 that *p* does not look

rational. More generally, if one has a higher-order defeater which is relevant to the tychometer's outputs, you should not assign to *p* the degree of confidence the tychometer seems to tell you you should.

Now, plausibly, there is some minimal degree of confidence one must have in the claim that one is misreading the tychometer in order to count as having a higher-order defeater. Let *m* be such that one must have a degree of beliefof at least *m* in the claim that you are misreading the tychometer if one is to have a relevant higher-order defeater. Consider the following case:

TYCHOMETER: Hercule says the following sentence into the TYCHOMETER: 'I will misread the numbers that appear on the tychometer's screen.' The tychometer displays the number m, and Hercule competently reads the screen as assigning *m* to the relevant claim, *MIS*.

What should Hercule do? Suppose he assigns a degree of belief of less than *m* to MIS.[88] Then something has gone wrong: he has failed to assign the degree of belief given by the tychometer, despite not having a relevant higher-order defeater (he and so he has assigned an irrational degree of confidence). Suppose he assigns a degree of belief of *m* to *MIS.* Then he has a relevant higher-order defeater, and so, in assigning *m*, he has assigned an irrational degree of confidence to *MIS.* It follows that, if he is to assign a rational degree of confidence to *MIS,* Hercule must assign to *MIS* a degree of confidence that is *higher than m.*

But things get even more peculiar if we turn our attention to a slightly more sophisticated tychometer: the tychometer 2.0. The tychometer 2.0 behaves exactly like the first tychometer, except that its screen has been cleverly engineered in such a way that it makes it much less likely that the numbers on the screen will be misread. But this new technology comes with a bug: with the first tychometer, if a user did misread the screen, they were just as likely to interpret the screen as showing a higher number than it actually did as to interpret the screen as showing a lower number than it actually did. But the tychometer 2.0's screen makes it impossible to misread the screen 'up': if someone misreads the screen, they invariably interpret the number it shows as lower than the screen it *actually* shows.

[88] This, of course, is only possible if $m > 0$.

Now, suppose Hercule has grown weary of misreading his tychometer, and bought a tychometer 2.0:

> TYCHOMETER 2.0: Hercule says the following sentence into the TYCHOMETER 2.0: 'I will misread the numbers that appears on the tychometer's screen.' The tychometer displays the number *m*, and Hercule competently reads the screen.

What should Hercule do?

This time, Hercule can be certain that the tychometer either says 'm' or some number smaller than 'm': if he has misread the screen, he is definitely 'downreading', rather than 'upreading'. But he is failing to assign a rational degree of belief if he assigns a degree of belief of *m* to MIS: if he has a degree of belief of *m* that he is misreading the tychometer, he has a relevant higher-order defeater, and should not have the degree of confidence that the tychometer appears to recommend to him. But if he has a degree of confidence of less than *m*, something has gone wrong, for he is failing to have the degree of confidence the

tychometer recommends to him, despite failing to have a higher-order defeater. But it does not seem that in this case, assigning a degree of confidence of more than *m* to *MIS* is a way of assigning a rational degree of confidence. For in this case, Hercule can be *sure* that, given his first order evidence, his degree of confidence in *MIS* should be no more than *m.* Consider: if one has evidence that one is not evaluating one's first order evidence correctly, it may well be rational, all things considered, to refrain from having the degrees of confidence which are *in fact* made rational by one's first-order evidence. But one may have evidence that one is not evaluating one's first-order evidence correctly *and* be rationally certain, all things considered, that correctly evaluating one's first order evidence will *not* have degrees of beliefwithin in a certain range R. In other words, *whether or not one has a higher-order defeater for p*, the degree of confidence one assigns to *p* should never be such that one is *sure* one's first-order evidence does not license that assignment.[89] How does this apply to Hercule? In Hercule's case, he is rationally certain that his first-order evidence requires him to avoid having a degree of belief of greater than *m* in *MIS.* If it is impermissible for him to have a degree of confidence which he is *certain* his first order

[89] Such a claim should be particularly appealing to integrationists.

evidence does not license, then it is not rational for him to have a degree of confidence of greater than *m* is MIS.[90]

The tension between SPC* and DEFEATISM, then, is not properly diagnosed as a function of any idiosyncratic feature of SPC*. Instead, the tension flows from (i) the fact that SPC* encodes higher-order defeat as capable of playing a particular theoretical role – a theoretical role that it must be able to play if it is to have the sort of interest its advocates claim – and (ii) the fact that agents may, on occasion, competently employ methods they are not generally competent with respect to, or know of a method that it is likely faulty, by relying on that very method.

4. Objections and Replies

In this section, I sketch some objections to my argument, and consider responses.

[90] Importantly, we cannot run an argument like this if one has a relevant higher-order defeater whenever one is less than certain that one is reading the tychometer correctly. If this is one's picture, then for any *p* fed into the tychometer, one will always be able to have a degree of confidence lesser than that recommended by the tychometer without ceasing to have a higher-order defeater. But I take it that under such circumstances, the category of higher-order defeat will do no work over and above that of higher-order evidence, and will be theoretically uninteresting.

4.1 Weird Deductions

In ORACLE Watson comes to believe *p* solely on the basis of deduction from his belief that p. One might be suspicious of the idea that one can believe a proposition *p* solely on the basis of deduction from p, and so question the viability of this paper's core example. In this section, I will speak to this worry in two ways. First, I will argue that, at least given natural ways of articulating what looks troubling about ORACLE-style cases, it is unclear that suspicion is warranted. Second, I will clarify the dialectical role of the example.

First, let us get clear about the structure of the putative problem. The idea that one cannot believe a proposition *p* solely on the basis of deduction from *p* typically flows from a conviction that if one is to come to believe *q* on the basis of deducing *q* from some believed content that p, one must retain one's initial belief that *p throughout* the deduction. Call this the *continuity* constraint on deduction.

The problem is as follows. In ORACLE, Watson starts with a belief, *b*, which is based solely on the oracle's testimony, and whose content is *SHERLOCK*. Watson performs a competent deduction and ends up with belief b', where b' is solely based on this deduction.

The content of b' is also *SHERLOCK*. Now assume that Watson never loses his initial belief b. It follows that Watson has a belief b' based solely on deduction, and a belief b based solely on the the oracle's testimony. On the assumption that if x is a is a belief of Watson's that *p* and y is a belief of Watson's that p, then x = y, it must be the case that b = b'. But b cannot both be solely based on testimony and solely based on deduction.[91]

This looks like a good argument. However, it depends on our interpreting the constraint that an agent retain his initial belief throughout the deduction in a particular way. Contrast the following two different ways of making more precise the idea that an agent must hang on to his initial belief throughout the deduction.

1. *If an agent A is to have deductively based belief that q on the basis of belief that p,* the time t at which A first believes *q* must be such that at t, A is (i) still in possession of his initial belief, and (ii) in possession of a new belief.

[91] Rejecting this assumption will not help the adversary of higher-order defeat. If we allow that we can individuate beliefs so finely that an agent can have two distinct beliefs with the same content, the structure of ORACLE can be re-cast in ways friendly to the integrationist. Consider: if Watson never loses his testimonially based knowledge that he is bad at deduction, the higher-order defeatist has an obvious and unproblematic story to tell about what goes on in ORACLE. Watson fails to have deductive knowledge that he is bad at deductions because his testimonially based knowledge that he is bad at deduction acts as a higher-order defeater for his deductively based belief in the same.

2. *If an agent A is to have deductively based belief that q on the basis of belief that p,* the time t at which A first believes *q* must be such that at t, A still believes that q.

The first cashing-out will clearly (see the above argument) make trouble for ORACLE style cases so long as an agent cannot simultaneously have two distinct beliefs that p. But what about the second cashing-out? It will not make similar trouble for ORACLE style cases. Where *p* = q, it will never be the case that A believes *q* without also believing p. Now, there may well be reasons to prefer (1) to (2). But if there are, they are not immediately obvious. As such, it does not *clearly* follow from the continuity constraint on deduction.

4.2.1 The Assimilationist Objection

The assimilationist draws attention to the similarities between the set-up in cases like ORACLE, and more familiar puzzles in the literature on the semantic paradoxes and other paradoxes of self-reference.[92]

[92] It is not clear that 'paradoxes of self-reference' picks out any theoretically interesting kind, given the different sorts of machinery and theoretical commitments that can generate (at least superficially) similar puzzles.

It is hard to think for very long about the sorts of puzzles exploited in this paper without suspecting sleight of hand. Cases like ORACLE are reminiscent of the semantic paradoxes and other paradoxes of self-reference (see, for example, Dummett, 1991; Prior, 1961; Russell, 1908). And once this has been noticed, it can be hard not to worry that this paper's arguments overreach somewhat, in coming to substantive epistemological conclusions from what are more properly thought of as puzzles in the philosophy of logic and language. The thought continues: these cases are not *really* puzzles in epistemology -- and do not require an epistemological solution -- but are more perspicuously thought of as one flavour of a very general philosophical puzzle. Once we are careful to treat them as continuous with these other puzzles, the call for adjustments in our epistemological theorising will appear quite premature.

In this section, I propose to take this worry seriously. I shall show that it is quite right to think of these puzzles as closely related to a cluster of paradoxes of self-reference. I shall trace the prospects for extending proposed solutions to these other paradoxes to the puzzles which have been the focus of this paper. Some do not naturally extend to these puzzles, or do not extend in ways that are friendly to the integrationist; those that do extend come with substantial costs. It is not untoward, then, to think that these puzzles require a distinctively

'epistemic' solution. Of course, I cannot show that there is *no* way of responding to paradoxes of self-reference which offers the integrationist better prospects – the literature on such paradoxes is vast. But I can show that the integrationist needs to do a good deal more than simply assume that these literatures will be friendly to their purposes.[93]

Suppose we encounter an interlocutor who presses the thought that puzzles like ORACLE are not 'really' epistemic puzzles. Let us get a little more precise about their charge. Suppose they want to assimilate the cases dealt with in this paper to the semantic paradoxes, where the semantic paradoxes are understood *narrowly,* either as paradoxes that are generated by treating 'is true' as a predicate, or as paradoxes that are tightly bound to 'distinctively semantic notions' (Studd, m.s.).[94]

First, articulating a satisfactory response to the semantic paradoxes is typically taken to require some strategy for identifying paradox-prone ('diseased') sentences (Bacon, 2015). Diseased sentences have been identified, among other things, with sentences which are introduced using an impredicative definition (Russell, 1903), which are ungrounded

[93]I am not going to consider responses that involve abandoning classical logic. I take it that it would be very surprising if liking higher-order defeat turned out to push us in the direction of non-classical logic.

[94] It is, as Studd notes, common to distinguish the semantic paradoxes (narrowly construed) from paradoxes of self-reference which are differently generated, such as the Russell-Zermelo paradox (m.s.).

(Kripke, 1975), and with sentences which fail to follow analytically from non-linguistic facts (McGree, 1990). If there is a culprit sentence in ORACLE-style cases, it is presumably 'I am bad at deduction'. But this sentence is not introduced using an impredicative definition, or a good candidate for being ungrounded. And if *any* sentences ever follow analytically from non-linguistic facts, 'I am bad at deduction' surely might.

Second, the classical logician's response to the semantic paradoxes seems very badly placed to generalise to the culprit sentence in ORACLE. If you want to hang on to classical logic in the face of the semantic paradoxes, you must endorse, at most, a *restricted* version of the T-schema (see Bacon, 2015). But it is deeply unappealing to think that the right response to ORACLE is to deny that 'I am bad at deduction' fails to obey the following schema:

T_i: 'I am θ' uttered by s is true iff s is θ.

But our interlocutor can refine his charge. The features of cases like ORACLE that catch his attention, he can say, have nothing to do with distinctively semantic notions; rather, it is their tendency towards self-reference which is noteworthy. And paradoxes of self-reference can get going without our treating 'is true' as a predicate, so long as we take ourselves to be

able to quantify over propositions. In an example inspired by Prior – call it ARTHUR – we are tempted to describe Arthur as believing only that nothing he believes is true. Using the resources of propositional quantification we can express this as the claim that Arthur believes only that *for all p, if Arthur believes that p then not-p* (see Prior, 1961). And these resources are well suited to express puzzles like ORACLE.

The structural similarities between ARTHUR and ORACLE should be fairly obvious. Arthur believes all his beliefs are false. Either this belief is true or false. Suppose it is true. Then all his beliefs are false, including the one we just supposed was true. So suppose it is false that all his beliefs are false. He must have at least one true belief. But he only has one belief. So his one belief must be true. Again: a contradiction.

It should be clear that if we want a unified solution to the problems posed by ORACLE and ARTHUR, said solution will not have anything to do with higher-order defeat. The sensible way to proceed, says our interlocutor, is to see what we can say about ARTHUR style cases and then do our best to apply that strategy to cases like ORACLE. What I will do now is look at what I take to be the most promising responses to cases like ARTHUR, and show that they do not naturally extend to cases like ORACLE.

4.2.2 Ramifiers

One strategy for dealing with ARTHUR style cases involves becoming a *ramifier*. Ramifiers, largely driven by considerations of paradox:

> [D]eny that there is a knowledge relation that can relate a person to propositions involving that very relation. According to the ramifier, there is a knowledge-zero relation that can relate people to "boring" propositions. And then there is a knowledge-one relation — and a belief-one relation — that can relate people to boring propositions and also to propositions involving the knowledge-zero relation — and the belief-zero relation — but not to propositions involving a higher relation. And so on (Fraser and Hawthorne, 2015).[95]

How might one extend the ramified approach to deal with cases like ORACLE? The natural way to extend the ramifier's claim, will be to deny that there is some belief forming method that has, as its output, propositions about that very method. So the ramifier will say

[95] For more on the ramified approach, see Bacon, Hawthorne & Uzquiano (2016) and Kaplan (1995).

that one cannot *deduce* propositions about deduction, or *remember* propositions about memory, or be told propositions about telling. Instead, one would be able to deduce-zero (remember-zero, be told-zero) 'boring' propositions, and be able to deduce-one (remember-one, be told-one) propositions about deduce-zero-ing.

What should we make of the ramifier's suggestion? Granted, the ramifier's approach is deeply counterintuitive even in its more limited incarnations. But this extension of it increases, significantly, the costs associated with the ramified perspective. Suppose we think it true that someone is good at deductions if they have performed only competent deductions for the last ten years. If Watson finds out that he has performed only competent deductions over the last ten years, and knows that someone is good at deductions if they have performed only competent deductions for the last ten years, it is overwhelmingly natural to think that Watson ought to be able to believe on the basis of deduction that he is good at deductions. But the ramifier cannot allow this (although they can allow that Watson can believe on the basis of deduction-ones that he is good at deduction-zeros.) So even if the ramifier provides a way to construe cases like ORACLE as non-paradoxical, becoming a ramifier will involve a substantial reconfiguration of our epistemic commitments. We should

be deeply suspicious of the idea that cases like ORACLE can be approached in any way that leaves epistemology untouched.

But there is another, more serious problem for the ramifier. Suppose we apply the ramifier's strategy to ORACLE, and say that Watson deduce-ones that they are bad at deduction-zeros. How does this help? Presumably some version of SPC* will be attractive for each of the ramifier's variants on deduction, including deduction-one. And if this is so, then if Watson fails to know that they are bad at deduction-zeros it must be because (given that they have competently* deduced-one it) they have a higher-order defeater for the claim. But either they lack a higher-order defeater (in which case the variant on SPC* is false) or they have one. And if they have one, the only plausible thing that could be doing the work of defeat is the claim that they are bad at deduction-zeros. But once we grant that beliefs about deduction-zeros can act as higher-order defeaters for beliefs that have been deduced-one, the ramifier's approach ceases to offer a way out of the paradox: a familiar pattern of contradiction will arise.

The alternative is that Watson ends up in the right sort of epistemic position with respect to the claim that they are bad at deduction-zero for the claim to act as a defeater. The ramifier

will point out that on their story, S did not *deduce-zero* that they are bad at deduction-zero. And this much is certainly true, but it is not clear that it is of much help to the fan of higher-order defeat. There is a dilemma for the ramifier. The first horn says that beliefs about deduction-zero can act as higher-order defeaters for beliefs formed on the basis of deduction-one. But if he goes for this horn, the ramifier will lose their claim to an integrationist-friendly story: ramification will come with the same troubling dynamics that have dogged the integrationist throughout this paper.

