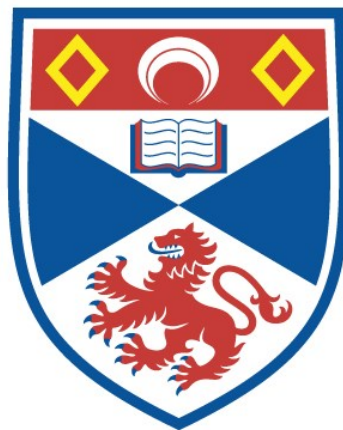


The epistemology of inquiry: individuals, groups & institutions

Lewis Dylan Ross

A thesis submitted for the degree of PhD
at the
University of St Andrews



2020

Full metadata for this thesis is available in
St Andrews Research Repository
at:

<https://research-repository.st-andrews.ac.uk/>

Identifier to use to cite or link to this thesis:

DOI: <https://doi.org/10.17630/sta/790>

This item is protected by original copyright

Candidate's declaration

I, Lewis Dylan Ross, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 59,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree.

I was admitted as a research student at the University of St Andrews in September 2015.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date 20/06/2019 Signature of candidate

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

20/06/2019

Date Signature of supervisor

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Lewis Dylan Ross, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

No embargo on print copy.

Electronic copy

No embargo on electronic copy.

Date

20/6/2019

Signature of candidate

Date

20/6/2019

Signature of supervisor

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Lewis Dylan Ross, hereby certify that no requirements to deposit original research data or digital outputs apply to this thesis and that, where appropriate, secondary data used have been referenced in the full text of my thesis.

Date 20/6/2019

Signature of candidate

The Epistemology of Inquiry:
Individuals, Groups & Institutions

Lewis D. Ross

Foreword

First and foremost: I'd like to register my appreciation for the good times to everybody I met as a research student. Although academic life certainly has its up and downs, the community in the small seaside town of St Andrews has been a great source of friendship, humour, absurdity, and enlightening conversations. I enjoyed my time thoroughly: so, to you all—cheers!

Secondly, I'd like to thank those who generously gave up their time and energy to talk philosophy with me. I am especially grateful to my primary supervisor, Jessica Brown, both for her patient and detailed comments on every aspect of this thesis and for her sensible and pragmatic advice. Many other people—particularly my friends in the Arché Research Centre—helped me think better over the years. Others simply led by example, acting as models of rigour and clarity. Collectively, I owe you all a debt.

Thirdly, I incurred various other debts (emotional, financial, and practical) during my PhD. The bulk of these are to my family and friends, while others are to itinerants who only passed through. I'm not sure how many of you will suppose that a thesis in philosophy is much recompense! I hope to repay these debts one way or another, even if only by passing on the favour.

Finally, I'd like to give special thanks to Tuuli and Sara, for putting up with me, and for years of support.

The ideas in this thesis have also partly been shaped by audiences and discussions at various institutions, including: Sheffield; Rutgers; Edinburgh; Pavia; National University of Ireland; Universidad Diego Portales; Cardiff; Rutgers; Graz; Aarhus; and in many conferences and workshops at St Andrews. In addition, various anonymous referees provided extremely—often painfully—instructive criticism of different manuscripts, helping me to sharpen my research considerably.

Abstract

Inquiring—roughly, the attempt to answer a question—is one of the most common intellectual undertakings. This thesis is comprised of a set of four essays that investigate the epistemology of inquiry. The first essay looks at one key epistemic state with which we end inquiry, viz understanding; the second essay examines collective inquiry; the third essay considers the normative role of curiosity in motivating inquiry; and the fourth essay discusses inquiry in the legal system. These essays will advance our understanding of what motivates inquiry, the mental states involved in inquiring, the sorts of norms that govern inquiry, and what continuities and discontinuities there are between how individuals and various types of collective inquire into different types of question.

Journal Acknowledgements

Thanks to the editors of the following journals for allowing me to reproduce published material in my thesis.

The paper ‘Is Understanding Reducible?’ in *Inquiry* contains much of the material in Essay One. <https://doi.org/10.1080/0020174X.2018.1562379>

The paper ‘How Intellectual Communities Progress’ forthcoming in *Episteme* contains much of the material in Essay Two.

The paper ‘The Virtue of Curiosity’ in *Episteme* contains some of the material in Essay Three. <https://doi.org/10.1017/epi.2018.31>

The paper ‘Rehabilitating Statistical Evidence’ forthcoming in *Philosophy and Phenomenological Research* contains much of the material in Essay Four.

Funder Acknowledgement

This thesis was completed with the help of a scholarship from the Carnegie Trust for the Universities of Scotland.

Table of Contents

PHILOSOPHICAL CONTEXT & INTRODUCTION	5
ESSAY ONE: At the End of Inquiry	10
ESSAY TWO: Collective Inquiry & Intellectual Progress	40
ESSAY THREE: The Virtue of Curiosity	66
ESSAY FOUR: Legal Inquiry & Statistical Evidence	94
CONCLUSION	119
BIBLIOGRAPHY	121

Introduction

This thesis is comprised of four essays that advance our understanding of the epistemology of inquiry. Although these essays stand independently of each other, they each explore different aspects of one of the most common intellectual undertakings: the attempt to answer a question. The first essay looks at one key epistemic state with which we end inquiry, viz understanding; the second essay examines collective inquiry; the third essay considers the normative role of curiosity in motivating inquiry; and the fourth essay discusses inquiry in the legal system. By the end, we will have a better idea of what motivates inquiry, the mental states involved in inquiring, the sorts of norms that govern inquiry, and what continuities and discontinuities there are between how individuals and various types of collective inquire. In this section, I provide some philosophical context for my thesis and then briefly introduce each essay.

Much analytic epistemology from the mid-20th century onwards was engaged in the project of investigating what is required in order to possess certain epistemic states: paradigmatically knowledge, but also warrant, justification, and latterly, evidence. This research programme was largely synchronic in nature, looking at particular time-slices of agents in an attempt to uncover what conditions must be met in order to for them possess these epistemic states. However, it is also possible to approach epistemic theorising from a more diachronic perspective by focusing on the extended process of inquiring into a question. This approach finds a place for analysing familiar epistemic states such as knowledge—e.g. by supposing them to be a terminus of inquiry—but adopts a broader outlook by asking about our motivations for seeking knowledge, about the mental states involved in moving from ignorance to knowledge, and considering how we should evaluate agents at different stages of inquiry.

This broader focus on inquiry coheres naturally with a number of other concerns that have become increasingly central to epistemology. One is the growth of work on *epistemic value*. In the last twenty years or so a healthy contemporary literature has emerged that goes beyond asking about the definition or analysis of epistemic states, by instead investigating why these states are worth having, and what distinctive value different epistemic states possess.¹ A focus on inquiry complements this literature because those providing an account of the value of different epistemic states will do well to explore how these states satisfy the different motivations we have for engaging in inquiry. A second concern that has been of growing interest is the relationship between *credences and beliefs*. While work on credence was once pursued largely in formal contexts without much overlap with less formal questions in epistemology, a growing body of literature now tries to clarify exactly how credences and beliefs are related.² A focus on inquiry is well suited to add to this research programme, because as we will see

¹ For instance see Haddock, Millar, and Pritchard (2009) for a representative collection.

² For example see Greco (2015) and Easwaran (2015) for conflicting views; Jackson (*forthcoming*) sets out the relevance of this debate for a range of other questions in epistemology.

below it has been claimed that beliefs play a distinctive ‘inquiry-settling’ role that credences do not. A final motivation is the rise of *social epistemology*. A great deal of fruitful work now asks how we can epistemically evaluate different collectives, whether and when we can ascribe epistemic states to groups, and how our positions on these questions relate to individual-level facts about a group’s members. Focusing on inquiry is a natural extension of this project, because it seems that groups can also be aptly described as engaging in inquiry—and indeed, there are various groups (such as courts and academic communities) that exist for the very purpose of answering different types of questions. Investigating the similarities and differences between individual and collective inquiry will help us better understand the epistemic lives of groups.

Given that the nature of inquiry and the norms which govern it will be central to the essays comprising this thesis, it is worth here setting out the assumptions about inquiry that I will maintain throughout. I set out those assumptions in this rest of the introduction before sketching the content of the subsequent parts of the thesis.

To begin theorising about inquiry, a good place to begin is with work of Jane Friedman who—in a pioneering series of papers—has provided an influential account of the nature of inquiry in addition to considering how it relates to other epistemological issues such as the nature of credence and belief.³ Central to Friedman’s view are what she calls the ‘interrogative attitudes’—that is, attitudes that take some question *Q* as their contents. Her suggestion that there are interrogative attitudes aiming at questions can be helpfully compared with the more familiar idea of propositional attitudes taking propositions as their contents. So, for example, we can say that wondering is always directed at some question *Q* in the same way that we typically suppose that belief is always directed at some proposition *p*. Possessing such attitudes, and hence having propositions or questions as the contents of one’s attitudes, does not require that you possess the concept of a proposition or a question.