The second horn says that beliefs about deduction-zero cannot act as defeaters for beliefs about deduction-one. Taking this horn means abandoning much of the spirit of integrationism. After all, Watson is not himself a ramifier: he thinks, as most of us do, that there is only one deduction-like relation that relates agents to propositions. So when Watson deduce-ones propositions about deduction-zero, things look to Watson 'from the inside' just as they would look to Watson if he had really deduced that he is bad at making deductions. To have things look this way is to be subject to a species of mental conflict: apparent dissonance between one's higher and lower epistemic attitudes. An integrationist who licenses this kind of mental conflict endorses a very ambivalence-friendly form of integrationism.

This is perhaps at its most stark in testimonial cases. Suppose it turns out that things are as the ramifier says, with respect to telling: 'boring' propositions get told-zero, propositions about telling-one get told-zero, and so on. The ramifier who goes for this second horn of the dilemma in effect endorses an account that has it that when I say to you 'You are highly likely to misinterpret anything I tell you', you can come to know, on my say-so, that you are highly likely to have misinterpreted my testimony. (For really, they say, regardless of how it seems to you, you have been told-one that you ought not to trust what I tell-zero you). But to think that you can come to know what I tell you, under these circumstances, is a significant departure from natural integrationist readings of the case, insofar as it licenses the mental dissonance involved in your believing on my say-so that you ought not to trust what I tell you. And once we allow the kind of mental dissonance at issue here, it gets harder to see what is wrong with the other cases of mental dissonance condemned by the integrationist. That is, grasping the second horn of the dilemma does avoid paradox but appears to come at the expense of much of the spirit associated with integrationism.

Maybe there are ways of refining the ramifier's account, or getting careful about the integrationist's commitments, that make the tensions sketched here look less troubling.

Regardless, it should be clear that even if paradoxes like ORACLE can be treated along the lines suggested by the ramifier, the ramifier's solution comes with substantial implications for epistemology. The lover of higher-order defeat cannot make any straightforward or blasé appeal to paradoxes elsewhere in the hope of making the problems posed by ORACLE look epistemologically insubstantial. To treat ORACLE by becoming a ramifier would require a significant revision of the integrationist's epistemic commitments.

4.2.3 Adding and Subtracting.

The two most promising strategies for dealing with cases like ARTHUR are the ADDER strategy and the SUBTRACTOR strategy (for discussion, see Fraser & Hawthorne, 2015). Let us take each strategy in turn, show how they handle cases like ARTHUR, and sketch the prospects for applying them to cases like ORACLE as well.

Dorr (m.s.) is the primary advocate of the ADDER strategy. The core idea of the ADDER strategy is *profusion*. In cases like ARTHUR, there is no paradox if Arthur *in fact* has at least one true belief. So long as we are only ever *tempted* to describe Arthur as believing only the proposition that everything he believes is false, or only ever *tempted* to describe

Arthur as believing only this proposition along with other false ones, we will not have a paradox. So, the ADDER suggests, we ought to think that Arthur, despite our being inclined to describe him as believing only that everything he believes is false, believes something *true* too (Fraser and Hawthorne, 2015). If Arthur has at least one true belief, we can happily suppose his belief that everything he believes is false is itself false. As Fraser and Hawthorne note, this strategy generalises to various other puzzles: if one writes on board X the words 'Every proposition expressed on board X is false' and writes down no other words, then one expresses, according to the ADDER, multiple propositions, at least one of which is true. All by itself, of course, logic cannot tell us what this extra true belief (or what these extra true beliefs) of Arthur's is (or are). But the strategy is at its most plausible if Arthur's 'extra' true belief is a belief whose content is introspectively indistinguishable from the content of Arthur's belief that everything he believes is false. This means the ADDER has to hand a natural account of why it is so tempting to describe Arthur as believing just one thing.

The SUBTRACTOR strategy is slightly different and has at its core a robust anti-disquotationalist tendency (see Bacon, m.s.). Instead of attributing to Arthur belief in the proposition that anything Arthur believes is false (as the ADDER does), the

SUBTRACTOR attributes to Arthur belief in a slightly different proposition, for example, belief in the proposition that anything Arthur *believes'* is false, where belief' is an attitude distinct from but *very similar* to belief. We can happily suppose, then, that Arthur's belief that all his beliefs' are false is vacuously true (perhaps he has no beliefs'), non-vacuously true (perhaps all his beliefs' are false, but some of his beliefs are true) or false (perhaps he has some true beliefs'). At any rate, we escape Prior's paradox.

What would the expansion of these strategies to the cases in this paper look like? Let us first look at the SUBTRACTOR strategy, and think through in some detail what mechanics a SUBTRACTOR envisages for cases like those we have been considering:

> HYPOXIA*: Watson begins at t_1 with a rational degree of confidence of *n* in the proposition that he has hypoxia ('*POXY*'). That is, let it be the case that $Cr_1(POXY) = n$. Watson competently deduces *POXY* from itself and comes, at t_2, to endorse *POXY* solely on the basis of this deduction.

Let it be the case that at t, *POXY* can act as a higher-order defeater only if $Cr_t(POXY) \geq n$. HYPOXIA* poses a problem for the higher-order defeatist because Watson either ends up

with $Cr_2(POXY) \geq n$, or $Cr_2(POXY) < n$. If the former, then Watson has a higher-order defeater for *POXY*, and so ought to have a degree of confidence of less than *n*. If the latter, then Watson lacks a higher-order defeater, and ought to have a degree of confidence of *n* or greater. Either way, we have a counterexample to RDB.

How does the SUBTRACTOR address this sort of case? The SUBTRACTOR will *deny* that Watson ends up with an attitude towards *POXY*. Rather, the SUBTRACTOR says, Watson will end up with an attitude towards some very similar content, *POXY'*, which is entailed by, but weaker than *POXY*. As such, *POXY'*, like *POXY*, may be competently deduced from *POXY*.

Let us examine the prospects for such a strategy. Suppose Watson's $Cr_2(POXY') < n$.This will not help much. For if $Cr_2(POXY') < n$, and *POXY'* has been competently deduced from *POXY*, and had a degree of belief assigned on this basis, RDB predicts that at t_2, $Cr_2(POXY)$ is less than *n* only if Watson has a relevant higher-order defeater. But considering that *POXY* functions as a relevant higher-order defeater at t only if $Cr_t(POXY) \geq n$, and that *POXY'* is a weaker claim that *POXY* – otherwise it could not be competently deduced from *POXY* – it is overwhelmingly natural to suppose than *POXY'*

can function as a relevant higher-order defeater at t_2 only if $Cr2(\textit{POXY}') > n$. If the SUBTRACTOR strategy is to avoid a familiar impasse, they must suppose that $Cr_2(\textit{POXY}') \underline{\geq} n$. But this may be a working strategy: RDB allows for $Cr_2(\textit{POXY}') \underline{\geq} n$, and there is no decisive reason to think that whenever $Cr_2(\textit{POXY}') \underline{\geq} n$, that Watson has a higher-order defeater relevant to *POXY'*. Indeed, there is pressure in the opposite direction.

Consider: it does not follow from the claim that (i) if, at t_2, Watson has a degree of confidence of at least *n* in *POXY*, he has a higher order defeater relevant to deduced content, that (ii) if, at t_2, Watson has a degree of confidence of at least *n* in a very similar but slightly weaker claim, he has a higher-order defeater relevant to deduced content. Given that *POXY'* is a slightly weaker claim than *POXY*, it is natural to think that the *m* such that at t, *POXY'* acts as a relevant higher-order defeater for deduced claims only if $Cr_t(\textit{POXY}) \underline{\geq} m$ is slightly larger than *n*. This is compatible with thinking that in general, where two contents are very similar and introspectively indistinguishable, they can act as defeaters in much the same way as each other. Suppose that there are two different colours, red and red*. The difference between these colours is so slight that ordinary people cannot tell the difference between red things and red* things, and shifts in meaning between 'red' and 'red*'

are introspectively undetectable. If one is looking at a bowl and it seems red to you, whether or not you have a defeater for this claim will not (generally) depend on whether you are told that the bowl is being irradiated with red light or that it is being irradiated with red* light. But note the 'in general' -- there *must* be cases in which these tiny shifts, despite being introspectively undetectable, do make a difference to whether these beliefs act as defeaters. We do not want to commit ourselves to the claim that for all *p* and *p'* where *p* and *p'* are very similar and are introspectively indistinguishable, and S is in the same epistemic position with respect to each, that if one defeats S's belief that *q,* so does the other. To so commit ourselves would be to expose ourselves to a sorieties series.

The SUBTRACTOR, then, has a way of explaining why, in HYPOXIA*, we are inclined to attribute a deductively based degree of confidence in *POXY* to Watson: he in fact has a deductively based degree of confidence in a very similar, perhaps introspectively indistinguishable contents. They also have a strategy for avoiding paradox: so long as $Cr_2(POXY')$ is *n,* or very slightly more than *n,* the reinterpreted case presents no problem for the integrationist.

The ADDER strategy is not quite so promising. How will the ADDER address HYPOXIA*? Typically, they will say that at t_2, Watson ends up with deductively based degrees of confidence of around *n* directed not only towards *POXY*, but also towards *POXY'*. The ADDER should not suggest that $Cr_2(POXY)$ and $Cr_2(POXY')$ are both less than *n*. If this is so, there will be no attitude that looks to be a good candidate for playing the role of defeater, and we will not avoid counterexamples to RDB. Similarly, the ADDER should not suggest that Watson's $Cr_2(POXY) \geq n$ – this will keep the core of the puzzle intact, regardless of what Watson's other attitudes might be. If the ADDER is to have a good strategy, it must be $Cr_2(POXY) < n$, and $Cr_2(POXY') \geq n$.

Let us be a little more specific: suppose the ADDER suggests that $Cr_2(POXY') = m$, where $m > n$. Either *m* is such that at t_2, *POXY* acts as a higher-order defeater for deduced claims where $Cr_2(POXY') \geq m$, or it is not. If it is *not* the case that *POXY'* acts as a relevant higher-order defeater at t_2, then we have no explanation for why $Cr_2(POXY) < n$. But if *POXY' does* act on *POXY* as a higher-order defeater at t_2, we have an explanation of why $Cr_2(POXY) < n$ (Watson has a relevant higher-order defeater) and no counterexample to RDB. For RDB allows that $Cr_2(POXY')$ may be greater than *n*. The trouble with this strategy is this. If we admit that where $Cr_2(POXY') = m$, *POXY'* acts as a higher-order

defeater for deduced claims, then it looks like *POXY'* will act as a defeater for *POXY'* as well as for *POXY*. And if this is so, the ADDER's picture looks susceptible to the same sorts of puzzles it is supposed to help us escape. (Note that this is not an argument that the ADDER strategy cannot be made good – just that there are *prima facie* problems with the strategy which the SUBTRACTOR avoids.

The SUBTRACTOR strategy, then, is the best bet for the defender of higher-order defeat who wants to resist the thought that there is a distinctively epistemic phenomenon here. But this account comes with costs: it means we cannot truthfully describe agents as having the attitudes they are most naturally described as having. It is far from obvious that this is a price worth paying in order to hang onto the ideology of higher-order defeat.

Conclusion and Outlook

I do not want to overstate the implications of this paper for the higher-order defeatist. The arguments of a single paper cannot hope to be decisive: there may be ways of amending the story told about higher-order defeat so that it is insusceptible to the kinds of problems I have

charted here. But the ball is now in the court of the defender. It is not my job to survey every possible strategy my adversary might use to patch up their favoured ideology.

Let us take a little more time to elaborate on the dialectical situation of the friend of higher-order defeat. Suppose someone argues that we should reject the ideology of *truth*, because endorsing an unrestricted T-schema leads to paradox. There are highly premature: surely we ought to *amend* the T-schema – restrict it in some way – rather than toss out the ideology of truth. But the project of actually specifying *how* the T-schema ought to be restricted is a good deal more difficult than noting that some restriction is necessary.[96]

Similarly, it is pretty *easy* to say the following: if we want to save the ideology of higher-order defeat, we need to tweak, or restrict in some way, the claims via which it is characterised. But the project of identifying the shape of such a restriction may prove to be a *good deal more difficult* than noting that some such restriction is necessary. The assumption ought not to be that it will be easy or straightforward to make the tweaks the higher-order defeatist wants. (And our starting point with respect to truth and knowledge and our starting point with respect to higher-order defeat are very different: truth is something we take

[96] See, for instance, Halbach (2006)

ourselves to have a pre-theoretic grip on; it is part of our basic picture of the world. Higher-order defeat is not part of a pre-theoretic package: it comes to us only once we start doing epistemology.)

This paper started off drawing a rough distinction between integrationist and ambivalent epistemologies, and identified higher-order defeat as an integrationist piece of ideology. The implications of my argument for the dispute between ambivalence and integrationism are far from straightforward. Even if we should do away with higher-order defeat, the integrationist need not tie their project to higher-order defeatism: maybe integrationist intuitions can be accommodated by some other, less brittle theoretical apparatus (anti-akratic constraints on rationality, for example.) But perhaps not.

Nevertheless, insofar as higher-order defeat codifies integrationist impulses in epistemology, its failings should nudge us in the direction of epistemic ambivalence. So far, the paper has adopted a critical register. I will finish in a more constructive mode, by saying something in favour of epistemic ambivalence.

Philosophy sometimes consists of arguments (perhaps less often than philosophers would like to think). But sometimes philosophy takes the form of more-or-less affectively laden invitations or suggestions to look at matters in a new or different light.

Epistemic ambivalence is only one sort of ambivalence; and not the only sort that has been looked at askance: a long tradition valorises the integrated mind (see Brunning, m.s.) Ambivalence, though, has found a recent apologist in Wallace, who directs our attention to the embeddedness and contingency of our various commitments. In one example, he considers the case of a young girl who falls pregnant and decides to carry the pregnancy to term and keep the child. His relationship with his child in later years becomes central to his life, and is a source of great joy and meaning to her. She is committed to regarding this child, and his relationship with his child, as a good. But on the other hand, when the mother looks back on his decision, she should be committed to judging that his decision was in some sense a foolish one. She is committed, then, and properly, to the upshot of his decision, whilst recognising that she should, perhaps, have chosen differently. In another case, a talented artist makes the decision to abandon his wife and child in order to pursue a career as a painter. He is able to achieve the great and valuable things that he does, as an artist, only because of this decision he made. But although 'his decision [to abandon his

family] was a necessary historical condition for the values that shape his later point of view and ground his affirmative attitude toward the life he has actually led', it seems that he ought not to affirm – at least not to the same, wholehearted degree – the decision to abandon his family (2013). Wallace sees the objects, relationships and practices we value and affirm as firmly embedded in and dependent upon wholes which we cannot affirm. A third case – 'the bourgeois predicament' – has a more global aspect. Addressing, one supposes, his fellow professional philosophers, Wallace notes:

> [T]he significant sources of affirmation for most of us are distinctively bourgeois attachments. They are projects and relationships that could exist only under conditions of comparative material advantage, and that presuppose both the opportunities and the freedom from material necessity that such advantage makes possible. The next observation I want to make is that the material circumstances that in this way condition our bourgeois attachments are constitutive parts of larger social and historical patterns that are impersonally lamentable (Wallace, 2013).

Ambivalence, in such cases, seems to be the right response – or so Wallace thinks. For there seems to be a sense in which I cannot – and indeed ought not – simply to give up on all of

my projects, friendships, work. But I cannot similarly affirm the institutions and processes that get me in touch with these goods – I cannot suppose that my attachment to certain projects 'swamps' the character of their genealogies.

One might, of course, disagree with Wallace's assessment of what such cases require of us – after all, Williams (1981) disagreed with Wallace's diagnosis of the first two cases, and Nagel, to some extent, takes issue with Wallace's account of 'the bourgeois predicament'. Both take Wallace to be unduly harsh (we can affirm more than Wallace supposes, they think). A different way of disagreeing would be to suppose that Wallace is too generous when he supposes that the painter is able to affirm his art, and the philosopher his bourgeois pursuits. But if affirmations can be appropriate or inappropriate, and we might appropriately affirm some relationship, practice, or object without appropriately affirming the conditions that allow us to get in contact with that relationship, practice, object, then we have to hand, I think, a promising way of looking at epistemic ambivalence. If we can think of knowledge as a species of affirmation – affirmation of a content which is in some sense apt – we should expect, if Wallace's diagnosis of our situation is right, that we will sometimes be in a position to aptly affirm contents *without* being in a position to aptly affirm the conditions

that get us in touch with these contents – that is, without being in a position to aptly affirm our belief forming mechanisms.

Acknowledgements.

Thanks to Lizzie Fricker, John Hawthorne, Alex Kaiserman, Nakul Krishna, Beau Madison Mount, Alex Roberts, Amia Srinivasan, Tim Williamson, and Andy Yu for discussion and comments on earlier drafts. Earlier drafts of this paper were presented in Oxford and Vienna; thanks go to my audience on both occasions.

References.

Bacon, A., 2015. Can the classical logician avoid the revenge paradoxes? *Philosophical Review* 124 (3): 299–352.

- m.s.. Radical Antidisquotationalism.

Bacon, A., Hawthorne, J. & Uzquiano, G., 2016. Higher-Order Free Logic and the Prior-Kaplan Paradox. In eds. Yli-Vakkuri, J. & McCullagh, M., *Williamson on Modality: Supplementary volume of the Canadian Journal of Philosophy.* Taylor & Francis.

Baker-Hytch, M. & Benton, M., 2015. Defeatism Defeated . *Philosophical Perspectives* 29 (1): 40-66

Bergmann, M., 2006. *Justification without awareness: A defense of epistemic externalism.* New York: Oxford University Press.

Bernecker, S., 2008. *The Metaphysics of Memory*. Dordrecht: Springer.

- 2011. Further thoughts on memory: replies to Schechtman, Adams, and Goldberg. *Philosophical Studies,* 153 (1): 109-121.

Blaauw M., 2008. Subject Sensitive Invariantism: In Memoriam. *Philosophical Quarterly* 58 (231): 318–325.

Burge, T., 1997. Interlocution, Perception, and Memory. *Philosophical Studies* 86 (1): 21–47.

Christensen, D., 2010a. Higher-Order Evidence. *Philosophy and Phenomenological Research,* 81 (1):185–215.

- 2010b. Rational Reflection. *Philosophical Perspectives* 24 (1):121–140.

Cohen, S., 1987. Knowledge, Context, and Social Standards. *Synthese,* 73 (1): 3–26.

De ridder, J., 2014. Why only Externalists can be Steadfast. *Erkennesis,* 79: 185–199.

Dorr, C., m.s.. Propositional Profusion and The Liar.

Dummett, M., 1991. Frege: Philosophy of Mathematics. London: Duckworth.

Elga, A., m.s. Lucky To Be Rational. https://www.princeton.edu/~adame/papers/bellingham-lucky.pdf.

Fraser, R. E. & Hawthorne, J., 2015. Cretan Deductions *Philosophical Perspectives* 29 (1): 163–178.

Fumerton, R., 1995. *Metaepistemology and Skepticism.* Maryland: Roman and Littlefield.

Goldman, A., 1986. *Epistemology and Cognition.* Cambridge, MA: Harvard University Press.

Greco, J., 2010. *Achieving knowledge: A virtue-theoretic account of epistemic normativity.* Cambridge: Cambridge University Press.

Halbach, V., 2006. How Not To State The T-Sentences. *Analysis* 66 (4): 276–280.

Horowitz, S., 2014. Epistemic Akrasia. *Noûs* 48: 718–744.

Kaplan, D., 1995. A problem in possible worlds semantics. In eds. Sinnott-Armstrong, W., Ra, D. & Ansher, N., Modality, Morality and Belief: Essays in Honor of Ruth Barcan Marcus. Cambridge: Cambridge University Press, pp. 41-52.

Kaplan, M., 2009. Williamson's Casual Approach to Probabilism. In eds. Greenough, P. & Pritchard, D., *Williamson on Knowledge*. Oxford: Oxford University Press, pp. 123-139.

Kripke, S., (1975). Outline of a theory of truth. *The Journal of Philosophy* 72 (19): 690-716.

Lackey, J., 2005. Memory as a Generative Epistemic Source. *Philosophy and Phenomenological Research* 70 (3): 636–658.

Lasonen-Aarnio, M., 2010. Unreasonable Knowledge. *Philosophical Perspectives* 24 (1): 1-21.

- 2014. Higher-order Evidence and The Limits of Defeat. *Philosophy and Phenomenological Research* 88 (2): 314-345.

McGee, V. M.,1990. *Truth, vagueness, and paradox: An essay on the logic of truth.* Indianapolis: Hackett Publishing Company Inc.

Nagel, T., 2014. An Invitation to Handwringing. *The London Review of Books* 36 (7).

Owens, D., 2002. Epistemic Akraisa. *The Monist* 85 (3): 381-97.

Plantinga, A., 1993. *Warrant and Proper Function.* Oxford: Oxford University Press.

Russell, B., 1903. *The principles of mathematics.* New York: WW Norton & Company.

- 1908. Mathematical logic as based on the theory of types. *American Journal of Mathematics* 30 (3): 222-262.

Schechter, J., 2013. Rational Self Doubt and the failure of closure. *Philosophical Studies* 163 (2): 428-452.

Studd, J., m.s.. *Absolute Generality.*

Wallace, R. J., 2013. *The View From Here: Affirmation, Attachment, and the Limits of Regret.* Oxford: Oxford University Press.

Wedgewood, R., 2011. 'Justified Inference.' *Synthese* 189 (2): 1–23.

Williams, B., 1981. *Moral Luck.* Cambridge: Cambridge University Press.

Williamson, T., 2000. *Knowledge and Its Limits.* Oxford: Oxford University Press.

- 2009. Replies to Critics, In eds. Greenough, P. & Pritchard, D., *Williamson on Knowledge.* Oxford: Oxford University Press, pp. 280–384.
- 2011. Improbable Knowing. In ed. Dougherty, T., *Evidentialism and its Discontents.* Oxford: Oxford University Press, pp. 146–164.

Supplement to Epistemic Ambivalence.

I began *Epistemic Ambivalence* by saying that it would be a mistake to run together the different ideologies favoured in different integrationist epistemologies. And in *Epistemic Ambivalence* I looked at one artefact of the integrationist picture: higher-order defeat. What I mean to do in this supplement is to sketch out some of the prospects for extending the arguments mobilised against higher-order defeat to other vestiges of the integrationist tendency in epistemology. In particular, I turn my attention to two different *anti-akratic requirements* (requirements which are the central focus of the final paper), and show how we can construct a case against each of these requirements that is structurally similar to the case made against higher-order defeat. These anti-akratic requirements are as follows:

> AAR: If A rationally believes that it is now rationally permissible for him to believe that *p,* it is rationally permissible for A to believe that *p*.

> AAR_e: It is impossible for A to rationally have high confidence both that *p* and that *p* is unlikely on her evidence.

The argument against higher-order defeat leant heavily on the plausibility of the thought that knowledge is closed under simple single premise deductions. If we are to construct structurally similar cases against anti-akratic constraints, we will need to find it similarly plausible that rational belief and rational high confidence are closed under simple single premise deductions. There are well-rehearsed objections to the claim that either rational belief or rational high confidence is closed under *multi-premise* deductions. Nevertheless, we might think that these worries do not extend to the single premise case and be tempted by some sort of single premise closure principle for rational belief. So long as one also likes an anti-akratic requirement on rationality, one will need to structure this principle so that it does not predict rational belief (that *q*, say) on the basis of deduction at the same time as rational belief that it is irrational to believe *q*.

Suppose one is inclined to endorse a principle along the following lines:

> SRC: If A rationally believes *p* and competently* deduces *q* from *p* and comes to believe *q* solely on that basis then A's belief that *q* is rational.

Suppose too that S starts off rationally believing the conjunction that *p and S cannot competently perform any deduction that exploits conjunction elimination.* (After all, one can be rationally highly confident that *p* where *p* is false.) Suppose S competently deduces that they cannot competently perform any deduction that exploits conjunction elimination ('*CON*'), and comes to believe said proposition solely on the basis of this competent deduction. We will have trouble whether or not we suppose that S ends up with rational belief that *CON*.

Suppose S rationally believes that *CON*. Then S rationally believes that all their deductions which exploit conjunction elimination are incompetent. But presumably they also rationally believe that they believe *CON* on the basis of an incompetent deduction, and as such, rationally believe that they ought not to believe *CON*. But then if they rationally believe they ought not to believe *CON,* they do not, by AAR, rationally believe that *CON,* and we have a contradiction.

So let us suppose, on the other hand, that S believes *CON* but is irrational to do so. SRC tells us that S rationally believes *CON* unless it is the case that S rationally believes that he is rationally forbidden to believe *CON*. So it must be the case that S believes he is rationally

forbidden to believe *CON* and is rational to do so. S knows he believes *CON* on the basis of a deduction exploiting conjunction elimination, so we have a natural explanation of how S has the relevant belief if he also has a rational belief that he is forbidden to believe things on the basis of conjunction elimination. But this is tantamount to admitting that he has a rational belief that *CON*. Again, a contradiction.

Similar problems emerge for AAR_e when it is conjoined with a plausible closure principle:

> $SRC_{e:}$ If S is rationally highly confident that *p* and competently* deduces *q* from *p* and comes to be highly confident that *q* solely on that basis, then S's high confidence in *q* is rational, so long as S ought not be highly confident that their deduction was incompetent.

It is natural to think that if *p* is highly probable on one's evidence and q is entailed by *p,* then *q* is highly probable on one's evidence. But suppose instead one ends up rationally believing that good evidence is *not* closed under entailment. Under such circumstances, one might, it seems, rationally believe the conjunction that *p and that one has little evidence for any proposition one deduces.* Suppose that from this conjunction you competently deduce that

any proposition you deduce is unlikely on your evidence. Call this proposition *D**. According to SRC_e, your high confidence in D* is rational.

Suppose, then, that as well as being rationally highly confident that *D**, you are rationally certain that you believe *D* on the basis of deduction.* Thus, you should be rationally highly confident that *D* is unlikely on your evidence.* But then we have a case in which you are highly confident both that *D** and that *D* is unlikely on your evidence*. And this is contrary to AAR_e.

It is worth noting, too, that we do not need anything nearly as strong as SRC_e to get a problem like this going. All we need is a much weaker claim:

> Close-SRC_e: If A is rationally confident to degree *n* that *p* and competently deduces *q* from *p* and comes to be confident to some degree m that q where m is either equal to or slightly less than *n* then A's high confidence in q is rational, *just so long as* A ought not be highly confident that they ought not believe that *q.*

This claim should be particularly attractive insofar as it is immune to critiques of closure principles that exploit the mechanics of 'long chains' of competent deductions of the sort levelled by Schechter (2013) at closure principles (see *Risk, Doubt and Transmission* for discussion). So long as you start off with really high confidence in the conjunction, we can suppose both that for every deduction you perform, you are a little less rationally confident in the conclusion than in the premises, and still think you end up with high confidence that *D** and high confidence that *D* is improbable on your evidence.* It may be, if not easy, tolerable to live with the thought that rational outright belief is not closed under competent deductions of the right sort, but it will be a good deal harder to live with the thought that one can competently deduce *q* from *p,* be rationally highly confident that *p,* and have one's rational degree of confidence that *q* be much smaller than one's rational degree of confidence that *p*.

It should, then, be clear that there are at least the relevant raw materials for an extension of my case against the ideology higher-order defeat to other artefacts of integrationist epistemology, and in particular, to different formulations of the anti-akratic requirement on rationality. But insofar as closure principles for rational belief are less precious to us than closure principles for knowledge, the case against AAR will be correspondingly less forceful.

It is worth wondering, too, whether these other artefacts of integrationist epistemology might be more readily combined with the ADDER or the SUBTRACTOR story than higher-order defeatism was. These strategies failed to combine with higher-order defeatism because small tweaks in a belief's content often make little difference to whether or not it will act as a defeater. Defeat conditions are too coarsely individuated for the tweaks in content posited by the ADDER and the SUBTRACTOR to make a sufficiently systematic difference to the troubling dynamics. Insofar as it seems that any two propositions p and p^* which are very similar will play very similar evidential roles, AAR_e does not look well placed to be rescued by the SUBTRACTOR. Things look a bit better for the proponent of AAR, for coarsened versions of AAR do not look entirely plausible:

> COARSE-AAR: In general, if A rationally believes that it is rationally permissible for him to believe that p, it is rationally permissible for A to believe that p^*, where p and p^* are very similar propositions.[97]

[97] Without the 'in general' COARSE-AAR would be flatly implausible. Suppose S rationally believes that she is rationally permitted to believe she is about 2 metres tall. She also rationally believes COARSE-AAR. From COARSE-AAR and the claim that S rationally believes that S is rationally permitted to believe she is about 2 metres tall it follows that S is rationally permitted to believe a proposition very similar to the claim that S is 2 metres tall, say, the claim that she is about 2.01 metres tall. And S can work this out. So she can rationally

But one might defend COARSE-AAR in the following way. So long as the differences between p and p^* are so tiny that they are not introspectively detectable, one will typically rationally believe of p* that it is the same proposition as p. And if one rationally believes of a belief that it has a content one is rationally permitted to believe, it looks like we should think the belief rationally permitted, even if the contents have been misidentified.

believe that she is rationally permitted to believe that she is 2.01 metres tall. And so she can work out from this along with COARSE-AAR hat she is rationally permitted to believe that she is 2.02 metres tall. And so on. But S ought not to be able to infer that she is permitted to believe she is 200 metres tall however many inferences of this kind she makes.

References.

Schechter, J., 2013. Rational Self-Doubt and the Failure of Closure. *Philosophical Studies* 163 (2):428–452.

Abstract – Tragedy and Rationality.

I argue that there are tensions between two appealing structural constraints on rationality: anti-akratic constraints, and what I call anti-tragic constraints. Anti-akratic constraints on rationality, popular with contemporary epistemologists, see dissonance between first and higher-order epistemic attitudes as a species of rational failure (see Greco (2014), Horowitz (2014), Littlejohn (2015), Scanlon (1998), Titelbaum (2015a)). Tragedies happen when we cannot avoid acting as we ought not; anti-tragic constraints codify the thought that so long as an agent is in a permissible state, there will be some permissible course of action available to them. . First, I use some cases from Arntzenius et al. (2004) to show how tasks involving series of infinitely many decisions interact with anti-tragic commitments for prudential rationality. Second, I show how these standard cases can be 'epistemicised', so as to generate interesting epistemological conclusions, as well as suggestive argumentative structures. Third, I show that cases with this structure reveal tensions between anti-tragic and anti-akratic constraints on rationality. Fourth, I show that these tensions cannot be avoided by endorsing what I call the 'moderate' anti-akratic view defended by Horowitz (2014). I conclude by sketching some different responses to the tensions I demonstrate.

Tragedy and Rationality.

Introduction.

We all know that the world is an unfair place. Those who act as they should not often end up better off than those who have done nothing wrong. As Thrasymachus insisted to Socrates, injustice is often more profitable than justice (1981). But fewer of us are convinced that the world has a *tragic* structure. Tragedies happen when we cannot avoid acting as we must not: Agamemnon, a 'previously guiltless man', finds himself in a tragic situation when he must either sacrifice Iphigenea, or leave his army becalmed, and 'both courses involve him in guilt' (Nussbaum, 1986). The tragedian insists that, whatever Agamemnon does, he will have done something he shouldn't have. The anti-tragic retort is that, however things appear, there must be some way for Agamemnon to avoid acting as he should not act. Here, then, are the core anti-tragic commitments:

> ANTI-TRAGIC: If you are in a permissible state at t_1, then at t_1, you can avoid acting impermissibly between t_1 and t_2.

FORBIDDEN: If you are in a permissible state at t_1 and an impermissible state at t_2, you have done something impermissible between t_1 and t_2.