There are a number of interrogative attitudes including familiar folk-psychological attitudes or states such as: wondering, curiosity, contemplation, deliberation, and suspension of belief. For Friedman, these interrogative attitudes—or ‘IAs’ for short—play a central role in inquiry; on her account, “someone inquiring at *t* has an IA at *t*”.⁴ She claims that this attitudinal component is essential to genuinely engage in inquiry. To see this, consider that we cannot say that someone is genuinely inquiring just in virtue of them performing certain actions. For instance, I might perform all of the same bodily actions as a detective (searching for clues, asking suspects about their alibi, writing out the evidence) and draw inferences accordingly—but, if I already know the answer to the *Q* ‘who committed the crime’ and am simply going through the motions in order to impress my supervisor, then I do not

³ See Friedman (2013a; 2013b; 2017; 2019).

⁴ Friedman (2019: 299).

count as genuinely inquiring into that question. So, having some IA, plausibly, is a necessary condition for engaging in inquiry.

However something we ought to note, that Friedman only gestures at briefly, is that having an IA is certainly not always *sufficient* to count as inquiring. In this vein, for example, we should acknowledge that being curious about a question does not mean that you are thereby inquiring into that question. After all, you can be curious about some Q even though you have no intention of finding out the answer to Q: for example, inquiring into Q might be very dangerous (e.g. <is Lorenzo a member of the mafia?>), or you might know that any inquiry would be fruitless because Q is simply impossible to settle (e.g. <how many different types of fowl did Henry VIII's chief courtier eat on December 9th, 1512?>). Hence, curiosity and inquiring are distinct states even though they clearly have an intimate connection. Here, I don't want to take a position on whether this applies to *all* of the interrogative attitudes. For instance, it seems much more plausible to suppose that contemplating Q entails inquiring into Q than it would to suppose that a similar entailment claim holds regarding curiosity and inquiry. Nonetheless, to count as inquiring, I think, one must *actively* be trying to answer some question Q at least in the sense of having embarked upon—and be committed to resuming—epistemic activity that will lead you to answer Q.⁵ At minimum, inquiring involves a type of commitment to directing one's cognitive effort towards considering a given question Q with a view to answering it. This explains why, for instance, mere curiosity does not entail inquiring; one can be curious about a question without any commitment to seeking an answer.

For Friedman, inquiry has a specific terminus: the acquisition of an *outright belief* that answers the question inquiry is directed at. The relevant contrast here is with credence or degrees of belief, and outright belief is supposed to encompass knowledge and other epistemic statuses that we can attribute to an outright belief.⁶ Friedman suggests that settling inquiry is one distinctive role that belief plays in our mental economy, writing that: “Belief is a settling attitude, it's a way of closing inquiry and more generally it's normatively incompatible with the IAs and inquiring; mere high credence lacks these properties.”⁷

This is where we should part ways with Friedman's analysis of inquiry. Although it may be generally true that inquiry ends with the acquisition of new belief, this is not always the case. The exception is

⁵ This gloss is intended to capture cases in which inquiry is extended over a long period of time. For example, I might be inquiring into whether it would for the best, all things considered, to move house. Making up my mind and gathering evidence might take place over many months. Nonetheless, it seems reasonable to suppose that this is all part of a single process of inquiry—rather than supposing that I inquire for a while, stop inquiring, start inquiring again, etc. However, for the purposes of this thesis, not much will turn on whether you share this judgement.

⁶ Although it is worth noting that, recently, Moss (2016) has argued for the view that credences can constitute knowledge.

⁷ Friedman (2019: 309).

found in scenarios in which inquiry ends when we merely increase our confidence in some proposition. Such cases come in two types: when inquiry ends as a result of increasing one's confidence beyond that required to hold an outright belief in a given proposition, and when inquiry ends as a result of becoming more confident without putting one in a sufficiently strong epistemic position to hold an outright belief.⁸

On plausible fallible epistemological assumptions, an agent can believe something without credence 1. That is, it is possible to believe and indeed know a proposition without being entirely certain about that proposition. With this in mind, we might expect that inquiry into some question *Q* (e.g. <whether *p*?>) can continue even if one already believes *p*. This might occur just because one wants to become even more confident that *p* answers *Q*—to move closer to certainty about *p*—despite already believing (or, indeed, knowing) *p*. Examples are easy to come by: you might believe that the restaurant is open, but want to make sure before making the journey across town, so you decide to re-open inquiry and gather further evidence by phoning to check. We can bolster this line of thought by pointing out that having an interrogative attitude towards *Q* is compatible with having a belief in the answer to *Q*—consider the following type of ascription: “I’m wondering whether she ever received my letter—I believe she did, but I’m not entirely sure.”⁹ This ascription suggests that settling inquiry can require becoming more confident in some proposition one already believes, and that continued inquiry is compatible with any attitude towards the answer to a given question short of subjective certainty.

In the other direction, one might envisage cases where inquiry ends by increasing one's confidence in *p* even without coming to believe *p* outright. This could occur in cases in which it would require a lot of evidence to bring yourself to believe *p* outright and you aren't sufficiently motivated to put in the required effort, or where doing so would be unreasonably difficult. For instance, you might wonder whether it is prudent, all things considered, to eat non-organic produce. This is a difficult question, and the evidence relevant to answering it is complex and difficult to evaluate. It is plausible, if you are not willing to devote an extended period of time to the complex task of weighing the arguments and empirical results in each direction, that you will end inquiry simply by becoming more confident that it is or is not safe—all without forming an outright belief either way. Indeed, not only is this possible, but it is arguably the normatively appropriate response. If you are committed to not considering a body of evidence comprehensively or thoroughly, it will be better to reach a qualified probabilistic belief ('probably *p*') rather than fully commit yourself one way or the other ('*p*') or (' $\neg p$ ').

⁸ A tempting objection: whenever when we become more confident and settle inquiry, won't this be in virtue of gaining a new belief—for example, a belief about a new piece of evidence? No. One might come to have a greater confidence whether *p*—thus settling your inquiry—by re-evaluating old evidence and ending up with a higher credence in *p*. For example, you might give less weight to far-fetched possibilities that count against *p*.

⁹ See recent work by Archer (2018) for a more detailed discussion of such linguistic data and associated issues.

With these remarks about inquiry in mind, I can now briefly introduce the essays comprising this thesis.

At the End of Inquiry. Both knowledge and understanding have been identified as the proper end of inquiry. The first section of my thesis attempts to clarify how these states relate. Focusing on the recently popular view that understanding can be reduced to what an agent knows, I develop a novel framework for testing such reductionist views, and then critique the idea that understanding reduces to knowledge. On this basis, I sketch an alternative account of understanding that focuses simply on the factive component of knowledge, rather than justification or anti-luck components. In so doing, I aim to show that focusing just on what an agent truly believes is both a theoretically well-motivated and robust approach to the epistemology of understanding.

Collective Inquiry & Intellectual Progress. There has recently been increasing interest in the epistemic lives of groups. The second section of my thesis considers the nature of collective inquiry. I sketch an account of collective inquiry, and aim to show why it is fruitful to adopt the theoretical perspective of collective inquiry. In particular, I apply this framework to a recently pressing question: how (and whether) intellectual communities—such as the scientific and philosophical community—make progress in their collective inquiries. Focusing on collective inquiry, so I claim, provides a more nuanced account of progress than simply looking at the presence or absence of epistemic states such as knowledge.

The Virtue of Curiosity. In recent years, there has been some scattered work in philosophy and psychology on curiosity. The third section of my thesis attempts to show that in addition to being one of the interrogative attitudes associated with inquiry, curiosity also has an important normative role to play in epistemology. After separating out a variety of different issues relating to curiosity and how they overlap with extant work, I develop a novel account of curiosity as an epistemic virtue and show that it is a central component of an intellectually virtuous character. In so doing, I also illustrate why considering this virtue sheds new light on broader issues within virtue epistemology.

Legal Inquiry & Statistical Evidence. The last decade has seen a spate of work at the intersection between epistemology and jurisprudence, with philosophers attempting to use familiar epistemic theories to resolve difficult legal questions. I argue that there are some crucial differences between the aims of legal inquiry and individual inquiry. Specifically, I show that individuals and courts are in different positions with respect to the possibility of adopted probabilistic beliefs in response to the questions they inquire into. On this basis, we should doubt whether concepts and arguments from individualistic epistemology neatly transpose to the philosophy of law. To exemplify this, I focus on a vexing issue regarding the role of statistical evidence in legal inquiry to show that legal inquiry is *sui generis*. An upshot of this argument is that recent critiques of statistical evidence in the law are, to some degree, unwarranted.