Of course, imperatives come in different flavours: there are oughts of rationality, as well as oughts of ethics, and rationality itself is polylithic: it is standard to distinguish between *epistemic* and *prudential* rationality.[98] I take it that we have a workably robust intuitive grasp of the distinction between epistemic imperatives and their non-epistemic analogues: when I am promised one million pounds if I believe that 2 is an odd number, there is a sense in which I ought to believe 2 is an odd number, and a sense in which I ought not. The first sense corresponds to prudential rationality, the latter to epistemic rationality.[99] Similarly, the language of permissibility and impermissibility comes in both prudential and epistemic varieties. Given my evidence, it is *epistemically* impermissible for me to believe the world is

[98] For example, Kelly distinguishes between epistemic and instrumental rationality thus: epistemic rationality is what one displays when one believes propositons that are well supported on one's evidence and refrains from believing propositions which are improabable on one's evidence; instrumental rationality is what one displays in taking the means to one's ends (2003).

[99] I take it that 'A ought to do such-and-such' is synonymous with 'A is required to do such-and-such'. To my ear, it sounds rather more natural to talk about what one ought to do, rather than what one is required to do, and so where sense is not at issue, I favour formulations that talk of oughts, rather than of requirements.

less than 6000 years old. But it may well be prudentially permissible (or even required) for me to believe this, if I will go to heaven only if I do, and to hell otherwise.[100]

We negotiate a diversity of normative structures; anti-tragic commitments come in correspondingly diverse forms. We can distinguish between ethical anti-tragedians, epistemic anti-tragedians, and so on. It is worth explicitly stating the 'epistemicised' versions of the core anti-tragic commitments:

> ANTI-TRAGIC$_e$: If you are in an epistemically permissible state at t_1, you can, at t_1, avoid acting epistemically impermissibly between t_1 and t_2.[101]

[100] There are delicate issues in the background here between belief and the will – you might think it can never be practically rational to believe that *p* because one never really has a choice of whether to believe *p* or not, in the way that one can have a choice of, say, whether or not to toss a coin. I put these complications aside; they are not important for the arguments of this paper.

[101] The anti-tragedian is not committed to the thought that an agent who has already acted impermissibly by t is able, at t, to avoid acting impermissibly in the future. An illustration: suppose you think agents are not rationally permitted to assign credences which violate the probability axioms. You violate the axioms by assigning a credence of 1 to both *p* and its negation. But having done this, you cannot avoid further violations: either you can assign a credence of less than 1 to their conjunction, or you can assign a credence of 1 to their conjunction. Either way you will be doing something impermissible: either you will be assigning credence 1 to a logical untruth, or you will be violating the probability calculus, which says that the right credence to have in p&q, where $Cr(p) = 1$ and $Cr(q) = 1$ is always 1.

FORBIDDEN$_e$: If you start off in a permissible belief state at $t_{1,}$ and end up in an impermissible belief state at $t_{2,}$ then you have done something epistemically impermissible between t_1 and t_2.

The epistemic anti-tragedian endorses one set of claims as to the structure of rationality. The project of this paper is to show, and to make vivid, some tensions between the anti-tragic impulse in epistemology and the claim that rationality has an anti-akratic structure. Anti-akratic constraints on rationality, popular with contemporary epistemologists, see dissonance between first and higher-order epistemic attitudes as a species of rational failure.[102] Here is one fairly standard statement of the constraint:

[102] Something like an anti-akratic constraint is defended or accepted by Greco (2014), Horowitz (2014), Littlejohn (2015), Scanlon (1998) and Titelbaum (2015a). Some anti-akratic constraints are formulated as wide scope rational requirements, others are formulated as narrow scope rational requirements. These differences can be significant (see Broome (2007), Titelbaum (2015b)) but are not very important for the arguments that follow.

Anti-Akratic Requirement (AAR): Any rationally permitted state in which I ought to have high confidence that, right now, I ought to have low confidence that p, is a state in which I do not, right now, believe that *p*.[103]

I argue that there are substantial tensions between anti-akratic and anti-tragic constraints on rationality. This should be a startling conclusion. Both constraints seem natural outgrowths of a picture of rationality as action guiding and tightly bound to the regulation of practical and theoretical reason. Defenders of anti-akratic constraints stress that akratic states license 'bad reasoning and irrational action' (Horowitz, 2013). And if there are cases in which,

[103] There are lots of worries one might have about AAR, and I suspect it impossible to formulate a version of AAR that avoids these worries entirely. One sort of worry is generated by the fact that one can have beliefs about one's own beliefs under different guises. Suppose you name one of your beliefs (*p*) 'Esmerelda'. Later on, you forget exactly which of your beliefs is Esmerelda, but an oracle tells you that Esmerelda is not rationally permitted. Your credence that any one of your beliefs is Esmerelda is pretty low. It seems fine for you to continue to believe that *p*. But then you rationally believe both Esmerelda and that you ought not to believe Esmerelda. A second kind of puzzle goes like this. Suppose there is a case in which high credence that *p* but not outright belief that *p* is permitted. (It is often thought that merely statistical evidence generates this sort of permissibility-structure: it is fine for me to have a high credence that you are guilty if 75% of defendants are guilty; it is not fine for me to believe you are guilty on this basis. See Buchak (2014) and Redmayne (2003) for discussion.) Suppose you rationally believe that high credence in *p* is permitted and you rationally believe that to believe *p* just is to have a high credence in *p*. It seems you can rationally believe that for you to believe that *p* is permitted. But then, by AAR, contrary to what we initially supposed, belief that *p* is permitted. This worry generalises. Suppose there is a case in which it is not permissible to believe *p* but it is permissible to disbelieve p's negation. You rightly believe that it is permissible to disbelieve p's negation, and rationally believe that to believe *p just is* to disbelieve p's negation. It seems you can rationally believe that it is permissible to believe *p*. But then it will not, contrary to our set up, be a case in which belief that *p* is impermissible. I do not think these worries are trivial. But this is not a paper particularly sympathetic to anti-akratic requirements, and none of these 'guise' flavoured worries relate closely to the troubling features of AAR I exploit here. A fuller treatment of these problems is a task for defenders of anti-akratic constraints.

through no fault of my own, there is no way for me to avoid acting irrationally, it is hard to see how rationality can be action guiding. Rinard (m.s.) notes that 'one central aspect of rationality is its guidance-giving role. Rationality is an ideal to which we aspire; the requirements of rationality are dictums to which we strive to conform.' But if rationality can tell me that there is nothing I am permitted to do, I cannot 'coherently aspire to conform to the requirements of rationality'.[104]

The plan for the rest of the paper is as follows. First, I use some standard cases to show how tasks involving series of infinitely many decisions interact with anti-tragic commitments for prudential rationality. Second, I show how these standard cases can be 'epistemicised', so as

[104] It is worth stressing the difference between the *epistemological* claim that we are often not in a position to know which actions are rationally permissible, and the tragic claim under consideration here, which says that there may be no rationally permissible courses of action available to faultless agents. The question of whether rationality has an anti-tragic structure is properly understood as orthogonal to the question of whether the norms of rationality are 'luminous', i.e. whether or not agents are always in a position to know whether some course of action is rationally permissible for them, for a tragedian might well think that whenever there is nothing an agent is permitted to do, said agent is always in a position to know that this is so. The differences between the two questions can be brought into sharp focus by considering their differing relations to a third question: that of what it takes for a norm to be action guiding. It has sometimes been argued that the norms of rationality will fail to be action guiding if they are not luminous (see Srinivasan, 2015). Srinivasan has argued that on a 'commonsense notion of guidance', norms can be action guiding without luminosity (2015). Her strategy, very roughly, is to say that norms can be action guiding so long as they, or something relevantly similar, are 'contextually luminous' (or luminous in certain contexts). Whether or not Srinivasan is right, there is obviously a *distinct* question as to whether the norms of rationality can be action guiding if there are cases in which they say that *nothing* we can do will be in accordance with them, and there is a sense in which rational dilemmas (cases in which nothing one can do is rationally permissible) stand in a deeper tension with a conception of rationality as action guiding than do anti-luminous epistemologies.

to generate interesting epistemological conclusions, as well as suggestive argumentative structures. Third, I show that cases with this structure reveal tensions between anti-tragic and anti-akratic constraints on rationality. Fourth, I show that these tensions cannot be avoided by endorsing what I call the 'moderate' anti-akratic view defended by Horowitz (2014). I conclude by sketching some different responses to the tensions I demonstrate.

1.1. Introducing Persephone, Hades & Co.

Arntzenius, Elga and Hawthorne tell the following story. You're Persephone, and you live on Mount Olympus, which you like a lot.[105] But you also like pomegranate. Hades has cut a pomegranate into an infinite number of pieces. If you eat a finite number of pieces of pomegranate, you get to stay on Mount Olympus, but if you eat infinitely many pieces, you will end up stuck in the underworld with Hades. Your first priority is to remain on Mount Olympus; your second is to eat as much pomegranate as you can. What does prudential rationality say you should do?[106]

[105] In Arntzenius et al., the story is about Eve and Satan. I have changed the names for the sake of consonance with the paper's Hellenic flavour. The structure of the example is unchanged.

[106] All claims about how Persephone ought to act, for the rest of this section, should be read as *prudential* oughts.

Arntzenius et al. answer: it matters how Hades asks you to make your choice. He might say: here are the pieces. Decide right now, at noon, how many you shall eat (50? 100? 7 million and one?). Call this the *synchronic* version of the puzzle. Or Hades might offer you the pieces one by one. At noon, he will say: 'Here is piece number one. Would you like to eat it?' Once you make a decision, Hades moves on to the second piece, at 12.30 and then the third piece, at 12.45, and so on, so that you will have been offered all the pieces by 1pm. Call this the *diachronic* version of the puzzle. There will, Arntzenius et al. point out, be no deep difference between the two versions of the puzzle if you are the sort of agent who can irrevocably bind herself to future courses of action.[107] Agents of this sort can make a binding decision concerning how many pieces of pomegranate they will eat regardless of how the pieces are presented to them. But you are probably not this sort of creature; most of us have future selves who will be able to revise the plans that our present selves have made. Let us assume, too, that every time-slice of Persephone knows that nothing she does *now* either makes any difference to the choices her future time-slices will make (she is not, for example,

[107] Not being able to leap 20 feet into the air is a failure of athleticism, not a failure of rationality, although that lack of athleticism determines *how one is able* to exercise one's rationality, as it determines which options one gets to pick between. Similarly, not being able to bind oneself irrevocably to future courses of action is not a failure of rationality, although it determines which options one gets to pick between, and determines how one can exercise one's rational abilities. A point along these lines is made by Arntzenius et al (2004). See also Hedden (2015b).

the sort of person who knows that if they have one piece they will probably end up gorging themselves on the rest of the pieces) nor gives her any new evidence concerning the choices her future time-slices will make.

Let us first look at the synchronic version of the puzzle. Anti-tragedians about prudential rationality will insist that Persephone can avoid acting impermissibly. Let us say that an *optimistic* anti-tragedian is one who insists (as do Arntzenius et al.) that Persephone can avoid acting impermissibly whilst also contriving to stay on Mount Olympus. Hades gives you the following argument that you should include piece #1 in your pomegranate-profile. 'Suppose that you include infinitely many other pieces of pomegranate. Then you will be stuck with me in the underworld whether or not you eat piece #1, and you may as well enjoy one more piece of pomegranate. Suppose that you include finitely many other pieces of pomegranate. Then this piece of pomegranate will not mean the difference between an infinite number of pieces and a finite number of pieces. But it will be pleasant to eat this piece of pomegranate. So whether you include infinitely many other pieces of pomegranate or finitely many other pieces of pomegranate, you should include piece #1. So you should include piece #1.' You are persuaded by this argument. Hades then moves on to piece #2. He gives you exactly the same argument for including piece #2, and you find it just as

persuasive. In the end, each piece of pomegranate is such that you are sure that it is better, regardless of what else you do, to include that piece of pomegranate in your profile than to exclude it.

Note that to be an anti-tragedian of any stripe (optimistic or not about Persephone's situation) you need to reject the claim that Persephone is permitted to pick a profile only if there is no better profile that she could have picked. For any profile Persephone picks is such that she could pick a better one: however large the finite number of pieces she chooses to include in her profile, she could have picked a profile including a larger finite number of pieces, and so could have got to eat more pomegranate. So if she is only permitted to pick a profile if there is no profile that is better, she is not permitted to pick any profile.

This is not a very appealing picture.[108] Consider an analogy. Suppose we have a carrot and a rabbit, Flopsy. Now, Flopsy loves carrot: the more carrot he gets to eat, the more pleasure

[108] It is, however, the picture given by a semantics for deontic modals on which '[d]eontic modals like *ought* or *have to* are necessity modals that claim that their prejacent is true in all of the best worlds. The corresponding possibility modal is *may*, which as the dual of the necessity modals claims that the prejacent is true in some of the best worlds (von Fintel, 2012.) On this picture, it is true that it is permitted that I pick some profile s only if some of the best worlds are worlds in which I pick an s-profile. Kratzer's (1977, 2012) semantics for deontic modals rejects, in rejecting the Limit Assumption, the assumption that there will always be some world which is either better than or just as good as every other world. Without the Limit Assumption

he experiences. And a world in which Flopsy experiences more pleasure is better than a world in which he experiences less. Things are set up like this: if I pick the number '1', Flopsy gets half a carrot. If I pick '2', Flopsy gets three quarters of a carrot. And if I pick '3', Flopsy gets seven eights of carrot, and so on: the bigger the number I pick, the more of the carrot Flopsy gets. But this means that for any number I pick, there is a number I could have picked that would have got Flopsy more carrot: for any *n* I pick, Flopsy would have got a bit more pleasure if I had picked n+1. But it is absurd to think that no matter which number I pick, I have acted impermissibly. A better thing to say is this: there are *some* numbers that I am permitted to choose *and* some that I am not permitted to choose, but it is very hard, often, to say whether any number is one or the other. Nevertheless, there are some numbers I can easily recognise that I ought not to pick (1, 2, 3) and some numbers I can recognise as permissible to pick (say, the largest known prime number).[109]

---

the semantics for necessity modals become more complicated: roughly, a sentence like 'It must be the case that p' will be true in a world *w* so long as *p* is true in any worlds that are better than w; and 'it is permissible that p' will be true in w just so long as it is not the case that it is required that not-p. But on these semantics, for any number *n* that you pick, it will be true that you are required to pick more than *n,* and so false that you are permitted to pick exactly *n.*If you really like a semantics along these lines, I will not be able to persuade you to give it up here. But I favour a model on which (i) a profile is permissible iff it features in a good enough ('satisficing') world, where a world can sometimes be satisficing even if there are better worlds, and so does not count as one of the *best* worlds, (ii) any best world is a satisficing world, and a profile is required if it is picked in every satisficing world, (iii) sometimes there are no satisficing worlds which are not best worlds, and (iv) sometimes it will be very hard to say whether a world is satisficing or not.

[109] It has been suggested that cases which are structurally similar to the case involving Flopsy, but in which the stakes are higher, generate quite different intuitions. Consider, for example, the following case, MURDER. In

Let us grant the anti-tragedian their verdict that Persephone can avoid acting impermissibly:

HAPPY: At noon, Persephone can avoid doing something impermissible between noon and 1pm.[110]

MURDER, Hades tells me he has picked out countably infinitely many people to be executed at noon. I can pick any natural number I want. Whichever n I pick, exactly n of the people presently marked for execution will be saved. Whichever number n I end up picking, it will be the case that I could have picked a larger number, and so, could have saved more people. And it is not obvious, the objector continues, that in this case, that I can avoid doing something impermissible. Now, certainly, the objector is correct that the raw intuitions elicited by MURDER are different to those elicited by the Flopsy case. But a few points should be made in reply. One point notes the following: regardless of the verdict we endorse in the case MURDER, the Flopsy case nevertheless puts pressure on the thought that one has acted impermissibly wherever one could have performed a better action. If one's pro-tragedy commitments flow from attraction to the thought that one has acted impermissibly wherever one could have performed a better action, the Flopsy case should put pressure on your pro-tragedy impulses, regardless of the verdict one ends up endorsing with respect to MURDER. A second, closely related point, is this: where the stakes are sufficiently high, one does not need to countenance cases in which there are infinitely many actions available to an agent in order to elicit get pro-tragedy intuitions. Consider Bernard Williams' (1994) classic case, in which twenty random people will be executed unless you yourself kill one of them, in which case, the other nineteen will be let off. This is, in one sense, a perfectly straightforward case: it certainly involves no fancy kind of supertask. But in another sense, the case is far from straightforward. Arguably, confronted with such a choice, there is *no way to avoid acting impermissibly*. If one is a committed tragedian, then, you probably do not need to ponder cases involving infinite series of actions to elicit intuitions favourable to your position. And if you are a committed anti-tragedian, it is not clear that cases involving infinitely many actions will pose any *special* problems. After all, any anti-tragedian will be committed to saying that actions which look, on the face of it, to be deplorable in some way, turn out to be permissible. It is unsurprising that there are special cases of this which involve infinitely many actions. The upshot of these considerations is as follows: if Persephone's case does turn out to be tragic, the tragedy will not flow from the *bare structure* of the case.

[110] The anti-tragedian need not endorse this claim if they are willing to say that Persephone is not in a permissible state at noon. Anti-tragedians should admit that if an agent has irrational preferences, they may not be able to avoid acting impermissibly. An example: if someone prefers A to B and B to C, but prefers c to A, they do something rationally impermissible if they pick A over c (if they prefer c to A they should pick C), and

Here are two more plausible claims about how Persephone should behave:

> RATIONALITY: When picking a complete profile, if taking some piece of pomegranate dominates not taking it, Persephone must pick a complete profile which includes that piece of pomegranate. [111]

> RATIONALITY*: If the only choice Persephone is making at noon is how many pieces of pomegranate to eat, and refraining from picking a profile including infinitely many pieces dominates picking a profile including infinitely many pieces, Persephone is rationally required to refrain from picking a profile including infinitely many pieces.

The anti-tragedian cannot have both. Let us see why. By RATIONALITY, every piece of pomegranate is such that you are required to pick a complete profile which includes that piece of pomegranate. But you are also sure that it is better to refrain from picking a profile

---

they do something impermissible if they pick *c* (if they prefer A to B and B to *c* they should pick A). But it is not appealing to think Persephone's starting preferences are irrational. See Hedden (2015b) for discussion.

[111] An option A *dominates* an option B if option A is better than A regardless of what the world is like.

which includes every piece of pomegranate. So by RATIONALITY* you are rationally required to refrain from including every piece in your profile. But then however many pieces of pomegranate you include, you do something impermissible. Suppose you include every piece of pomegranate. You were required to refrain from including every piece, so you did something impermissible. Suppose you do not include every piece. Then, given that every piece was such that you were required to include it, there was some piece that you were required to include but didn't. So you did something impermissible. But then you cannot avoid doing something impermissible.[112] And if we are *optimistic* anti-tragedians, we will prefer RATIONALITY* to RATIONALITY: it tells us that Persephone is *not* required to select a profile with an infinite number of pieces, and so that she is not required to end up with Hades in the underworld.

[112] Argument: if you really want to hang onto ANTI-TRAGEDY, RATIONALITY* will have to go. Consider: suppose Persephone considers picking a profile with exactly 70 pieces in it. She can be sure there is something better: picking a profile with exactly 71 pieces in it. And she can be sure there is something better than *this*: picking a profile with 72 pieces in it. And so on for every finite number. Counter-argument: refraining from picking a profile with exactly 70 pieces in it does not *dominate* picking a profile with exactly 70 pieces in it. For there are multiple ways a world in which one refrains from picking exactly 70 pieces might be: you might pick *fewer* than 70 pieces. This would not be better than picking exactly 70. But you might pick infinitely more than 70 pieces, in which case you get expelled from Mount Olympus, which is *worse* than picking exactly 70. So refraining from picking exactly 70 pieces is not better than picking exactly 70 pieces *regardless of what else happens*. It *is* true that we could not have both ANTI-TRAGEDY and BETTER, where BETTER says that if you are sure there is some a better option than φ-ing, you are required not to φ. Because for any finite number *n* Persephone can be sure that there is a better option than picking a profile including exactly *n* pieces: the option of picking a profile including n+1 pieces, for example.

For the optimistic anti-tragedian, the lesson of the synchronic case -- the lesson Arntzenius et. al. draw from it, at least -- is this. When you are deciding which profile to choose, you are making infinitely many choices: you are choosing whether to include piece #1 in your profile, and choosing whether to include piece #2, and choosing whether to include piece #3, and so on. So when one is making infinitely many choices, that doing such-and-such is a dominant choice does not mean that you are required to do such-and-such. If it did, Persephone would be required to eat every piece of pomegranate, even though there are clearly better options available (Arntzenius et al., 2004). In other words, the lesson of the synchronic case is that we should weaken RATIONALITY, so that it only tells us to pick dominant options when we are selecting between finitely many options:

> RATIONALITY$_f$: When making finitely many choices at t, if taking piece #n dominates not taking it, you are, at t, rationally required to take piece #n.[113]

So much for the synchronic case; what of the diachronic? The diachronic case holds two different lessons for the anti-tragedian. First, here is a plausible (and optimistic) thought: if

[113] There are background questions here about how to individuation, but I take it that everyone can agree that it is overwhelmingly natural to describe Persephone in the synchronic case as choosing between infinitely many options. See Buchack (2013) for discussion.

you start off in an OK situation, and you do something that improves your situation, doing *that* does not make you worse off. (You might think this more than plausible -- perhaps it verges on the tautological.) Here is a natural extension of this thought: if you start off in an OK situation, and you do countably many things, each of which makes you better off, performing this sequence of actions does not make you worse off. Call this COUNTABLE-RATIONALITY.

> COUNTABLE-RATIONALITY: If you perform a sequence of countably many actions, each of which improves your situation, you do not become worse off in virtue of performing this sequence of actions.

But anyone who likes HAPPY and RATIONALITY* must reject COUNTABLE-RATIONALITY. Suppose Hades offers you piece #1. Eating piece #1 improves your situation -- it does not expel you from Mount Olympus, and you get some pomegranate. Similarly, eating piece #2 improves your situation -- it does not expel you from Mount Olympus, and you get some pomegranate. For any #n, eating piece *n* will improve your situation. For either, before eating #n, you have only eaten a finite number of pieces of pomegranate, or you have already eaten an infinite number of pieces of pomegranate. If the

former, then eating piece #n will not be enough to expel you from Mount Olympus, so it does not make you worse off. And you get some more pomegranate than you had before, so it makes you better off. If you have already eaten an infinite number of pieces, you will be stuck in the underworld no matter what you do. So eating piece #n will not make you worse off. And it will get you an extra piece of delicious pomegranate, so it will make you better off. Each piece of pomegranate, then, is such that eating it will improve your situation (Arntzenius et al, 2004). But if you eat every piece of pomegranate, you will be expelled from Mount Olympus, and be much worse off than you were when you were on Mount Olympus before you had any pomegranate at all. But eating *every* piece of pomegranate makes your situation much worse (Arntzenius et al, 2004). And so at 1pm, when Persephone has eaten every piece of pomegranate, she would have been better off if she had acted prudentially impermissibly, i.e. had acted contrary to what prudential rationality told her to do. We already knew, thanks to Thrasymachus, that injustice can be profitable. But now it seems, startlingly, that prudential irrationality can be profitable too.

Second, in the diachronic case, the pairing of $\text{RATIONALITY}_{\text{f}}$ with anti-tragic commitments gets us the conclusion that Persephone ought to eat every piece of pomegranate (see Arntzenius et al., 2004). We can begin by noting that Persephone

certainly cannot wriggle out of eating every piece of pomegranate using the same strategy as in the synchronic case. In the synchronic case, the key to getting out of eating every piece of pomegranate lay in the insight that at the point at which you are choosing whether or not to eat an infinite number of pieces of pomegranate, refraining from picking a profile with infinitely many pieces dominates picking a profile with infinitely many pieces. But in the diachronic case, there is no point at which you have the choice between eating a finite number of pieces and eating an infinite number of pieces. For every piece of pomegranate that you are offered, either you will have already eaten a finite or an infinite number of pieces. And if you have already eaten only a finite number of pieces, choosing to eat *this* piece is not a choice between eating a finite or an infinite number of pieces, for whether or not you eat it, you will have eaten only a finite number of pieces. But if you have already eaten an infinite number of pieces, any choice you get is not a choice of whether or not to eat an infinite number of pieces, because whatever you choose to do, you will already have eaten an infinite number of pieces. So Persephone cannot get out of eating every piece of pomegranate in the diachronic version of the case in the same way as she was able to get out of eating every piece in the synchronic version of the case.

Another way to capture the difference between the synchronic and the diachronic case is to note that in the synchronic case, it is natural to say that you are making infinitely many choices at noon. But in the diachronic case, there is no point at which you are making infinitely many choices. At each point, it is natural to say, there are just two different things you can do: eat a certain piece of pomegranate, or not eat it. But a weakened version of RATIONALITY, which tells us to pick our dominant options when picking between finitely many options, will tell us, for each piece of pomegranate, that we are rationally required to eat it: $\text{RATIONALITY}_{\text{f tells}}$ us that every time Persephone has to choose between eating a piece of pomegranate and not eating it, Persephone is rationally required to eat it (Arntzenius et al, 2004).[114] If you want to resist the thought that Persephone is

[114] As Arntzenius et al. point out, the case with Hades and the pomegranate shares a structure with interpersonal prisoner's dilemmas, in which a group is worse off in the case where every member of the group picks their dominant options than the group is in the case where at least some members of the group fail to pick their dominant option. Arntzenius et al. advocate treating Hades' Pomegranate in the same way prisoner's dilemmas are standardly treated (see Hedden 2015b), in which it is acknowledged that each member of the group would prefer that at least one member fail to select a dominant option (just as each time-slice of Persephone would prefer that some of her time-slices decline pomegranate) but it is denied that any member of the group has acted irrationally, or that the group has acted irrationally, when each of its members select its dominant options. There are alternative treatments of interpersonal prisoner's dilemmas which rely on what Bacharach (2006) calls *agency transformation.* Agency transformation means 'that the agent conceives the situation – not as a decision problem for individual agents – but as a decision making problem for the group conceived as an agent' (Hakli, Miller & Tuomela, 2010). Group agents can decide which action is to be undertaken by each of its members, because they can 'specify the strategies of all the agents in the group' (Hakli, Miller & Tuomela, 2010). Let a *diachronic group* be a group agent composed of different time-slices of the same agent. If each of Persephone's time-slices conceives of Hades' Pomegranate as a problem for the diachronic group Persephone, we might be able to extend the agency transformation strategy even to diachronic cases in which Persephone cannot self-bind.

required to expel herself from Mount Olympus, there are a few different ways you might go. None of them are appealing for the anti-tragedian.

You might read the case as a *reductio* of RATIONALITY$_f$, and say that at least one piece of pomegranate must be such that Persephone is *not* required to eat it, even though every piece is such that eating it dominates not eating it. This does not look like a good way to go. Exactly the same considerations are in play with respect to every piece of pomegranate. But it is hard to see how differences of rational permissibility might emerge without there being different considerations in play. So either every piece is such that it is impermissible to eat it, or every piece is such that it is permissible to eat it. But the former suggestion is clearly false: suppose Persephone eats 100 pieces and then, before she can make any other choices, a stray gust of Olympian wind blows all the other pieces away. Persephone has not done anything impermissible just by eating those 100 pieces.

Second, you might keep RATIONALITY$_f$, but be suspicious of agglomeration principles:

COUNTABLE-AGGLOMERATION: If each of a countable series of actions is such that you ought to perform it, you are required to perform every member of the series.[115]

Denying COUNTABLE-AGGLOMERATION will allow you to say that although every piece of pomegranate is such that Persephone is prudentially required to eat it, Persephone is prudentially required to refrain from eating an infinite number of pieces of pomegranate.

The trouble is, it is hard to argue like this without imperilling HAPPY. By RATONALITY$_f$ every piece of pomegranate is such that Persephone must eat it. So if she fails to eat any piece of pomegranate, she has done something she shouldn't. But to refrain

[115] The corresponding principle for permissibility ('COUNTABLE-PERMISSIBILITY') would tell us that if each of a countable series of actions is such that you are permitted to perform it, you are permitted to perform every member of the series. COUNTABLE PERMISSIBILITY is less appealing than COUNTABLE-AGGLOMERATION insofar as agglomeration principles for permissibility are generally less appealing than they are for requirements. In general, that I am permitted to do such and such and that I am permitted to do so and so does not mean that I am permitted to do both. For example, I might be permitted to press the red button and permitted to press the green button but not permitted to press both (maybe if one gets pressed, Flopsy gets a carrot, but if both get pressed, Flopsy the rabbit dies). But suppose we think of two actions as being in some series of actions only if no time at which one is happening is a time at which the other is happening. Then it is not clear that such cases tell against a principle which says that if I perform a series of finitely many actions, each of which is permissible, I have done nothing impermissible. For suppose I press the green button at noon and the red button at ten past, and kill Flopsy: the natural thing to say is that once I pressed the green button, it was impermissible for me to press the red button, so I have not performed a series of actions where each action is permissible.

from finishing the pomegranate Persephone must fail to eat at least one piece of pomegranate. So at noon, however much pomegranate Persephone will end up eating, she will do something she should not. So at noon Persephone cannot avoid doing something she should not between noon and 1pm.[116] But by HAPPY, at noon Persephone *can* avoid acting impermissibly between noon and 1pm. What if we say that Persephone is *permitted* to refrain from eating every piece, but not that she is *required* to refrain from eating every piece? This similarly imperils anti-tragic commitments. But by RATIONALITY$_f$, if Persephone does anything other than eat every piece, she has done something impermissible. So there is no way for her to leave the pomegranate unfinished without having done something impermissible.

A third way to go is to say that when Persephone finishes the pomegranate, Persephone has not done anything impermissible, but that she is nevertheless in an impermissible state. But this goes against one of our core anti-tragic constraints. If FORBIDDEN is true, Persephone cannot have started in a permissible state at noon and be in an impermissible state at 1pm unless she has done something impermissible between noon and 1pm.

[116] To be clear, I am not arguing that everyone must accept COUNTABLE-AGGLOMERATION, just that the anti-tragedian ought to accept it.

Let us take stock. If you like anti-tragic constraints on rationality, along with $\text{RATIONALITY}_{f,}$ you will get the conclusion that Persephone is rationally required to eat every piece of pomegranate in the diachronic version of the case. You will end up with what we might call a hyper-Thrasymachian conclusion. If Thrasymachus' slogan is 'injustice is profitable', hyper-Thrasymachus' slogan is 'prudential irrationality is profitable'. For the anti-tragic position grants that at noon, Persephone can avoid doing something she shouldn't between noon and 1pm. And it grants that so long as she avoids doing what she shouldn't, she avoids being in an impermissible state. But it ends up committed to the thought that, sometimes, doing just what prudential rationality says you *should* do, can make you *really badly off*, and that prudential rationality can compel you to act in ways that make you worse off (along some axis) than you would be if you had acted impermissibly. This will be a misfortune, certainly, but not all misfortune is tragedy: Persephone will be able to comfort herself in the underworld with the thought that she acted perfectly rationally.

2. Epistemicising Hades' Pomegranate.

Hades' Pomegranate makes vivid some of the difficulties embedded in the anti-tragic position. It shows us one way in which prudential rationality can require an agent to make themselves much worse off than they would be if they were to act irrationally. Cases with a similar structure can make similarly vivid some sobering consequences of the anti-tragic attitude in epistemology. Let us remind ourselves of the epistemicised anti-tragic commitments:

ANTI-TRAGIC$_e$: If you are in an epistemically permissible state at t_1, you can, at t_1, avoid acting epistemically impermissibly between t_1 and t_2.

FORBIDDEN$_e$: If you start off in a permissible belief state at t_1 and end up in an impermissible belief state at t_2 then you have done something epistemically impermissible between t_1 and t_2.

COUNTABLE-AGGLOMERATION$_e$: If each of a countable series of actions is such that you epistemically ought to perform it, you epistemically ought to perform every member of the series.[117]

2.2. Hades' Salmon.

Some decisions improve our epistemic position; others worsen it. Plausibly:

EPISTEMIC-RATIONALITY: When your only choice at t is the choice of whether or not to Θ, and, regardless of what else you do, Θ-ing will improve your epistemic position more than refraining from Θ –ing, then, at t, you epistemically ought to Θ.