Essay One

At the End of Inquiry

0. Introduction

A prominent theme in epistemic theorising since around the start of the millennium has been the (re)discovery of the epistemic state of understanding. Although notions sometimes translated as ‘understanding’ appeared in ancient philosophy, only in the last few decades have epistemologists attempted to use the tools of analytic philosophy to precisely distinguish the nature of understanding from other epistemic states such as knowledge and belief. There have been two key (and related) motivations for this resurgent focus on understanding. Firstly, it has been suggested that understanding is somehow importantly different from knowledge (e.g. Zagzebski 2001; Kvanvig 2003; Elgin 2007). Secondly, it has been claimed that understanding has some distinctive evaluative importance in epistemology—such as being of particular epistemic value, or being a fitting state with which to end inquiry. One thought that unified both of these motivations is the idea that understanding necessarily involves a particularly rich grasp of a body of explanatory facts or relations, while knowledge can be atomistic and even superficial in the sense of merely concerning single propositions. In this vein, Kvanvig tells us that while “[a] head full of trivia and detail is an amazing thing, [it is] nothing to be compared with the reach and sweep of a person of understanding, so if knowledge is a good thing, understanding is even better.”¹⁰

In tandem to this work in epistemology, understanding has also come to occupy an increasingly important role in other areas of philosophy. Philosophers of science have been interested in the relationship between explanation and understanding for some time now, and the tendency to elicit understanding is often recognised as an important virtue of a scientific theory or model.¹¹ Further afield, moral philosophers have also appealed to understanding as a way to elucidate certain normative intuitions. For example, it has been suggested that the possession of understanding (rather than mere knowledge) explains when we are morally praiseworthy for acting in certain ways, and it has been claimed the lack of understanding might provide an explanation for cases in which we should not defer to the testimony of others concerning matters of right and wrong.¹²

Despite this proliferation of work, there is still a striking lack of consensus on the nature of understanding. In the search for a satisfactory account, the most prominent debate has been how exactly understanding relates to knowing. As I mentioned already, the initial resurgence in theorising about understanding was motivated by apparent *divergences* between the conditions under which we attribute understanding and knowledge. However, more recently, there has been a spate of work motivating *knowledge based* theories of understanding—accounts that seek to explain the nature of understanding by appealing to facts about what an agent knows (particularly Kelp 2014, 2015, 2017; Sliwa 2015,

¹⁰ Kvanvig (2013: 186).

¹¹ For instance, see Strevens (2013) and de Regt (2017) for recent contributions.

¹² For discussion, see Hills (2010; 2013; *forthcoming*).

2017; Riaz 2015; and see Lawler 2016 for discussion). Given that proponents of knowledge based accounts deny the supposed differences between understanding and knowledge, they are undercutting one of the primary reasons that epistemologists started to become interested in understanding in the first place. Indeed, if understanding does not diverge from knowledge in the way that was suggested, this may have interesting ramifications for the further theoretical uses to which understanding has been put—for instance, in debates on testimony and moral worth. Therefore, we will do well to get clear on the plausibility of these knowledge based accounts of understanding.

In this essay I will advance this debate by arguing that while knowledge based accounts of understanding are incorrect, the basic insight behind these accounts provides a very promising way to think about the nature of understanding. Specifically, I will take a critical look at an attractive framework for theorising about understanding endorsed by many proponents of knowledge based accounts—namely ‘reductionist’ frameworks that aim to reduce understanding to other more familiar epistemic states. After situating this framework in relation to other prominent views of understanding (§1), I will develop a natural way to test such reductionist theories (§2), before using this method to outline a counterexample to a recently influential view on which understanding reduces to a body of knowledge (§3). Then, I will outline my own preferred form of reductionism on which understanding reduces to a body of true beliefs (§4). After suggesting that this theory can account for the gradability of understanding (§5), I close by defending this theory against one of the most powerful objections that it faces—namely, the suggestion that the role of falsehoods in the success of scientific theorising make it implausible to suppose that understanding is a factive notion (§6).

1. Knowledge, Understanding & Reductionism

‘Understanding’ features in various philosophically interesting locutions. Most recent work focuses on understanding-*why* (e.g. Jones understands why the war was lost) and *objectual* understanding (e.g. Juanita understands Polynesian culture).¹³ I will focus primarily on *understanding why*.¹⁴ I will not much be concerned with ‘linguistic understanding’, such as when one understands the meaning of a word or sentence.

The relationship between understanding and knowing has been a central topic in theorising about understanding. The earliest knowledge based theories of understanding-why, theories that aim to

¹³ Grimm (2012) summarises early work on understanding. On understanding-why see, inter alia, Sliwa (2015; 2017), Hills (2015), Pritchard (2010), Grimm (2006). On objectual understanding see, inter alia, Kvanvig (2003), Kelp (2017), Elgin (2007), Zagzebski (2001).

¹⁴ See Kvanvig (2003) and Khalifa (2013) for discussion of whether objectual understanding can be reduced to understanding why something is the case. Later on in §6 when I discuss Elgin’s non-factive theory, I will look at examples of objectual understanding. In so doing, I hope to show that my approach deals with both understanding why and objectual understanding in a satisfactory manner but without taking a definitive stance on which is the more fundamental notion.

account for understanding in terms of knowledge, defended the following view: *S understands-why p iff S knows-why p*.¹⁵ We might call these views ‘equivalence’ theories, because they claimed that understanding is simply equivalent to some item of knowledge. This position proved unable to explain the *gradability* of understanding—the fact that, unlike an orthodox conception of knowledge, understanding comes in degrees.^{16,17} We can readily conceive of two agents who both share knowledge why *p* yet are in asymmetrical epistemic positions regarding *p*. A well-known example adapted from Pritchard (2010) exemplifies. Your house burns down due to faulty wiring: *S* is informed of this proposition through extremely reliable testimony, while *S** deduces it from a raft of other beliefs after investigating the scene and applying their experience as an electrician. Plausibly, both know why the house burned down—on standard account, by knowing *that* it burned down due to faulty wiring—yet *S** possesses greater understanding. Equating knowing and understanding cannot explain such epistemic asymmetries.¹⁸

Recent work has outlined a more sophisticated knowledge based approach, utilising the idea that understanding can be *reductively explained* in terms of knowledge.¹⁹ This position offers resources with which to explain the gradability of understanding. To use an analogy due to Sliwa (2017): just as reductive physicalists about the mental can say that a subject experiences a mental state in virtue of a *complex* underlying physical state, the reductionist about understanding can attempt to explain the extent of an agent’s understanding by appealing to *a complex body of relevant and interconnected knowledge*. Such views permit the judgement that two agents can both know why *p* while understanding why *p* to different degrees—epistemic asymmetries can be explained by pointing at differences in other things the agents know. So, while two agents might know a house burned down due to faulty wiring, asymmetries in understanding can be explained by showing that the agent with greater understanding has more knowledge about (for example) how wires can be faulty, why houses are combustible, etc. In short, explaining the gradability of understanding is an essential requirement for any plausible theory and theories that reduce understanding to a body of knowledge have good claim to fulfil this requirement.

¹⁵ E.g. Sliwa (2015), Riaz (2015).

¹⁶ This is pressed effectively by Lawler (2016).

¹⁷ It also prompted debate on the relationship between understanding and epistemic luck. See Pritchard (2010), Hills (2015), Sliwa (2017: 526-527) and Kelp (2017: 261-3).

¹⁸ Grimm (2006; 2014) is also sympathetic to a sort of equivalence view, but places emphasis on knowledge of a special kind: knowledge of the “modal relatedness of the terms of the causal relata” (i.e. knowing that *p* causes *q* in such a way that one grasps the modal relationship between *p* and *q*). I won’t discuss further specifics of Grimm’s view here, but it seems like the same issue arises—two agents might possess such knowledge but be in asymmetrical epistemic positions. In response, it would be open for Grimm to appeal to a reductionist theory as I outline below.

¹⁹ Most prominently, Sliwa (2015; 2017) and Kelp (2014; 2015; 2017).

Although we have introduced reductionism by considering a knowledge based version, the reductionist approach offers a more general framework that one might use to theorise about understanding without focusing on knowledge. Indeed, with varying levels of plausibility, one could propose any mental or epistemic state as the reductive base of understanding. Let ϕ be a placeholder for some state, and we get the following very general gloss on reductionism:

Reductionism: ϕ is all there is to understanding—the extent to which an agent understands why p can be explained wholly in virtue of ϕ they possess.²⁰

In addition to the knowledge based view, there are a variety of other possible presifications of **Reductionism** that might have some plausibility. For instance, one might defend a theory on which understanding reduces just to a body of justified true beliefs; just to a body of true beliefs; just to a body of beliefs that have a sufficient degree of verisimilitude; and so forth.²¹ These are all *univocal* presifications of **Reductionism**—they reduce understanding to only one type of epistemic state. Of course, on pain of implausibility, this is not to say that someone who defends a justified true belief view would claim that *only* justified true beliefs rather than items of knowledge can constitute one’s understanding. It simply means that, to the extent that some piece of knowledge *does* figure in one’s understanding, this is in virtue of the fact that knowing that p entails justifiably and truly believing that p . Furthermore, it is also perfectly consistent for there to be a *pluralist* presification of **Reductionism** on which understanding reduces to more than one type of epistemic state, even though it is harder to immediately see the motivation for such a view.

Reductionism offers a different theoretical starting-point to two other approaches to theorising about nature of understanding that have predominated in the literature, namely views focusing on grasping and views focusing on abilities. However, as I discuss after presenting these views and some apparent difficulties facing them, reductionist views are in fact compatible with each of these approaches in principle, even though it is not committed to endorsing either of them.

Compare **Reductionism** with:

²⁰ Kelp (2017: 252) points out the need for a ‘well-connectedness’ proviso in reductionist theories. For instance, an agent might have considerable knowledge but be unable to connect or draw upon it simultaneously, to the detriment of their understanding. I take this proviso as implied.

²¹ Beyond the knowledge based variant, which has been explicitly defended, it is less easy to confidently attribute other **Reductionist** views to particular theorists. This is because, as I discuss below, whether other extant views can rightly be called reductionist depends on whether other theoretical posits which they take to be relevant to understanding—such as grasping and ability—can themselves be reduced to some epistemic state. And extant theories often do not address this issue. However, other views emphasising the idea of understanding being grounded in particular epistemic states include Kvanvig (2003), who emphasises the connection between *objectual* understanding and a body of justified (viz. internally coherent), broadly factive beliefs.

Grasping: In order to possess understanding, it is necessary to stand in a particular cognitive relationship—that of *grasping*—towards the object of understanding.

Under this proposal, understanding consists (at least in part) in possessing a sui generis mental state that differs from merely believing or knowing. One popular suggestion—defended by Bourget (2007)—is that grasping has a distinctive phenomenology belief and knowledge lack. However, a precise account of the nature of grasping has proved an elusive task, with even proponents of such a view conceding that “the question of the nature of [grasping] is perhaps the deepest in all philosophy.”²² Reductionist theories can avoid the deep waters of analysing grasping by claiming that there is nothing epistemically basic about understanding—rather, on reductionist views, understanding can be wholly constituted by facts about familiar epistemic states such as knowledge.

Grasping accounts also face further difficulties. Firstly, it has been suggested that even if (some) instances of understanding appear to come with a distinctive phenomenology, so do certain instances of propositional knowledge—for instance, it can seem appropriate to talk of ‘seeing’ or ‘grasping’ the truth of some analytically true mathematical proposition (see Grimm 2006). If this is the case, understanding-why cannot be individuated or distinguished from propositional knowledge by any grasping phenomenology. Secondly, it has been suggested that grasping accounts struggle to account for the way in which understanding is often non-occurrent or latent; it is unclear how non-occurrent understanding can be analysed in terms of possessing some phenomenally distinctive mental state.²³ Reductionist accounts avoid both of these issues insofar as knowledge requires no distinctive phenomenology and one can standardly be taken to know something non-occurrently or latently.

A second approach to understanding is the following:

Ability: One possesses understanding (at least partly) in virtue of possessing certain *abilities* regarding the object of one’s understanding.²⁴

There is clearly an intricate relationship between understanding and possessing abilities. Proponents of an ability-view claim that possessing certain intellectual abilities is not merely symptomatic of understanding, but rather *constitutive* of its very nature. One prominent ability-view, suggested by Hills (2017), is that understanding consists in the ability to perform certain reasoning tasks relating to the object of understanding—viz. to give and follow explanations, be able to draw conclusions, and so forth.

²² Strevens (2013: 511)

²³ On this basis, Lynch (2017) suggests that grasping might be central to *coming to understand* but not constitutive of understanding itself. One unexplored response might appeal to the *disposition* to grasp.

²⁴ Hills (2017) provides an influential statement of an ability-view of understanding why.

As the nature of abilities are far from settled, the eventual plausibility of ability-views will depend on what account of ability they are wedded to. Reductionist views can avoid the issue of the contested nature of ability by reducing understanding to more familiar epistemic states such as knowledge. A general worry with ability-views is that they might be overly restrictive in some ways and overly permissive in others. On the former front, Sliwa worries that defining understanding in terms of reasoning abilities leaves out the essential role that first-personal experience and relying on the advice of others can play in achieving particular instances of understanding. For example, it might be possible to conceive of cases where one's first-personal experience affords one valuable insights into some phenomenon—for instance of a moral phenomena such as understanding why some behaviour that amounts to sexual harassment is wrong—even though one lacks a sophisticated ability to reason regarding this phenomena.²⁵ Relatedly, Lipton (2009) discusses the possibility of 'inarticulate understanding' where one can acquire understanding by relying on images or models—he appeals to the example of using an orrery—without being able to articulate an explanation in propositional terms. In the other direction, one might worry that merely possessing abilities such as drawing conclusions or providing explanations licenses problematically liberal ascriptions of understanding. For instance, a chess computer can draw conclusions about complex positions (e.g. which side has the stronger position) and explain why this is so (king safety, piece mobility, etc.). However, it is far from clear whether such a mechanical system can possess the epistemic state of understanding, where the target phenomena is the same epistemic state we attribute to human agents. By contrast, reductionist views focusing on knowledge have resources to avoid these problems. As Sliwa puts it, knowledge is 'multiply realisable'—instances of knowledge (and hence degrees of understanding) can be acquired through first-personal experience, using models, relying on advice, and so forth. There is no *necessary* requirement that an instance of knowledge either results from, or constitutes, reasoning-abilities. Such theories also allow us to avoid overly-liberal understanding ascriptions by withholding knowledge attributions to non-intentional structures such as machines. For instance, one might deny that such structures can possess *beliefs* and consequently deny them knowledge. At very least, **Reductionism** enables the attractive position that ties the plausibility of attributing understanding to machines to our position on attributing them other epistemic states such as belief or knowledge.

While we have seen that **Reductionism** initially seems to be a rival to **Grasping** and **Ability**—and proponents of knowledge based variants of **Reductionism** often motivate it as such—**Reductionism** is actually compatible with each of these dominant views in the literature on understanding. This is because (at least *prima facie*) it is possible to affirm or reject any combination of **Reductionism**, **Grasping** and **Ability**. For instance, one might hold that understanding is constituted by certain abilities. This is consistent with it also being the case that knowledge is all there is to an instance of

²⁵ See Sliwa (2017) for some discussion of this idea.

understanding, just if one also holds that the abilities in question are themselves metaphysically dependent on the agent's body of knowledge. The same thought applies to grasping. Of course, one might have particular views of grasping or the nature of ability that rule out some combinations of **Reductionism**, **Grasping**, and **Ability**. For instance, if you thought that abilities are irreducible to knowledge; then you would deny the conjunction of a knowledge based version of **Reductionism** and **Ability**.²⁶ In principle, **Reductionism** is a very ecumenical framework with which to theorise about the nature of understanding. Going forward, I will focus on identifying the most plausible version of **Reductionism** as a metaphysical theory about the nature of understanding while remaining neutral on the plausibility of **Grasping** and **Ability**.

2. How to test reductionism

Having sketched the nature of reductionist theories about understanding why, my aim now is to develop a natural way to test presifications of such theories. In so doing I will be recommending a slight refocusing of the debate, but in a way that I suggest will afford us with a clearer grip on the underlying metaphysics of understanding.

Much extant discussion concerns ascriptions of *outright understanding*—i.e. involve cases about which we are asked to judge whether it is true or false that some subject *S* “understands why *p*” simpliciter. However, I propose is that it will be more fruitful to attend to *degrees of understanding* rather than judgements of outright understanding when testing various reductionist theories.²⁷ The easiest way to execute this approach will be to look at *comparative judgements* about whether one subject possesses a greater or lesser degree of understanding than another. The comparison subject can either be a separate person, or the same person across different times. Shifting the focus to degrees of understanding does not involve any controversial assumptions; it simply relies on the natural thought that if an agent's understanding why *p* outright is constituted by (e.g.) their knowledge, then their degree of understanding why *p* is also constituted by (e.g.) their knowledge.

But why should we prefer this approach? I propose that this focus has a number of independent advantages, while also providing us with a straightforward methodology with which to test reductionist theories of understanding. The first advantage stems from the fact that the gradability of understanding is a distinctive feature of understanding that sets it apart from knowledge; focusing on degrees of understanding zeros in on this central feature. Moreover, as our earlier discussion of ‘equivalence’ theories illustrated (at the beginning of §1), accounting for epistemic asymmetries between the degrees of understanding possessed by different agents constitutes a crucial test for a satisfactory theory of

²⁶ This issue leads quickly into difficult territory; the plausibility of reducing abilities to knowledge will depend, in part, on one's view of knowledge-how. The nature of knowledge-how and how it relates to propositional knowledge is a vexing issue in contemporary epistemology.

²⁷ Another epistemologist who has focused on degrees of understanding is Kelp (2017).

understanding—focusing on comparative judgements ensures that we keep this desideratum firmly in mind. A second advantage is methodological. Attending to degrees of understanding rather than outright understanding allows us to examine very fine-grained differences between epistemic subjects. For, if we suppose that there will often be a considerable difference between the epistemic positions of subjects who outright understand why something is the case compared with those who lack understanding, then any comparison between these subjects would require us to keep track of a variety of cognitive changes. This raises the *prima facie* worry that it might not always be obvious exactly which of these changes our understanding attributions were tracking. Looking at degrees of understanding instead allows us to attend to smaller, incremental epistemic changes—thus allowing us to more confidently trace back our intuitive judgements about understanding to very specific changes in a subject’s epistemic position.

Thirdly, focusing on degrees of understanding controls for unhelpful theoretical noise due to context-sensitivity influencing judgements about outright understanding. It seems plausible that the standards for, say, understanding why climate change is influenced by human behaviour might vary between a high-school class and an academic colloquium. Focusing on comparative judgements about degrees of understanding bypasses such complications. For example, while the threshold for being ‘tall’ might vary between a basketball court and a nursery school, judgements as to whether LeBron James is taller than a small child remain true across contexts. Similarly, while the standards for outright ascriptions of understanding why the climate change is influenced by human behaviour might vary between a high-school class and an academic colloquium, the comparative judgement about whether the high-school student or the professor has better understanding does not vary. By excluding contextual variables, the focus on comparative gradable judgements facilitates a cleaner focus on the underlying metaphysics of understanding. And, as a final advantage, the comparative approach I suggest allows us to issue instructive verdicts in cases where two agents *both* understand outright at some salient context, but possess different degrees of understanding.