[117] I noted above that COUNTABLE-PERMISSIBILITY is substantially less appealing as a principle than COUNTABLE-AGGLOMERATION. An epistemicised version of COUNTABLE-PERMISSIBILITY, together with ANTI-TRAGEDY$_e$, would have significant epistemological consequences. For example, it looks to be incompatible with what Kelly (2014) calls 'intrapersonal permissivism' – the view that given a certain body of evidence there can be a range of mutually exclusive attitudes that I am rationally permitted to take towards *p*. Suppose Glaucon is permitted to have any credence smaller than 0.8 and greater than or equal to 0.7. Suppose Glaucon starts off at Cr(p) = 0.7. At 12.30 he adjusts his credence to 0.75. At 12.45, he adjusts it to 0.775. And so on. If Glaucon makes every upwards adjustment, he will end up with Cr(p) = 0.8.The permissivist has no easy way to resist the thought that each credence adjustment will permissible. If we liked COUNTABLE-PERMISSIBILITY, we could conclude that he is permitted to make all of the credence adjustments, and so, by ANTI-TRAGEDY, that he is permitted his credence of Cr(p)=0.8. Thankfully for those inclined towards permissivist views, COUNTABLE-PERMISSIBILITY is not compelling, for reasons discussed above.

Thrasymachus insists to Socrates that injustice is profitable; and just as we can epistemicise anti-tragic commitments, we can epistemicise the dispute between Thrasymachus and Socrates. Our ventriloquised Thrasymachus will insist that *epistemic irrationality is profitable:* that acting as one is epistemically required not to act can improve one's epistemic position more than acting as one epistemically ought to act. Socrates will demur:

> ANTI-THRASYMACHUS: No one is ever epistemically required to worsen their epistemic position.

But no one with anti-tragic commitments will be able to endorse both ANTI-THRASYMACHUS and EPISTEMIC-RATIONALITY. Suppose you are Hercules. Just before noon, you have just eaten exactly 1 kg of salmon, and you believe, just before noon, that you have eaten at least 1 kg of salmon today.[118] Let your degree of confidence just before noon, in the claim that you have eaten at least1kg of salmon today, be n, where $0 < n < 1$.

[118] Let this be the strongest proposition you have any doxastic attitude towards.

Plausibly, you will be in a better epistemic position at 1pm – at least with respect to the claim that you have eaten at least 1kg of salmon today – if your rational degree of confidence in said claim is higher at 1pm than it is just before noon. More generally, for any time t from noon onwards, it is plausible that you will be in a better epistemic position at t – at least with respect to the claim that you have eaten at least 1kg of salmon today – if your rational degree of confidence in said claim at t is higher than it was just before noon. Enter Hades. Hades has with him 1 kg of salmon sliced into a countably infinitely many pieces (the first piece is 500g, the second piece 250g, the third piece 125g, and so on). At noon, Hades will offer you piece #1, at half past noon, if you accepted piece #1, he will offer you piece #2, at quarter to one, if you accepted piece #2, he will offer you piece #3, and so on, so that by 1pm, so long as there is no piece of salmon you have refused, Hades will have offered you every piece of salmon. Every time you eat a piece of salmon, your rational degree of confidence that you have eaten at least 1 kg of salmon will slightly increase. For example, suppose you accept Hades' offer of piece #1. If, at noon you eat piece #1, you will have eaten 1.5 kg of salmon, and your degree of confidence that you have eaten at least 1kg of salmon will increase to $n+k$. Let k = 0.1. Suppose you go on to eat piece #2. You will then have eaten, at half past noon, 1.75kg of salmon. Your degree of confidence that you have eaten at least 1kg of salmon will increase to $n + 1.5k$ (and so, where k = 0.1, your degree of

confidence will increase to n+0.15). In general, if you eat piece #m, your degree of confidence will increase by $k/2^m$, and once you have eaten m pieces, your degree of confidence will have increased by $\sum_{m=0}^{m} k/(2^m)$. If you eat every piece, at 1pm, your degree of confidence that you have eaten at least 1kg of salmon today, given that $\sum_{m=0}^{\infty} 0.1/(2^m) = 0.2$, will be $n + 0.2$.

Here's the rub: if you eat infinitely many pieces of salmon, you will become seized with an irrational conviction that you can breathe underwater (after all, you will have eaten a *lot* of salmon!) Regardless of whether your rational degree of confidence that you have eaten at least 1 kg of salmon is higher or lower, it is clearly epistemically better to lack an irrational conviction that you can breathe underwater than to have such an irrational conviction.

Eating piece #1 clearly improves your epistemic position regardless of what else you do. If you eat finitely many other pieces, you eating piece #1 will get you a higher rational degree of confidence that you have eaten at least 1 kg of salmon today. And if you eat infinitely many other pieces, you will be stuck with your irrational belief that you can breathe underwater in any case, so you may as well up your rational degree of confidence that you have eaten at least 1 kg of salmon today. We can run the same argument for piece #2, and

piece #3, and so on. Each piece of salmon will be such that it improves your epistemic position to eat it more than not eating it will improve your epistemic position. So by EPISTEMIC-RATIONALITY, every piece of salmon is such that you are required to eat it. And so by familiar anti-tragic reasoning, if there is any way for you to avoid doing something you ought not to do, it involves eating every piece of salmon. So the anti-tragedian must think that you epistemically ought to finish the slab of salmon. The trouble is, if you end up eating *every* piece of salmon, you will end up in a much worse epistemic position than you were before you had eaten any salmon, for if you finish the salmon, you will irrationally believe that you can breathe underwater! But there is no piece of salmon such that the decision to eat that piece of salmon worsens your epistemic position. Think of it this way: which piece could it be? The only plausible answer is: the piece prior to which you had not eaten an infinite number of pieces of salmon, and after which, you have eaten an infinite number of pieces of salmon. But there is no such piece. An anti-tragedian who likes EPISTEMIC-RATIONALITY cannot have ANTI-THRASYMACHUS. They must admit that epistemic rationality can require agents to worsen their epistemic position.

2.2. Knowledge Banking.

It is worth taking some time to see how a different case (modelled on another case from Arntzenius et al., (2004)) can reinforce the Thrasymachian conclusion.[119] Suppose you think that knowledge is very epistemically valuable. If you have this kind of view, it is natural for you to think that if you are offered a choice between *n* pieces of knowledge and n+1 pieces of knowledge, then, *ceteris paribus*, you should always pick getting n+1 bits of knowledge:

> KNOWLEDGE: If your only choice at t is between *n* bits of knowledge and n+1 bits of knowledge, then, *ceteris paribus*, at t. you epistemically ought to choose n+1 bits of knowledge. [120]

[119] The original case is something like this. Suppose at noon Donald is given ten roubles labelled 1-10, and at half past noon, has the choice of sticking with what he has or getting another 10 roubles (labelled 11-20) and burning note number 1. Donald should choose getting the roubles and burning: that way he gets an extra nine roubles. But he is offered another choice fifteen minutes later: stick with what he has, or get another ten roubles (labelled 21-30) and burn note number 2. And so on, until 1pm. If Donald never chooses to stick with what he has, he will end up with nothing. For eventually, he will burn every rouble he gets. And so long as it is true that Donald has a rouble at 1pm only if Donald has received some rouble and not burnt it by 1pm, if he has burnt every rouble he gets by 1pm, at 1pm he will have no roubles.

[120] Some remarks about this principle are worth making. (i) The *ceteris paribus* clause is doing a lot of work here. Maybe you think it is better for you to know Newton's first law than which is the 63rd word to appear in *Harry Potter and the Philosopher's Stone.* The *ceteris paribus* clause controls for this. (ii) Maybe you are worried about the cogency of the ideology of 'bits of knowledge'. The worries prompted by this ideology seem akin to the worries prompted by attempts to count and individuate beliefs; Pritchard notes that it is often tacitly assumed by epistemologists that *some* sense can be made of the notion of 'counting beliefs' (2015). For discussion see Dennett (1966), Pritchard (2014) and Treanor (2014). (iii) As it is, KNOWLEDGE makes no reference to whether you are keeping the knowledge in question, or gaining new knowledge. It may be that there is some epistemic badness associated with losing knowledge which is not outweighed by just any (however small) net epistemic gain. If this is so, then the principle KNOWLEDGE will be false, as it may be that sometimes you ought to stick with what you have rather than lose it, even if you would be compensated

Suppose Persephone is about to be presented with a countable sequence of choices such that by 1pm, she might have no new knowledge, or she might have lots of new knowledge, depending on the choices she makes. If you think knowledge is very epistemically valuable, the following claim will be appealing:

> END-POINT: If Persephone ends up with no new knowledge at 1pm, she has done something she epistemically ought not to have done.

Hades comes along and, at noon, gives Persephone a new bit of knowledge. Then at half past noon, he gives Persephone a choice between 'sticking' or 'adding'. He tells Persephone this: 'either you get to keep that one bit of knowledge ('k_1') I just gave you ('stick'), or you

---

with more than you originally had. This is relatively plausible. But what is not plausible is that losing knowledge is so bad that if you have a choice between sticking with what you have or losing what you have and getting *lots* more knowledge, then it is fine to stick with what you have. In other words, even if KNOWLEDGE is false, something like KNOWLEDGE' will be true:

> KNOWLEDGE': If you have a choice between *n* bits of knowledge and n+100 bits of knowledge, then, *ceteris paribus*, you are epistemically required to choose n+1 bits of knowledge.

But so long as KNOWLEDGE' is true, the same dynamics will be in play as when we exploit KNOWLEDGE. Suppose you learn that k_1, and then you have the choice of sticking, or getting 100 new pieces of knowledge $k_2 \ldots k_{101}$, and forgetting k_1, and then the choice of sticking or getting 100 new pieces of knowledge and forgetting k2, and so on. To end up with new knowledge by adding every time, there would need to be *n* such that you never get the choice of whether to forget k_n. But so long as there are enough choices, for any *n,* you will eventually have the choice of forgetting k_n.

get two new bits of knowledge, but lose k_1 ('add')'. [121] Then at 12.45 he tells her: 'either you get to keep what you have ('stick'), or you forget k_2 and k_3, but get three new bits of knowledge ('add').' Then seven and a half minute later, he tells her 'either you get to keep what you have ('stick'), or you lose the three bits of knowledge you now have and get four new bits ('add')'. And so on: Hades will always offer Persephone the choice of keeping the *n* bits of knowledge she has (stick), or losing all *n* bits of knowledge she has, but getting n+1 new pieces of knowledge (add). By KNOWLEDGE, then, every time she has a choice, she is required to add. But then, if she does this, will Persephone have acquired any knowledge at 1pm?[122]

We can argue like this. Every piece of knowledge she has gained will also have been destroyed. For there to be some knowledge that has not been lost, there would need to be some *n* such that she could not be offered n+1 new pieces of knowledge. But there is no

[121] Perhaps you want a concrete example of the sort of series of choices Persephone might be asked to make. imagine Persephone is choosing between: know whether *n* is odd or even, or know whether n+1 and whether n+2 is odd or even, for some arbitrary n.

[122] It is worth noting that a superficially similar series of choices would have very different consequences. Suppose Persephone gets a piece of knowledge at noon and at half past has the choice of sticking with what she has, or getting two new pieces of knowledge and then having Hades toss a coin to decide which of the new pieces he will destroy. At 12.45, she gets the option of sticking with what she has, or getting two new pieces of knowledge and then having Hades toss a coin to decide which of the new pieces he will destroy. And so on. If Persephone chooses to add each time given this series of choices, she will end up with an infinite number of pieces of knowledge. See Arntzenius et al. (2004) for discussion.

such *n.* But then, if Persephone has will lose any knowledge she gains between noon and 1pm, she will end up with no new knowledge at 1pm. [123] So by END-POINT, Persephone will have done something she should not by 1pm. What can the anti-tragedian say? If Persephone is required to perform each of the countable sequences of adding, but does something impermissible if she adds every time, then Persephone cannot avoid doing something wrong. So if Persephone can avoid doing something wrong, and we like KNOWLEDGE, then END-POINT needs to go. Epistemic rationality can require Persephone to end up with no new knowledge at 1pm; in order to end up with any new knowledge at all, she must act irrationally. We have another triumph for Thrasymachus.

3.1. Akrasia and Tragedy.

So far, we have drawn out some interesting consequences of anti-tragic commitments in epistemology. In this section, I show that there are tensions between anti-tragic and anti-akratic requirements. I show that an agent can start off in a permissible belief state and end up in an akratic state *without performing any epistemically impermissible act.* It follows that

[123] I am assuming that Persephone knows at 1pm something she did not know at noon only if she has received a new bit of knowledge and failed to lose it between noon and 1pm. This is analogous to the assumption in Arntzenius et al. (2014) that Donald has some roubles at 1pm only if he has not burnt every rouble.

we must reject the claim that akratic states are always impermissible, or become tragedians, either by admitting that agents can end up in rationally impermissible states without doing anything wrong, or by admitting that 'guiltless agents' in permissible states may be forced to choose between two ways of acting impermissibly.

3.2. Hades' Encyclopaedic Scroll

Suppose that you read a book, and the author tells you 'I have done my best to check everything I say here. I make no claim in this book of which I am not highly confident. But I am confident that I have made some mistake *somewhere*: it is overwhelmingly probable that despite my best efforts, some claim I make here is false.' In such a case, you ought to be highly confident of each claim in the book, but also be highly confident that their conjunction is false. This is the paradox of the preface (see Makinson, 1965), and it comes in different flavours. And here it is in one of those other flavours: suppose you know that with books with more than 100 pages, if you read every page, you are highly likely to end up misremembering something you have read. For every individual claim *c* that you believe on the basis of reading the book, you should be highly confident in *c*. But you should be highly

confident that you have misremembered what was written on at least one of the scroll's feet. It is this second flavour of case that is of interest to us here.

Hades knows all about the paradox of the preface, and decides to exploit it to put Apollo in a difficult situation. Hades has a scroll which is countably infinitely many feet long. And Hades knows a conjunction with countably infinitely many conjuncts, *INF*. *INF* is not something that Apollo knows at noon – indeed, at noon, Apollo ought to have low confidence in *INF*. Hades has written a sentence expressing this conjunction on his scroll (each conjunct of the sentences takes up a foot of scroll).

Let each foot of scroll be labelled with one of the natural numbers. The first foot is labelled foot #1, the second foot, foot #2, the third foot, foot #3, and so on. At noon, Hades will begin by offering Apollo the chance to read foot #1. Then at half past noon, Hades will offer Apollo the chance to read foot #2. (Recall that a scroll can be rolled up again, as well as unrolled.) Then at 12.45, Hades will offer Apollo the chance of reading foot #3. And so on, until by 1pm, Apollo will have been offered the chance to read every foot of scroll. Let us label, too, the conjuncts written on the scroll: let p_1 be the conjunct written on foot #1, p_2, be the conjunct written on foot #2, and so on. For any foot of scroll #n, let the conjunct

written on foot #n be p_n. Now, Apollo's memory is pretty good, but it is not perfect. Let us then, also have a way of talking about the conjuncts that Apollo will seem, at a given time, to remember reading. These may not always be identical with the conjuncts he *actually* read. Where p_n is a conjunct that Apollo has read on the scroll, let $m_{t(p_n)}$ be the proposition such that, at time t, Apollo seems to remember reading q at foot *n.* Now, let us make the following continuity assumption about the relation between which feet of scroll Apollo reads *by* 1pm and what he seems to recall reading at 1pm.

> CON_1: Let *c* be 1pm. If, by 1pm, Apollo has read countably many conjuncts p_1, p_2, p_3,….then at 1pm, Apollo will seem to remember reading $m_{c(}$ $p_{1)}$, $m_{c(p_2)}$, $m_{c(p_3)}$,…, such that the cardinality of the set of conjuncts that Apollo has read by 1pm will be the same as the cardinality of the set of conjuncts which Apollo will seem to remember reading at 1pm.