In addition to these advantages, focusing on degrees of understanding allows us to identify some commitments of reductionist theories about understanding that have not been fully unpacked in the literature. Once we have noted these commitments, they can then be used to provide a useful testing procedure for the particular version of reductionism we are examining. The commitments I have in mind can be best appreciated if we consider a particular presification of reductionism, such as the knowledge based version. Consider: if the extent to which an agent understands is to be fully explained by what the agent knows, this entails a variety of theses about what can bring an agent to possess a *different degree* of understanding. In terms of Sliwa’s metaphor—just as the physicalist claims that mental changes must be explained by some physical change, a knowledge based reductionist about understanding must hold that changes in understanding must be explained by changes in what is known. This line of thought yields three principles that any reductionist theory of understanding will require a

version of.²⁸ (Let ϕ be a placeholder for some epistemic state, and ϕp be a placeholder for being in that state with relation to some proposition).

IMPROVEMENT: Only acquiring new (or better-connecting) ϕ can directly improve your understanding; acquiring epistemic states other than ϕ cannot directly improve your understanding.²⁹

CONVERSION: If ϕp (partly) constitutes your understanding, your understanding will be impoverished if your relation towards p is replaced by some other relation that does not entail ϕp .

IMPOVERISHMENT: Only losing ϕ (or connections between ϕ) can directly impoverish your understanding; losing epistemic states other than ϕ cannot directly impoverish your understanding.

These principles can be used to test specific reductionist theories in the following way: substitute ϕ for whichever epistemic state the reductionist claims is the basis of understanding. This will give us three principles that are constitutive of the particular reductionist theory in question. And if any of these principles turn out to be false, then so does the reductionist theory that entails them.

3. Against knowledge based reductionism

Having outlined a methodology to test reductionist theories, we can now turn to scrutinise the knowledge based variant. This will be the initial focus because it is the most popular version of reductionism that has been explicitly defended in the literature. My rejection of the knowledge based

²⁸ An anonymous reviewer suggested that Kelp's view, on which understanding is measured against approximating maximally systematic knowledge, might avoid being committed to such principles. For instance, it might be the case that an agent who believes that p (where p is some true proposition) better approximates maximally systematic knowledge than an agent who has no doxastic attitude towards p . I think this is a promising route for a reductionist to explore. However, pursuing an approach that is not committed to the principles I outline risks abandoning the metaphysical ambitions of reductionism. For, if a knowledge based reductionist theory of understanding holds that understanding can improve in virtue of acquiring a non-knowledgeable belief, then it is not clear that such a theory is offering a statement about the *metaphysical dependence* of understanding on knowledge.

²⁹ I include the qualification regarding 'direct' improvement (and below, impoverishment) of understanding to respect the fact that reductionists are interested in metaphysical rather than merely incidental relationships between understanding and other epistemic states. Thus, **Improvement** and **Impoverishment** are not violated by cases where a subject acquires or loses some epistemic state other than ϕ and this brings about an *indirect or mediated change* in their understanding. For instance, an indirect impoverishment might occur when a subject loses an unrelated belief and this happens to start a psychological process (e.g. a crisis of confidence) that results in their losing further knowledge relevant to the object of understanding. We should regard the loss of the initial belief as an indirect cause and the loss of the knowledge as the direct cause of the change in understanding. Hence, the principle cannot be violated by the former.

view will then be used to prompt an alternative version of reductionism focusing on what is truly believed. The following theory will be our target:

URK: Knowledge is all there is to understanding—the extent to which an agent understands-why p can be explained wholly in virtue of what they know.³⁰

And, substituting knowledge for our epistemic placeholder, we get the following three principles:

K IMPROVEMENT: Only acquiring new (or better-connecting) knowledge can directly improve your understanding; acquiring epistemic states other than knowledge cannot directly improve your understanding.

K CONVERSION: If knowing p (partly) constitutes your understanding, your understanding will be impoverished if your relation towards p is replaced by some other relation that does not entail knowing p .

K IMPOVERISHMENT: Only losing knowledge (or connections between knowledge) can directly impoverish your understanding; losing epistemic states other than knowledge cannot directly impoverish your understanding.

To anticipate my argument, I will defend the following thought—while reductionism is compelling, the exclusive focus on knowledge is unwarranted. If we accept that the degree to which one understands why p (where p is some substantive fact or phenomena) depends on a network of other cognitive commitments, it is not plausible to hold that *every* commitment that contributes to an agent's understanding will be a known proposition. This thought seems especially compelling when it comes to a high level of understanding regarding some complex phenomenon. A historian's understanding of why the Roman Empire collapsed will be comprised by a vast range of her beliefs and assumptions. Each member of this vast network has a variety of aetiologies—inference, abduction, testimony of varying shades of reliability, and so forth. On pain of an overly liberal epistemology, I think we should reject the suggestion that each of these commitments must be known. This will open the space for an alternative version of reductionism that instead focuses on what is truly believed, regardless of whether those true beliefs are knowledgeable.

In order to press my objection I will scrutinise an example of moral understanding recently discussed by Sliwa (2017), a prominent defender of **URK**. The instance in question is understanding <why eating meat is wrong>.³¹ Sliwa claims that cases in which someone better understands why meat-eating is

³⁰ Again, I take a well-connectedness proviso as implied.

³¹ A salient question is whether moral understanding might be *sui generis* compared to understanding some non-evaluative state of affairs. Here, I follow the contours of the existing debate which discusses examples of moral understanding as evidence for general theses about understanding why. Addressing the difference between the epistemology of evaluative and non-evaluative domains is outwith the scope of this essay.

wrong than someone else are explained by that person having *more knowledge of the things that make meat-eating wrong*: for example, knowing that the meat industry involves maltreatment of animals, and/or knowing that is bad for the environment. This is a plausible idea—knowing more typically results in greater understanding. However, I think this reductionist move forces a broader perspective than its proponents have appreciated. If we accept that one’s degree of understanding why *p* is explained by certain epistemic relationships towards various relevant considerations, then there will turn out to be many such relevant considerations—and it will not turn out to be plausible to suppose that every epistemic relation towards these is that of knowledgeability. A case will elucidate.

3.1 Putative Counterexample

As we are focusing on comparative judgements about understanding, it will be helpful to have an initial reference point. I assume the following characters to be epistemic duplicates: they start in the same epistemic position and have identical dispositions to respond to new information.³²

Annie, Brian & Carl: Each of these characters are fairly unimaginative inquirers, but they do have some knowledge. For example, they know that the environment is fragile and sensitive to human behaviour, that methane is a gas that (somehow) is produced by certain industries such as petroleum manufacturers, that methane is bad for the environment because it (somehow) heats the globe, and that supporting practices (e.g. the petroleum industry) that harm the environment is morally wrong. They don’t currently think that the meat industry harms animals or the environment.

Now, consider one’s epistemic judgements should the following occur:

Annie: Annie sees a headline in the *New York Times* (a paper she knows is highly reliable) that reads—“MEAT INDUSTRY RESPONSIBLE FOR MORE METHANE THAN PREVIOUSLY THOUGHT”. Without reading on, she comes to believe <the meat industry produces methane> and on this basis infers that <the meat industry is morally problematic in the same way as the petroleum industry>.

Brian: Brian finds himself surfing a conspiracy website—featuring articles about fake Moon Landings and Paul McCartney dying in 1966. He sees a headline that says: “MEAT INDUSTRY METHANE DESTROYING THE PLANET: CLICK HERE FOR MORE”. Without clicking, he comes to believe <the meat industry produces methane> and on this basis infers that <the meat industry is morally problematic in the same way as the petroleum industry>.

³² This is an artificiality, but not a vicious one. The same point could be made by taking one character and discussing counterfactual scenarios.

Carl: Nothing happens. Carl retains his ignorance regarding the ethics of eating meat, neither believing that it harms animals nor is bad for the environment.

3.2 Discussion

The case is set up such that the following is the case:

- (1) While Annie's belief that the <the meat industry produces methane> is knowledgeable, Brian's belief that <the meat industry produces methane> is not knowledgeable.

Annie acquires knowledge that the meat industry produces methane by trusting reliable testimony from a trustworthy source. Brian, on the other hand, puts his faith in an untrustworthy source that happens to get it right on this occasion (we can further suppose that the author of the article lacks knowledge, and that the article itself contains a variety of falsehoods). A natural way to substantiate the intuition that Brian lacks knowledge that the meat industry produces methane is by supposing that his belief, while true, is not justified.³³ This can be vindicated on either an externalist or internalist conception of justification. On the former, we can merely note that Brian has not used a reliable belief-forming procedure; there are close worlds in which the headline would have contained falsehoods and been believed by Brian regardless. On an internalist conception we might add to the case that Brian, although credulous, has been warned before that he ought not to trust such websites but errs by failing to bear this in mind when forming his belief.

In addition, by stipulation it is true that:

- (2) All three started at epistemic parity.

Further, I argue it is intuitive that:

- (3) Annie and Brian *gained* understanding of why eating meat is wrong while Carl did not.

Annie and Brian both gain the belief that the meat industry is morally problematic, and infer this belief from the truth that it produces environmentally harmful methane, and their background belief that supporting environmentally harmful practices is wrong. Both can now provide a rudimentary correct explanation of why eating meat is wrong. Carl, by contrast, has no such belief and can cite no reasons in favour of the morally problematic status of eating meat. Comparatively, Annie and Brian have better understanding than Carl. Furthermore, I suggest:

- (4) Annie and Brian understand why eating meat is wrong equally well.

³³ A different way to press this thought might appeal to the fact that Brian benefits from environmental *epistemic luck* in believing the headline containing truth rather than directing his attention towards the various false headlines surrounding it. See footnote 17 for references on understanding and luck. (However, these discussions were pitched largely at early 'equivalence' theories; not the reductionist views we target here.)

If you accept (3), it is very difficult to deny (4). As they started at epistemic parity, share identical epistemic dispositions, and acquired beliefs that have *exactly the same content*, the natural assumption should be that they now understand why eating meat is wrong to the same extent. We can bolster this thought by appealing to the sorts of abilities they have regarding the target propositions, if we assume that such abilities are diagnostic even if not constitutive of understanding. For instance, each will possess identical dispositions to explain why you should not eat meat and to justify actions performed on the basis of their new moral belief. Moreover, as they represent the world in exactly the same way, with the only discernible difference between them being the aetiology of one of their beliefs, we can further assume that they have the same internal ‘grasp’ of why eating meat is wrong.

Taken together, (1) – (3) spell trouble for **URK**. For, it seems like we have a case in which Brian’s understanding why *p* improved in virtue of acquiring a true belief that he does not know. In the case, Brian’s belief that <the meat industry produces methane> is central to his degree of understanding—without this belief, he would revert back to his former epistemic position and lose the increased understanding we attributed him. Indeed, Brian understands why eating meat is wrong *better* than somebody who simply takes it on reliable testimony that eating meat is wrong because it is bad for the environment, where this person lacks any beliefs about methane, noxious gases, the globe heating up, and so forth. The degree to which one understands why eating meat is wrong depends not only on whether one believes that it harms the environment, but also on why one believes that it harms the environment. Given this, it is hard to avoid the conclusion that Brian’s increased understanding is partly constituted by his non-knowlegeable true belief that <the meat industry produces methane>. Thus, we have found a putative counterexample to **URK**—the thesis that understanding reduces just to what is known—by showing that the principle of **K IMPROVEMENT** is false; one can increase one’s understanding (partly) in virtue of acquiring a non-knowlegeable true belief. Once we grant this, it is easy to see why **K IMPOVERISHMENT** is also false. Were Brian to stop believing that the meat-industry produces methane, his understanding would deteriorate in virtue of losing a commitment that falls short of knowledge.

Moreover, (4) states that Brian is on epistemic parity with Annie, who *knows* the same content—<the meat industry produces methane>—that Brian merely truly believes. Once we grant this, **K CONVERSION** becomes dubious. **K CONVERSION** predicts that Annie’s understanding would deteriorate if she lost knowledge but retained belief in the fact that the meat industry produces methane. But this seems doubtful if we grant that Brian understands equally well despite not knowing this same content. **K CONVERSION** might occur, for instance, within any framework under which knowledge can be *defeated* or can be degraded via exposure to *high-stakes*.³⁴ For example, Annie could lose knowledge that the meat industry produces methane via *normative defeat*—when there is some

³⁴ Hills (2015) provides discussion of the relationship between epistemic defeat and understanding.

additional belief she ought to have entertained that would either conflict with, or undermine, her initial belief.³⁵ Consider a case where Annie learns that the edition of the *New York Times* she consulted was full of errors, but pigheadedly continues believing regardless. On various theories of defeat, Annie would lose knowledge that the meat industry produces methane. But, I suggest, losing this knowledge (while maintaining true belief) would not impoverish her understanding of why eating meat is wrong.

3.3 Objections Considered

Proponents of **URK** might try to claim, despite the intuitions the foregoing case elicits, that knowledge nonetheless plays an indispensable explanatory role in accounting for differences in understanding. There are a number of possible strategies for attempting to do so, of which only some can be considered here. I will take it to be implausible to suggest that both Annie and Bob have—like Carl—no understanding whatsoever. Furthermore, insofar as Annie and Brian believe exactly the same content, and they intuitively understand why eating meat is wrong to the same degree, I assume we prefer a unified explanation of the epistemic improvement they have made.³⁶ As such, the best response will be to attempt to show that their improved understanding is explicable by **URK**—viz. that there is something they know that explains our judgements.

Conditional Knowledge? An initial objection might run as follows: the understanding we attribute can be fully explained by *conditional knowledge* of the form <if the meat industry produces methane, then it would be bad for the environment, hence ...>. One might press this objection by supposing that the testimony acquired by Annie and Brian only *better connects* knowledge they already possess, and by suggesting that just this conditional knowledge explains their ability to infer that eating meat is morally problematic.

However, this objection is untenable. We stipulated that all three characters have this conditional knowledge at the outset (consider their attitude towards the petroleum industry) and yet Annie and Brian improve their understanding without acquiring new conditional knowledge. The issue is that, by itself, conditional knowledge doesn't suffice to explain understanding-attributions regarding a *specific* phenomenon. To exemplify: imagine that, unbeknownst to anybody, domestic cats produced vast quantities of methane. You and I *know* that methane is bad for the environment, and *know* that if supporting the domestic cat industry produced lots of methane, then keeping cats would be bad for environment and hence morally deleterious. But if it were discovered tomorrow that domestic cats produce lots of methane, we would not say that 'we understood why keeping domestic cats is wrong all along'. At best, conditional knowledge only puts one in a position to understand.

³⁵ Some knowledge-first epistemologists argue against orthodox theories of defeat. See Lasonen-Aarnio (2014) for a good example. My argument stands or falls independent of my discussion of **K CONVERSION** and defeat—undermining **K IMPROVEMENT** or **K IMPOVERISHMENT** suffices to refute **URK**.

³⁶ Although I will consider and reject a disjunctive explanation, at the end of this section.

Knowledge-acquisition? A second more powerful objection is that Annie and Brian improve their understanding because they both come to know something Carl does not. Namely, *they both come to know* either that <the meat industry is bad for the environment> (MIBE) or <eating meat is morally wrong> (EMMW). And, on this objection, only this knowledge (part) constitutes their understanding rather than any mere true belief.

In the vignette, I avoided the question of whether the characters acquire knowledge of (MIBE) or (EMMW). In the case of Brian, affirming this would rely on the dubious assumption that we invariably gain *knowledge* by making inferences from non-known premises. If Brian is to know (MIBE), this would involve claiming that Brian knows <the meat industry is bad for the environment> via the use of a premise he knows <if an industry produces methane then it is bad for environment> and a (true) premise that he does not know <the meat industry produces methane>. What we should say about (EMMW) is substantially more complex and will depend on what we should say about (MIBE); as such, I will only focus on the latter.

Whether we can gain knowledge through mixed inferences where some premises are known and some are not is a difficult topic.³⁷ However, for our purposes, we can simply note that the schema this objection relies upon is not knowledge-producing *in general*. To see this, consider the following schema (let K = knowledge and let B = truly believed):

- (1) **K:** If X then Y
- (2) **B:** X

- (3) **K:** Y

It is not hard to appreciate that there are many instances of this schema where the subject does not acquire knowledge. For example:

- (1) I know that if the Butler's prints are on the knife, then the Butler is the murderer.
- (2) I believe, without evidence, that the Butler's prints are on the knife.

- (3) I deduce that the Butler is the murderer.

Clearly, I do not know the Butler is the murderer here even if it happens to be true. So, it is far from obvious that we should grant that Brian *knows* the meat industry is bad for the environment or that eating meat is morally wrong. Certainly, given his credulity, his route to these beliefs does not seem to have the properties of being reliable, counterfactually robust, or excluding salient error-possibilities. The proponent of **URK** needs to show that there is no way to describe the case such that Brian fails to acquire knowledge that the meat industry is bad for the environment or that eating meat is morally

³⁷ There is a literature on this issue stemming from Warfield (2005). However, the cases found in that debate are structurally different from those at issue here, often focusing on false premises.

wrong. This seems like a potentially quixotic battle, but I leave it as a challenge for proponents of that theory.

The second more fundamental problem with this objection is that it isn't clear that attributing some extra piece of knowledge suffices to defuse our initial worries. As we have pressed, the challenge is to explain epistemic asymmetries, and the strategy adopted by **URK** was to say that an agent's understanding is merely a matter of what they know. However, even if we accept that Annie and Brian both know either (MIBE) or (EMMW), we can conceive of further cases where some other agent has better understanding of why eating meat is wrong in virtue of having a more fine-grained appreciation of why the meat industry is bad for the environment. **URK** will again have to identify and appeal to further knowledge to explain this. But the puzzle can be reiterated again, taking cases where an even more extensive and complicated body of belief is at issue. And so on. The worry is that the proponent of **URK** ends up being committed to insisting that every single one of the many beliefs relevant to explaining any conceivable epistemic asymmetry towards one target fact must count as knowledge, or they will be unable to explain such epistemic asymmetries. This is a dilemma: either **URK** cannot fulfil its explanatory ambitions, or it ends up licensing overly liberal knowledge-ascriptions.