More formally, let *c* be 1pm, and let *f* be a function that takes us from time to the natural numbers in such a way that $f(t) = m$, where *m* is the number of conjuncts believed by Apollo at *t*. CON_1 guarantees that us that $f(c) \geq \lim_{t \to c} f(t)$. No matter how much scroll Apollo

ends up reading by 1pm, there will be no foot of scroll *n* she has read such that there is not, at 1pm, some proposition which she takes herself to have read at foot *n*.

Hades explains that it is important to remember that, so long as Apollo reads more than 1 foot of scroll, he will end up *having read a conjunction.* After all, Hades has not just written countably many non-conjunctive claims on his scroll; he has written *one very long conjunction*. For example, if by 12.40 Apollo has read both foot #1 and foot #2, by 12.40, Apollo will have read the conjunction p_1 & p_2. Let us say, then, that Apollo satisfies the following additivity constraint:

> CON_2: If, at some time t, Apollo takes himself to remember reading countably many conjuncts $m_t(p_1)$, $m_t(p_2)$, $m_t(p_3)$,…, then, at t, Apollo believes their conjunction c, just so long as there is no conjunction c' with (i) more conjuncts than *c* and (ii) no conjunct c* which is not such that Apollo takes himself to remember reading c*.[124]

[124] Special thanks go to Timothy Williamson for discussion on these points.

Let us consider a case in which Apollo reads finitely many feet of scroll by 1pm. Say he has, by 1pm, read exactly *n* feet of scroll, and not turned down any foot of scroll. From $CON_{1,}$ it follows that at 1pm, there are *n* propositions Apollo takes himself to remember reading, $m_c(p_1)$, $m_c(p_2)$, $m_c(p_3)$,…$m_c(p_n)$. From CON_2 it follows that at 1pm, Apollo believes the conjunction $m_c(p_1)$ & $m_c(p_2)$ & $m_c(p_3)$,…$m_c(p_n)$. Let us consider a case in which Apollo reads countably infinitely many feet of scroll by 1pm. It follows from CON_1 that at 1pm, Apollo takes himself to remember reading countably many conjuncts. It follows from CON_2 that at 1pm, Apollo believes some conjunction *c* with countably many conjuncts. (No conjunction with countably many conjuncts will have more conjuncts than *c*.)

Now, there are two different versions of the puzzle: a finitary version and an infinitary version. In the finitary version, Apollo knows that he is generally pretty good at accurately remembering what he has read. But, Satan tells him, if he reads more than 999 feet of the scroll, it will release a strange gas which interferes with Apollo's memory. Once the gas is released, no matter how many feet Apollo reads, it becomes very likely that, by 1pm, Apollo will end up misremembering at least one of the conjuncts written on the scroll. That said, there is a limit to how badly the gas can affect Apollo's memory: the worst it can do is make

Apollo misremember what is written on a single foot of the scroll (it is hard to find gases that interfere too badly with an Olympian's memory).

Things are set up similarly in the infinitary version, except that the gas is released not if more than 999 feet of the scroll are read, but only if *infinitely* many feet of the scroll are read. Again, once the gas is released, it becomes very likely that Apollo will end up misremembering one of the conjuncts written on the scroll. But again, there is a limit to how badly the gas can affect Apollo: the worst it can do is make Apollo misremember what is written on a single foot.

3.2. Hades' Finite Encyclopaedic Scroll

In the finitary case, there are two different ways that things might go. On the one hand, Apollo might not read more than 999 feet by 1pm. On the other, Apollo might read more than 999 feet by 1pm. And so long as Apollo has read at most 999 feet by 1pm, everything will be fine.

But suppose instead that Apollo reads more than 999 feet of scroll by 1pm. Let us suppose that by 1pm, Apollo has read exactly n feet of scroll, where n is some finitely large number

which is larger than 999, and that Apollo has refused no foot of scroll. Then at 1pm, Apollo will believe some conjunction $m_c(p_n)$ & $m_c(p_{n-1})$ & $m_c(p_{n-2})$….$m_c(p_1)$. Call this conjunction 'CON'. Let's suppose that against the odds, although the gas has been activated, Apollo has managed not to misremember anything he has read.

AAR tells us that any permissible state in which Apollo ought to be highly confident that he should have low confidence that *p* is a state in which he does not believe that *p.* So if the state that Apollo is in at 1pm is a state in which he ought to be highly confident that he should have low confidence that *CON,* and Apollo believes *CON,* then Apollo is in an impermissible state. And it seems quite right that at 1pm, Apollo ought to be highly confident that he should have low confidence that CON. For Apollo can easily tell how many feet he has read: each foot of the scroll comes marked with a number. And Apollo was rationally highly confident at noon that if he read 1000 or more feet, he would have misread what was written on one of them. Just the process of reading *n* feet of scroll ought not to alter Apollo's degree of confidence that this is so. So he was rationally highly confident at noon that if he read 1000 or more feet he would be very likely to misremember one of them. So at 1pm, Apollo ought to be very highly confident that he has misremembered one

foot of scroll, and so, should be very highly confident that he did not in fact read one of CON's conjuncts.

Of course, there will be cases in which one ought to be highly confident that one did not read *p* and yet in which one should not be highly confident that one should have low confidence in *p*. Suppose I seem to have read in the newspapers that Viktor Krum and Hermione Granger are dating. A few days later, my gossip-columnist friend tells me that they have discovered a tremendous secret, unknown to anyone else in the press: that Viktor Krum and Hermione Granger are dating! My friend generally knows a good deal more than me concerning, and is very reliable on the subject of, who is dating who and who knows what about who is dating who, and I know that I may well have misremembered what I read in the paper (I know that I am forever mixing up Viktor Krum with other famous quidditch players). In this case, I should be both (i) highly confident that I did not read *p*, and (ii) be highly confident that *p*. And I ought not to think this high confidence misplaced. But many cases are not like this: suppose I mention to my friend that I have read in the papers that Viktor Krum and Hermione Granger are dating, and she says that she is sure no such story has been published. In this case -- a case in which I should have something other than low confidence that *p* only if I have read that *p* -- I should be highly confident that if I

did not read that *p*, I should have low confidence that *p*. We can apply this lesson to preface cases. Suppose we have a conjunction c_1 & c_2…&…c_n, and you are very highly confident, of each conjunct *c* that you should have low confidence in *c* unless you read that *c*. So long as you are very highly confident that you did not read one of the conjuncts, you should be highly confident that you should have low confidence in the conjunction. [125] We can, in turn, apply this lesson to Apollo's situation, with the result that Apollo ought to be highly confident that he should have low confidence in CON. Consider: Apollo should be very highly confident that he did not read one of CON's conjuncts, and none of the conjuncts of *CON* are such that if Apollo did not read them, Apollo should have more than low confidence in them. So Apollo ought to be very highly confident, of each conjunct of *CON,* that unless he read that conjunct, he should have low confidence in it. The result is that at 1pm, Apollo ought to be highly confident that he should have low confidence in CON. So by AAR, he is in a permissible state only if he fails to believe that CON.

The instincts elicited by Hades' Finite Encyclopaedic Scroll dovetail very nicely with the instincts of the anti-tragedian and the anti-akratic. It is highly intuitive to think that Apollo

[125] That you ought to be highly confident that *p* and that if *p* than q does not guarantee that you ought to be highly confident that q. But that you ought to be very highly confident that *p* and that if *p* then q does guarantee that you ought to be highly confident that q.

ought not to read more than 999 feet, and that if he does, he will end up in an impermissible state; anti-akratic constraints give us a sharp and satisfying way of accounting for what is wrong with the belief state that Apollo ends up in if he reads more than 999 feet of scroll before 1pm. Things are less friendly to the pairing of anti-tragic and anti-akratic commitments in the infinitary case. Let us see why.

3.4. Hades' Infinite Encyclopaedic Scroll.

Hades asks Apollo, 'Suppose that you are sure the conjunction I have written on my scroll is true. Every foot of the scroll is such reading it will improve your epistemic position more than not reading it, whatever else you do. And I will offer you the chance of reading each foot of the scroll, one by one. Do you think that every time I offer you the chance to read a foot, you should (that is an epistemic should) read it?' Apollo ponders the question, and then replies, 'Yes'. As such, Apollo is inclined to endorse the following compelling claim:

> REQUIRED: If the only choice Apollo has at t is whether or not to read foot #n, and, no matter what else he does, reading foot #n will improve Apollo's epistemic

position more than not reading it, then at t, Apollo epistemically ought to read foot #n.[126]

Apollo is sure, recall, that if he reads only finitely many feet of scroll, he should be pretty confident that he will not end up misremembering anything. But he is sure that if he reads infinitely many feet of scroll, he is extremely likely to end up, by 1pm, misremembering what is written on one of the feet. Then Hades asks Apollo whether he would like to see the first foot of scroll. He argues like this: 'Look, however many other feet you end up reading, you ought to read this one. If you end up reading only finitely many other feet, then reading this one will not be enough to make the difference between infinitely and finitely many feet of scroll, and it will clearly be better for your epistemic positon to have read this foot rather than not. And if you end up reading infinitely many other feet of scroll, then reading this one will not make you any *more* likely to misremember anything -- the most the gas can do is ruin your memory of a single foot. But c*eteris paribus*, it is clearly better to read this foot of scroll than not. So by REQUIRED, you ought to read this foot of scroll.' Apollo is persuaded, and reads the first foot. Half an hour later, Apollo gets the chance to read foot #2. Hades uses the same argument, and once again, Apollo is persuaded. Apollo, then, is

[126] This is a special case of EPISTEMIC-RATIONALITY.

persuaded that he epistemically ought to read every foot, as every foot is such that reading it will improve his epistemic position more than not reading it, regardless of what else he does. So by 1pm, Apollo will have read countably many feet of scroll. As discussed above, from CON_1 and CON_2, it follows that at 1pm, Apollo believes a conjunction with countably infinitely many conjuncts. Call this conjunction *INF*. Just as we supposed in the finite case that, although it was very likely that Apollo would misremember something after reading 1000 feet of scroll, he actually managed not to misremember anything, let us suppose here that although it was very likely that after reading infinitely many feet of scroll, Apollo would misread something he read, he has actually managed not to.

By the lights of the anti-akratic, Apollo cannot avoid either (i) doing something he should not or (ii) ending up in an impermissible state by 1pm. AAR tells us that any permissible state in which Apollo ought to be highly confident that he should have low confidence that *INF* is a state in which Apollo does not believe that *INF*. By REQUIRED, at 1pm, Apollo will have done something wrong unless he believes that a conjunction with countably infinitely many conjuncts, as to avoid believing a conjunction with countably infinitely many conjuncts, he would have had to refuse to look at at least one foot of scroll, contrary to what REQUIRED demands. But it seems that if Apollo has looked at infinitely many

feet of scroll, then he ought to be highly confident that he should have low confidence in *INF.* The argument here is exactly parallel to the argument in the finite case that, at 1pm, Apollo ought to have high confidence that he should have low confidence that CON. But it is worth going through each stage of the argument in turn. We are supposing that Apollo has read every foot of the scroll. Under these circumstances, Apollo can easily tell that he has read infinitely many feet of scroll: there will be no scroll left to read! And Apollo was sure at noon that if he read infinitely many feet by 1pm, he would very likely end up misremembering a conjunct by 1pm. Just the process of reading infinitely many feet of scroll, none of which say anything to the contrary, ought not to alter Apollo's degree of confidence that this is so. So if he was rationally sure at noon that if he read infinitely many feet he would be very likely to end up misremembering something, he should be sure at 1pm that this is so. So at 1pm, Apollo ought to be very highly confident that he has ended up misremembering something, and so ought to be highly confident that he did not read one of the conjuncts of *INF.* And it is also the case that Apollo ought to be very highly confident, of each of INF's conjuncts *c*, that unless he read that *c*, he should have low confidence that c. So he ought to be highly confident that he should have a low degree of confidence that *INF.* But then by AAR, if at 1pm Apollo ought to be highly confident that he should have a low degree of confidence that *INF,* and he does believe that *INF,* Apollo is

in an impermissible state. But then, by FORBIDDEN, he has done something he ought not to have done between noon and 1pm. But then by REQUIRED, however many feet he reads, he will do something impermissible: if he reads all of them, he will have done something he should not, and if he does not read all of them, he will have done something he should not. But then, given that the only thing Apollo could have done differently between noon and 1pm was read a different number of feet, Apollo cannot avoid acting impermissibly. And so Apollo finds himself in a tragic situation.

Let us return to this paper's central drama, ventriloquized using the figures of Socrates and Thrasymachus. Thrasymachus uses his slogan 'injustice is profitable' to articulate his scepticism that doing as one ought will leave one better off (or at least as well off) than doing otherwise. By looking at the Hades' pomegranate case, we saw that Thrasymachus' claim has something to it when it comes to prudential rationality: Persephone will be better off (on Mount Olympus) if she acts irrationally than she will be (in the underworld) if she acts exactly as she ought. And by looking at Knowledge Banking and Hades' Salmon, we saw that the same Thrasymachian structure is present in epistemic rationality too: acting just as one epistemically ought to can make one epistemically worse off than acting as one ought not. This sort of structure makes for misfortune, but not for tragedy: to be badly off is not

always to be in an impermissible state, and to be forced to choose between acting as one should, or being well off along some axis, is not to be forced to act impermissibly. But once we combine the insight that epistemic rationality has a broadly Thrasymachian structure with a commitment to anti-akratic constraints on rationality, we have the materials for genuine epistemic tragedy.

My project here has been to show, and to make vivid the claim, that there are tensions between anti-akratic and anti-tragic constraints. It has not been my project to say which of the constraints we should keep, and which we should do away with. But I am happy to say this much: insofar as anti-tragic constraints have considerable appeal, the tensions charted here should be considered a significant cost for fans of anti-akratic constraints.[127] In the next section I show that these tensions cannot be avoided by adopting more moderate anti-akratic views.

4.1. Moderate Anti-Akrasia.

[127] It would be interesting to see if other structural constraints on rationality are susceptible to similar arguments, and perhaps consoling for the anti-tragedian if they can find some partners in crime.

So far we have been talking in a rather coarse grained fashion about anti-akratic commitments. But as with many things, anti-akratic commitments come in degrees. Horowitz, for example, is broadly sympathetic to anti-akratic constraints on rationality, but acknowledges that there are some cases in which akratic states are permissible:

> DARTBOARD: You have a large, blank dartboard. When you throw a dart at the board, it can only land at grid points, which are spaced one inch apart along the horizontal and vertical axes. (It can only land at grid points because the dartboard is magnetic, and it's only magnetized at those points.) Although you are pretty good at picking out where the dart has landed, you are rationally highly confident that your discrimination is not perfect: in particular, you are confident that when you judge where the dart has landed, you might mistake its position for one of the points an inch away (i.e. directly above, below, to the left, or to the right). You are also confident that, wherever the dart lands, you will know that it has not landed at any point farther away than one of those four. You throw a dart, and it lands on a point somewhere close to the middle of the board (2014).[128]

[128] This case is closely modelled on Williamson's (2014) clock case. See also Williamson (2011).

Say the dart lands on <3,3>. If it landed on <3,3>, you can be sure that it landed on one of <3,3>, <3,2>, <3,4>, <2,3>, <4,3>. And given that you should be roughly just as confident that the dart landed on <3,3> as that it landed on any other one of the five points, it seems that it is rational for you to have high confidence that it landed on one of <3,2>, <3,4>, <2,3>, <4,3>. Call this proposition 'RING'.

But we should also recognise the following: that if the dart has landed on any of the four points around <3,3> and not on <3,3> itself, it would not be rational for me to have high confidence that RING. Suppose it landed on <3,2>. Then I can be sure that the dart hit one of <3,2>, <4,2>, <2,2>, <3,1>, or <3,3>. But if I ought to be sure of this, my degree of confidence that RING ought to be low. So if RING is true, then I ought to have a low degree of confidence in RING.

Now, Horowitz counts herself as a defender of anti- akratic constraints, but grants that cases like DARTBOARD generate genuine exceptions to them. In DARTBOARD, you ought to have high confidence that your evidence does not support RING (and so, presumably,

you should have high confidence that you ought to have low confidence that RING). But it does not seem irrational for you to have high confidence in RING. But Horowitz thinks that this case exhibits 'non-standard' features, and that in virtue of these non-standard features, anti-akratic constraints do not apply to it. Call Horowitz a 'moderate' anti-akratic. My project in this section of the paper is to show that being a moderate anti-akratic is not a way of avoiding the tensions between anti-akratic constraints and anti-tragic constraints. The features that Horowitz claims make for exceptions to anti-akratic constraints are not in play in Hades' Encyclopaedic Scroll.

Let us explore the first so-called 'non-standard' feature of DARTBOARD. In DARTBOARD, says Horowitz, you do not know what your evidence is. Your evidence in DARTBOARD is, roughly speaking, how things look to you, and your evidence is enough to make high confidence in RING rational. And we grant, in DARTBOARD, that the strongest claim you can have rational high confidence in is something pretty weak, like RING. It is also the case that by reflecting on your own epistemic capabilities, you can come to know that things will look a certain way only if the dart landed on <3,3>. Suppose that you are in a position to know (or to be rationally highly confident), when the dart lands on <3,3>, how things look to you. Then you could reason like this: things look just so. But

they look just so only if the dart landed on <3,3>. So the dart landed on <3,3>. And so it would not be the case that the strongest proposition you are in a position to have rationally high confidence in is RING. It follows that in DARTBOARD, you are not in a position to know how things look to you, and so, are not in a position to know what your evidence is. Call this feature of DARTBOARD its *opacity*.

The second 'non-standard' feature of DARTBOARD, claims Horowitz, is the presence of 'non-truth guiding evidence'. In standard cases, she thinks our evidence is truth-guiding:

> '[E]vidence usually points to the truth: when it justifies high confidence in a proposition, that proposition is usually true, and when it justifies low confidence in a proposition, that proposition is usually false.[129] If a detective's first-order evidence points to a particular suspect, that suspect is usually guilty. If it points away from a particular suspect, that suspect is usually innocent. In Dartboard, however, the evidence is *not* truth-guiding, at least with respect to propositions like Ring.

[129] This formulation might be read as suggesting temporalism about propositions. I suspect that it is not supposed to.

> Instead, it is *falsity*-guiding. It supports high confidence in Ring when Ring is false—that is, when the dart landed at <3,3>. And it supports low confidence in Ring when Ring is true—that is, when the dart landed at <3,2>, <2,3>, <4,3>,or<3,4>. This is an unusual feature of Dartboard. And it is only because of this unusual feature that epistemic akrasia seems rational in Dartboard. You should think that you should have low confidence in Ring precisely *because* you should think Ring is probably true—and because your evidence is falsity-guiding with respect to Ring' (2014).

In other words, in RING, in most of the cases in which things look the right sort of way for you to have rational high confidence in RING, RING is false, and in most cases in which RING is true, the way things look will not support high confidence in RING. Call this feature of DARTBOARD its *evidential inversion.*

There are two different ways we might argue that opacity and evidential inversion are not features of Hades' Encyclopaedic Scroll. One sort of argument involves specifying what Apollo's evidence is in Hades' Infinite Encyclopaedic Scroll, and then determining whether opacity or evidential inversion are features of the case. For someone who is thinking as

Horowitz seems to be thinking about evidence, the natural specification will be: *p* is part of Apollo's evidence so long as Apollo seems to remember reading *p.* But given this way of thinking about evidence, it does not look promising to suppose opacity or evidential inversion to be features of Hades' Infinite Encyclopaedic Scroll. There does not seem to be any particular reason to think that Apollo is not in a position to know which propositions he seems to remember reading. And given this way of thinking about evidence, the evidence Apollo gets from reading the scroll looks like a good candidate for being truth-guiding. For Apollo's memory is accurate in the vast majority of cases, so in most cases in which Apollo seems to remember reading *p, p* is true.. But a second sort of argument depends less vitally on our specifying a conception of evidence to work with. This second argument points out that the sort of evidence Apollo acquires in the infinite version of Hades' Encyclopaedic Scroll is the same as the sort of evidence Apollo gets in the finite version. So if Apollo is in a situation which features opacity or which features evidential inversion in the finitist version of the case, he is in a situation which features it in both the the finitist and the infinitist versions of the case. But insofar as someone sympathetic to Horowitz thinks that the finite version of the case does not generate an exception to anti-akratic constraints, they should think that opacity and evidential inversion are *not* features of the finite version of Hades' Encyclopaedic Scroll. And so someone sympathetic to Horowitz should think that the finite

version of Hades' Encyclopaedic Scroll does not have her non-standard features: surely if any akratic states are impermissible, the state Apollo ends up in if he reads more than 999 feet of scroll pages is surely one of them! But then, if the same sort of evidence is in play in the finite and the infinite versions of the case, neither opacity nor evidential inversion will be features of Hades Infinite Encyclopaedic Scroll.

In my concluding remarks, I will sketch some of the different ways one might respond to tensions between anti-akratic and anti-tragic constraints.

Conclusion.

Anti-akratic constraints articulate synchronic norms of rationality: they say something about how the attitudes of a particular agent's time-slice should hang together. Anti-tragic constraints – FORBIDDEN, in particular -- articulate diachronic claims: they say something about how the state of my present time-slice must be, given the states and the behaviour of my previous time-slices. That there turn out to be tensions between these two sorts of claim, then, will look to have implications for a dialectic that, so far, has been just out of our sightline.

Time-slice epistemologists defend the view that there are no 'essentially' diachronic norms of rationality (see Hedden (2015a, 2015b,) and Moss (2015) for motivation; Hlobil (2015) for critique).[130] Time-slice epistemologists think that what it is for one time-slice of an agent to be rational depends only on what is happening at that time-slice, and that the relations between different time-slices of a single agent are not importantly different, for the purposes of evaluation, from the relations between single time-slices of distinct agents. (As such, time-slice epistemologists reject conditionalization, which tells me that how my attitudes ought to be at one time depend on what *my* attitudes were earlier (see Hedden, 2015a).) Buying into the spirit of time-slice rationality may be one way to make becoming a tragedian look more palatable. If the relationships between my present time-slice and my past time-slices are not importantly different, as far as evaluating the permissibility of my present state goes, than the relationships between my present time-slices and your present time-slices, then the sort of essentially diachronic claim being articulated by something like FORBIDDEN will start to look misguided.[131] Anyone who wants to hang on to AAR,

[130] See also Christensen (2004), Foley (1993) and Lam (m.s.)

[131] Hedden (2015b) argues against the general strategy of deeming attitudes irrational because they lead to cases of 'diachronic tragedy', which occurs when one performs a sequence of actions S1 such that at all times you prefer performing some other possible sequence of actions S2 over performing S1. Being trapped in a Sisyphean loop looks a lot like a diachronic tragedy.

then, should consider themselves as having some motivation to endorse the time-slice epistemologist's picture.

Analytic epistemology is sometimes criticised -- both from within and without -- for its tendency to 'idealise', and to model the epistemic lives of agents in abstraction from their cognitive shortcomings and social positions. For example, subjective Bayesianism has been 'widely criticised for [postulating] norms on degrees of belief… [that] cannot be followed by human agents, and hence have no normative force for beings like us.' (Staffel, 2015). And Fricker (1998) observes: 'the tradition [of analytic epistemology] provides us with a clinically asocial conception of the knowing subject, with the result that epistemology tends to proceed as if socio-political considerations were utterly irrelevant to it.' If one is sympathetic to these critiques, one might be wondering why beings like us – beings incapable of the sorts of supertasks exploited in these paper's core examples – ought to care much about the epistemic predicaments of creatures quite unlike ourselves. If you are implacably of the opinion that theorising about cognitively idealised agents is of no real epistemic interest -- what we want, after all, out of epistemology, is a theory that applies to beings like us, living

.

the sorts of lives we lead -- I will not be able to persuade you here to change your proclivities. But there might be a few points of interest, even for those who are so inclined, about the relationship between 'ideal' and 'non-ideal' epistemology. Exactly what these points of interest are will depend on whether you think anti-akratic or anti-tragic constraints will be the one to give.

Suppose you go for the first option. This involves rejecting anti-akratic constraints for beings like Apollo. Either you can keep them for creatures like us, or have different norms of rationality for us and for Apollo.

It is often assumed that if different norms apply to cognitively limited beings like us than apply to beings with fewer cognitive limitations, then the norms that apply to the latter will be more demanding than the norms that apply to the former. We will be incapable of acting in accordance with norms that more epistemically perfect beings can act in accordance with, and so, norms will apply to *them* that do not apply to us. The exigency of epistemic normativity will correlate with degree of cognitive perfection, and the norms posited by 'ideal epistemology' will need to be slackened if they are to have any binding force or normative bite for creatures like us. This neat moral will be disrupted if we think the right

conclusion to draw from Hades' Encyclopaedic Scroll is that beings like Apollo are sometimes permitted to occupy akratic states, but that beings like us never are. The anti-akratic constraints that apply to us will be tighter than the anti-akratic constraints that apply to Olympian Gods.

It is hard, though, to see that keeping anti-akratic constraints for us and getting rid of them for Apollo is particularly satisfactory. Presumably, a large part of the reason we want anti-akratic constraints, for ourselves in the first place, is that akratic states, in Horowitz's terms, license bad reasoning and irrational action. But it seems that this sort of argument should apply just as well to Apollo as to ourselves. If we can acknowledge the force of those arguments for Apollo and yet refrain from imposing an anti-akratic constraint on him, why can we not do the same for ourselves? If Apollo stands in a different relationship to anti-akratic constraints than we do, it would be nice to have a story that explains why there is this difference, rather than one which merely frames the difference as a brute consequence of Apollo's greater susceptibility to epistemic exploitation.

Suppose you go for the second option, and think that anti-tragic constraints should be the ones to go. This rejection comes with another decision point: reject anti-tragic constraints

tout court, or reject them for Apollonians and keep them for humans. If you want to keep anti-tragic constraints for humans, part of your motivation is probably that you take seriously the thought that rationality's anti-tragic structure is an important part of what makes rationality action guiding. And, at a first glance, it looks rather natural to think that which norms are action guiding for a creature depends on their cognitive capacities. The trouble is that it is very natural to think that for any given set of norms, and whether or not those norms are action guiding for a creature, will depend on the epistemic and cognitive capacities of that creature. But it is also very natural to think that where a norm fails to be action guiding for me, but does not fail to be action guiding for you, this is because you are more epistemically or cognitively astute than me. Take the norm: don't step on ants. This norm might be action guiding for you but not for me because you are better than me at not stepping on ants. But the case of anti-tragic constraint is not like this. Instead, rationality can play an action guiding role for us *because*, rather than in spite of, our cognitive limitations. Our cognitive limitations come with perks: they make us invulnerable to certain kinds of epistemic tragedy that less limited creatures are vulnerable to. That is, the discontinuities between ideal and non-ideal epistemology that Hades' Encyclopaedic Scroll reveals, present in surprising ways, even if the fact that there are discontinuities in the first place is not surprising.

Acknowledgements.

Thanks to Matt Clark, Lizzie Fricker, David Glick, John Hawthorne, Annina Loets, and Martin Pickup for discussion and comments on earlier drafts of this paper. Special thanks go to Beau Madison Mount.

References.

Arntzenius, F., Elga, A., & Hawthorne, J., 2004. Bayesianism, Infinite Decisions, and Binding. *Mind* 113 (450): 251 -- 283.

Bacharach, M., 2006. Beyond Individual Choice: Teams and Frames in Game Theory, ed. Gold, *n.*& Sudgeon, R.. Princeton: Princeton University Press.

Broome, J., 2007. Wide or narrow scope? *Mind* 116 (462): 359–370.

Brown, B., 1999. Adjunction and Aggregation. *Noûs* 33 (2): 273-283.

Buchak, L., 2013. Risk and Rationality. Oxford: Oxford University Press.

- 2014. Belief, Credence, and Norms. *Philosophical Studies* 169 (2): 285-311.

Christensen, D., 2004. *Putting Logic in its Place.* Oxford: Clarendon Press.

Dennett, D. C., 1996. *The Intentional Stance.* London: MIT Press.

Fricker, M., 1998. Rational Authority and Social Power: Towards a Truly Social Epistemology. *Proceedings of the Aristotelian Society, sup. vol.* 98 (2): 159-177.

Greco, D., 2014. A Puzzle About Epistemic Akrasia. *Philosophical Studies* 167 (2): 201-219.

Hakli, R., Miller, K., & Tuomela, R., 2010. Two Kinds of We-Reasoning. *Economics and Philosophy* 26: 291-320.

Hedden, B., 2015a. Time-slice Rationality. *Mind* 124 (494): 449-491.

- 2015b. Options and Diachronic Tragedy. *Philosophy and Phenomenological Research* 90 (2): 423-451.

Hlobil, U., 2015. There Are Diachronic Norms of Rationality. *Thought* 4 (1): 38-45.

Horowitz, S., 2014. Epistemic Akrasia. *Noûs* 48 (4):718-744.

Kelly, T., 2003. Epistemic Rationality as Instrumental Rationality: A Critique. *Philosophy and Phenomenological Research* 66 (3): 612–640.

– 2014. Evidence Can Be Permissive. In eds. Steup, M., Turri, J. & and Sosa, E.. *Contemporary Debates in Epistemology.* Wiley-Blackwell: 298-311.

Kratzer, A., 1977. What must and can must and can mean. *Linguistics and Philosophy* 1(3). 337–355. http://dx.doi.org/10.1007/BF00353453.

– 2012. *Modals and Conditionals*. Oxford: Oxford University Press.

Kyburg, H. E., 1961. Probability and Rationality. *Philosophical Quarterly* 11 (44):193-200.

Lam, B., m.s.. The Dynamic Foundations of Epistemic Rationality. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.592.9508&rep=rep1&type=pdf.

Littlejohn, C., 2015. Stop Making Sense: A Puzzle About Rationality. *Philosophy and Phenomenological Research.* doi: 10.1111/phpr.12271.

Makinson, D., 1965. Paradox of the Preface. *Analysis* 25: 205-207

Moss, S., 2015. Time-slice epistemology and action under indeterminacy. *Oxford Studies in Epistemology* 5. http://dx.doi.org/10.1093/acprof:oso/9780198722762.003.0006.

Nussabaum, M., 1986. The Fragility of Goodness. Cambridge: Cambridge University Press.

Plato, 1981. *The Republic,* trans. Grube, G.M.A.. London: Pan Books.

Pritchard, D., 2014. Truth as The Fundamental Epistemic Good. In eds. Matheson, J. & Vitz, R., *The Ethics of Belief : Individual and Social*. Oxford: Oxford University Press, pp. 113-127.

– 2015. *Epistemic Angst.* Oxford: Oxford University Press.

Redmayne, M., 2003. Objective Probability and The Assessment of Evidence. *Law, Probability and Risk* 2: 275–294.

Rinard, S., m.s.. 'Reasoning One's Way Out of Scepticism' http://philpapers.org/archive/RINROW.pdf..

Ryan, S., 1996. The Epistemic Virtues of Consistency. *Synthese, 109* (2): 121-141.

Scanlon,T. M., 1998. *What We Owe To Each Other*. Cambridge, MA: Harvard University Press.

Srinivasan, A., 2015. Normativity Without Cartesian Privilege. *Philosophical Issues* 25 (1): 273–299.

- 2016. The Archimedean Urge. *Philosophical Perspectives* 29 (1): 325-362.

Staffel, J., 2015. Measuring the Overall Incoherence of Credence Functions. *Synthese* 192 (5): 1467-1493.

Titelbaum, M., 2015a. Rationality's Fixed Point. O*xford Studies in Epistemology* 5*:* 253–294.

- 2015b. How to derive a narrow scope requirement from wide scope requirements. *Philosophical Studies* 172 (2): 535–542.

Treanor, N., 2014. Trivial Truths and the Aim of Inquiry. *Philosophy and Phenomenological Research* 89(3): 552-559.

von Fintel, K., 2012. The Best We Can Expect to Get. Unpublished paper for a session on Deontic Modals at the Central APA, February 17, 2012.

Williams, B., 1994. Jim and The Indians. In ed. Singer, P. Ethics. Oxford: Oxford University Press.

Williamson, T., 2011. Improbable Knowing. In ed. Dougherty, T. *Evidentialism and its Discontents*. Oxford: Oxford University Press.

- 2014. Very Improbable Knowing. *Erkenntnis* 79 (5): 971-999.