Proximity to knowledge? A further objection might be that Annie's understanding improves because she knows some of the propositions discussed above, while Brian's improves because—by acquiring a true belief—he comes *closer to knowing* an important explanatory fact about the ethics of eating meat.

This objection is problematic insofar as it cannot straightforwardly treat as correct the intuition that both Annie and Brian understand equally well. However, even if one is willing to grant this on broader theoretical grounds, the proximity to knowledge objection is independently dubious. We can imagine different agents using a variety of similar, yet differentially reliable processes, to form exactly the same belief that figures in their understanding of why something is the case. The objection we are considering would yield the verdict that understanding increases incrementally with the reliability of the belief-forming process used. I lack a knock-down argument against such a position, but register my doubt that it is plausible to suppose that our attributions of understanding are sensitive to facts such as that one of the beliefs contributing to one's understanding was formed using an incrementally more or less reliable process, with no difference in the content of what is believed.

*A disjunctive explanation?*³⁸ When introducing constraints on possible objections, I suggested that we want a unified explanation of the epistemic improvement made by Annie and Brian. However, one might question whether excluding disjunctive explanations was really well-motivated. For instance, demanding a unified explanation of Annie and Brian's epistemic improvement rules out responses like the following: Annie understands better in virtue of knowing *p*, while Brian understands better in virtue

³⁸ I am grateful to an anonymous referee for comments and suggestions that were invaluable for this section.

of knowing some different body of knowledge q , where q is self-regarding knowledge such as: <according to a source he trusts, the meat industry produces methane> and <it is very likely on his evidence that the meat industry is morally problematic in the same way as the petroleum industry>. I'll end by briefly arguing against this sort of disjunctive explanation.

An initial reason to be sceptical of such an approach is that it is far from clear that appealing to self-regarding epistemic attributions really provides much succour to the defender of **URK**. We might conceive of a case in which Brian *merely believes rather than knows* that his evidence supports his conclusions about the meat industry. For instance, we could suppose that Brian has just had a conversation with a sensible friend who warns him against trusting unreliable websites, but that Brian irrationally disregards this evidence when deciding whether or not to believe the conspiracy website about the meat industry. I think this would be bad news for our evaluation of Brian's intellectual character, but would not impoverish his object-level understanding of the meat industry.

Even if our intuitions go hazy when we consider cases involving higher-order beliefs about our evidence, there are further reasons to prefer a unified approach, and avoid appealing to such self-regarding knowledge. Consider the following: (i) intuitively, Annie and Brian have equal degrees of understanding—so, on the knowledge based account, we should expect that they know the same amount about the meat industry, yet (ii) if it is legitimate to ascribe self-regarding knowledge to Brian, it is surely also legitimate to ascribe it to Annie, so she will possess further knowledge in addition to her object-level knowledge about the meat industry. And so we would yet again be faced with the burden of explaining why they appear to understand equally well, despite asymmetries in what they know. An anonymous referee suggested to me that, at this juncture, defenders of **URK** might attempt to reject the judgement that Annie and Brian understand equally well by appealing to the context-sensitivity of understanding attributions. One such attempt might try to identify a context at which the outright judgement 'A and B understand equally well' comes out false. A second such attempt might claim that gradable attributions can themselves be context-sensitive. For instance, in the context of discussing an unequal race, it might be permissible to make assertions like: "A is as far from the starting point as B and both are further than C" even if A is 20.4km away from X while B is 19.7km away. Thus, one might wonder whether attributions of the same degree of understanding might be permissible in some contexts but not in others. These possibilities create the worry that the apparent equality in understanding might tacitly depend on assessing this statement at a less demanding context—perhaps at a more demanding context, this judgement might be false. However, if such contextual factors did explain our intuition about parity between Annie and Brian, one would have to suppose that the context at which we were evaluating Annie and Brian earlier involved less demanding standards of accuracy than other readily conceivable standards. But it is hard to see why this would be so. After all, the context of the present essay is an examination of the relationship between knowledge and understanding, where we are taking particular care to provide the most accurate description of the epistemic position of the characters we

discuss. We are not in a situation in which loose-talk is permissible because subtle distinctions are unimportant (such as when identifying the likely winner of an uneven race). Rather, we were explicitly attending to the subtleties of the case. So I am not tempted by this alternative explanation.

4. An alternative reductionism?

The inability of **URK** to explain the counterexample outlined earlier—and the consequent failure of **IMPROVEMENT**, **IMPOVERISHMENT** and **CONVERSION**—resulted from attempting to reduce understanding *exclusively* to knowledge. Nonetheless, reductionism remains an appealing framework for theorising about understanding. Hence, we should investigate whether it is possible to pursue a reductionist programme for understanding without the exclusive focus on knowledge. After all, there is no necessary connection between suggesting that (i) understanding is a complex epistemic state constituted by other epistemic states and (ii) knowledge is all there is to any instance of understanding. Indeed, as we noted earlier, there a variety of possible reductionisms about understanding. Some alternative approaches could also be *univocal*; i.e. reducing understanding to one type of epistemic state, while others are *pluralist*; i.e. reducing understanding to multiple epistemic states. A parallel: one might have a theory of forests claiming that a forest is metaphysically constituted by trees, or one might have a theory on which forests are constituted by trees, shrubs, and grasses—both are reductionist.

We found that an agent’s understanding of why something is the case can rest on what is known *and* truly believed. There are two reactions to this datum: one is to endorse a theory on which understanding univocally reduces to a body of true belief (knowing, after all, is just a mode of truly believing).³⁹ Another is to advance a pluralist theory on which understanding requires some knowledge, but tolerates non-knowledgeable true belief. The latter approach is an interesting one that I will not explore in detail. Rather, I wish to tentatively defend the former approach—the idea that understanding reduces to a body of true beliefs.⁴⁰ I think that we can motivate a truth based version of reductionism about understanding as an approach well worth taking seriously:

URT: True belief is all there is to understanding—the extent to which an agent understands-why *p* can be explained wholly in virtue of what they truly believe.⁴¹

Before defending this approach, it is important to note that it is hard to conceive of any case where a normal inquirer who possesses understanding on the basis of a body of true beliefs will lack knowledge of each of these beliefs. For instance, any understanding of why eating meat is wrong will depend on very basic beliefs about, e.g., what eating is, what animals are, and so forth. It is difficult, perhaps

³⁹ This is true on an orthodox conception whereby knowledge is a property of belief. But it is also easily accommodated on a knowledge-first framework whereby knowing is primitive but entails true belief.

⁴⁰ As mentioned earlier, the factivity of understanding is controversial and below I tackle some objections to factive approaches to understanding.

⁴¹ As with knowledge based views, I again take a ‘well-connectedness’ proviso to be implied.

impossible, to envisage scenarios in which an agent possesses true beliefs about such things without having *any* knowledge whatsoever—even limited inquirers will know a great deal of what they truly believe via the justification they acquire through appreciating rudimentary coherence relations between their beliefs, by relying on testimony, or by the use of reliable processes such as perception or deduction. By the same token, as fallible inquirers, we will also inevitably have true beliefs that fail to count as knowledge for various reasons. For instance: sometimes our true beliefs are insufficiently justified to count as knowledge, and sometimes they benefit from substantial epistemic luck. In practice, most cases of understanding will be mixed cases involving both true belief and knowledge.

However, in favour of my univocal truth based view, I suggest that it is not essential that any *particular* proposition be known rather than truly believed. So, if we imagine two agents who share a body of true beliefs about *p* with *exactly the same content*, **URT** predicts that they will understand why *p* to the same degree. This is so, I suggest, even if one agent possesses less knowledge than a second agent because of factors that do not vary the content of what they truly believe. Examples of such factors that preclude knowledge without entailing changes elsewhere in the content of one's body of belief might include: being in a Gettier-case with respect to some proposition, unwittingly relying on an unreliable testifier, being victim of a normative defeater, and so forth. In a slogan—identical true belief entails identical understanding.

The most immediate objection to this view is that it sounds like a problematically liberal view of understanding. For example, isn't it deeply implausible to attribute understanding to an agent who has just luckily stumbled across the right views? In response to this worry, I suggest that any initial discomfort about attributing understanding to an agent who lacks knowledge will often be explained by the natural thought that a lack of knowledge will often affect the *content* of one's overall body of beliefs. For example, we would rightly balk at ascribing any moral understanding to someone who formed their moral beliefs using tarot cards. Of course, one cannot come to *know* moral truths on the basis of such a hopeless belief-forming procedure. However, more significantly, the content of the epistemic commitments held by an agent who used such a procedure—even if it caused them to endorse a claim like 'eating meat is morally wrong'—would be rather different from the content of what is believed by an agent who believes that eating meat is wrong in virtue of reflecting upon their other beliefs about the environmental consequences of supporting the animal farming industry. Quite simply, these agents will believe different things. Similarly, I suggest that the natural thought that greater knowledge yields greater understanding is only correct when there is some additional content that the agent in question believes—such as knowing the connection between some evidence *E* and the target phenomena. Knowing something on the basis of using a reliable process does not afford greater understanding compared to forming exactly the same belief by (unwittingly) using an unreliable process, unless using the reliable process affords you with some additional true content. With these points in mind, the truth based approach becomes less radical than it first appears. It simply claims that *if we hold completely*

fixed the content of an agent's overall body of belief, it is irrelevant to an agent's degree of understanding whether any single belief within this broader network is accorded the status of knowledge or not.

URT, the truth based theory I have outlined, inherits all of the advantages of reductionist theories such as **URK** over rival approaches such as **Grasping** or **Ability**. To briefly recap, the main advantages were that: knowledge is a familiar and well-theorised epistemic state; knowledge can be non-occurrent; knowledge is multiply realisable (it can have a diverse aetiology and need not *entail* certain abilities); and knowledge-ascriptions, unlike abilities, can be confined to familiar epistemic subjects (e.g. human believers) and exclude non-agential structures. A theory that identifies true belief as the reductive base of understanding can cite the very same marks in its favour. Indeed, on many metaphysical frameworks, the reason that knowledge possesses the aforementioned features is because *knowledge is a property of belief*. Belief is a heavily theorised epistemic state, is often attributed non-occurrently, is multiply realisable, and—on many accounts—has a cognitive element that is absent in the case of non-agential structures such as machines. In this sense, to the extent that knowledge based approaches are preferable to rival accounts, so is the truth based account.

5. Understanding better and worse

Recall a central desiderata for reductionist views: explaining the gradability of understanding. Knowledge based reductionism offered an ostensibly simple and clear diagnosis, claiming that agents understand better (or worse) when they know more (or less). A natural question, then, is how a truth based approach deals with this issue?

On any plausible account of understanding, there will be a very close link between understanding why something is the case and accurately representing explanations for why that thing is the case.⁴² This is the common link between **URK** and **URT**—both theories recognise the importance of accuracy in one's epistemic position towards the target phenomena. But what explains *the degree* to which some true belief contributes to our understanding? After all, there are many beliefs that one can have about a given subject-matter that, while true, are rather peripheral to an understanding of it. Answering this question is not straightforward. However, to make some progress, we can take our cue from the response to it outlined by proponents of the knowledge based view.

While **URK** holds that agents understand better when they know more, clearly not all pieces of knowledge make an equal contribution to how well you understand why something is the case. As proponents acknowledge, we should not think of 'knowing more' in quantitative or numerical terms—

⁴² Hazlett (2017) discusses the idea that understanding just is correctly representing explanations for different phenomena. He remains neutral on the issue we discuss here: what epistemic states—i.e. the bearers of representations—understanding reduces to.

it is a fool's errand to attempt to measure how much an agent knows by *counting* known propositions. An intuitive example: someone might be inquiring into whether—(i) there are more than 100 beans of coffee in the jar, or (ii) the earth is roughly spherical rather than flat. There is a sense in which settling the latter question brings about a more extensive cognitive change than the former; in effect, you come to know more by knowing that the earth is roughly spherical than by learning that there are 103 beans in the jar. Similarly, someone with false beliefs about the number of beans in a coffee jar seems less misguided than someone with erroneous beliefs about the shape of the earth.

To explain these judgements, Sliwa draws on a theoretical model due to Treanor (2013) that suggests measuring knowledge in terms of a similarity metric—one knows more when one's representations are closer to the way the world actually is. Applied to our example, the idea would be that while a false belief about the number of coffee beans is only a minor deviation from reality, wrongly thinking the earth is flat substantially reduces the overall fidelity of your commitments. Sliwa's presification of this model notes that explanations for why something is the case come at different levels of generality. Her claim is that an agent with greater understanding will, by knowing more, cognitively privilege more fine-grained (correct) explanations and rule out misconceived explanations. This is best appreciated with an example: someone might learn, via testimony, that eating meat is morally wrong because factory farming is cruel. They will have less understanding than another agent who learns this same fact by looking at documentary footage and forming their own view. The former will privilege a less fine-grained explanation. For instance, they will not have ruled out believing that the problem with factory farming is simply the animals lacking the ability to go outdoors. This is consistent with the misconception that factory-farmed animals enjoy a comfortable indoor habitat. In contrast, the agent who has seen the footage and thus acquired more knowledge, will privilege a more fine-grained explanation for the cruelty of factory-farming—the cramped space, the unsanitary conditions, the distress of the animals, and so forth. The epistemic space left open by their commitments will not include the possibility that the animals enjoy a comfortable habitat. In this sense, someone who has first-personal experience will have a greater understanding by knowing explanations of a finer grain.

If we accept this rough line of thought on which understanding consists in similarity between one's representations and the world, and that this can be usefully measured by looking at the generality and specificity of the (correct) explanations one (is inclined to) endorse and reject, then **URT**, the truth based account of understanding, can co-opt exactly the same resources. To see this, return to the case we used against **URK**: we had two agents who came to hold that eating meat is wrong in virtue of believing—in one case knowledgably, in the other merely truly believing—that the meat industry produces methane. The only difference between these agents was the aetiology of one of the beliefs that (partly) constituted their understanding—any other further beliefs or misconceptions they had were shared in common. In such a case, it seems compelling to suppose that the content of their cognitive commitments reflects reality equally accurately, and that both (are inclined to) accept and reject exactly

the same explanations. Both agents, therefore, privilege explanations for why eating meat is wrong with exactly the same fineness of grain. This dovetails nicely with our prediction that they understand to the same degree. In other words, the theoretical resources used by the knowledge based view to offer a way of measuring different degrees of understanding are consistent with the truth based approach I defend here.

6. Understanding is Factive

My defence of a truth based version of reductionism explicitly raises a question we have largely sidestepped by focusing on views that take truth and knowledge as the building blocks of understanding—namely, must we accept a factive view of understanding? Attempting to get clear on the relationship between understanding and truth has been another central debate about the nature of understanding. Indeed, the denial of a factive view of understanding is probably the most powerful objection to the truth based version of reductionism I have outlined. It is nonetheless worth noting that a critique of the factivity of understanding is entirely consistent with taking a reductionist approach to understanding. A tenable theory is that understanding reduces to a body of beliefs that possess some property other than truth; this would still be a reductionist theory that eschews any essential reference to the idea of ability or grasping. However I will now attempt to defuse the objection that understanding is not a factive state, arguing that we can accommodate many of the insights prompting non-factive conceptions of understanding on the **URT** framework defended above.

6.1 Elgin against the factivity of understanding

To focus the discussion, I will chiefly consider the criticisms levelled by Catherine Elgin in a series of papers and in her recent book *True Enough*, where she develops a sophisticated approach to understanding that aims to ‘dethrone’ the centrality of truth.⁴³ Key to her arguments are the thought that a strictly factive notion of understanding cannot account for the manifest success of scientific theorising, including the widespread role of idealisations and ‘true enough’ inaccuracies that are characteristic of the upwards trajectory of scientific understanding. I will close the essay by arguing that we can vindicate Elgin’s insights about the success of scientific theorising within a strictly factive conception of understanding. In particular, I focus on her argument that only a non-factive theory of understanding explains why falsehoods can help us move from ignorance towards a sophisticated understanding of our subject-matter. I argue that, given very modest assumptions about the contents of belief, we can defuse her argument and explain why judicious falsehoods can play an important role in scientific education while maintaining a factive conception of understanding.

⁴³ For other work on the relationship between understanding and truth, see, amongst others, Grimm (2012) and Rancourt (2017) for useful discussion.

Elgin’s way of putting the factivity view she argues against is amenable to the sort of reductionist approach to understanding we have been discussing thus far because she focuses on instances of understanding which are *comprised* by propositional commitments. According to Elgin, a factive view of understanding amounts to the following:

Factivity: “If ‘understanding’ is factive, all or most of the propositional commitments that comprise a genuine understanding are true”.⁴⁴

Elgin’s discussion primarily trades in attributions of *objectual* understanding rather than, as we have focused on so far, understanding why something is the case. My critique, in order to track Elgin’s discussion of particular examples, will also shift to focus on objectual understanding. I do not think that this shift is in any way problematic for my purposes here. Khalifa (2013) argues that objectual understanding reduces to understanding why something is the case. If this is right, then the choice of which to focus on is a matter of convenience rather than of deep philosophical significance. However, even without assuming this, we should note that Elgin chooses to focus on objectual understanding because she believes it best illustrates the need for a non-factive approach⁴⁵; so, if we can argue against her account on its own terms, then this is a solid basis for generalising our results to understanding why. And finally, it is in general not particularly difficult to convert attributions of objectual understanding into very similar attributions of understanding why. For instance, Elgin discusses objectual understanding of ‘human evolution’. This might naturally be converted into an attribution of understanding why by focusing on an agent’s understanding of why humans evolved as they did.

A chief concern for Elgin in denying **Factivity** is to undermine a foundational view in epistemology that she dubs ‘veritism’, namely the idea that truth-conduciveness is the only appropriate standard of assessment for epistemic policies, practices, and their products.⁴⁶ Rather, Elgin wants to replace veritism with a view on which understanding rather than truth is the central aim of many of our epistemic practices, including scientific theorising, where understanding does not conform to **Factivity** but rather is matter of being able to ‘exploit information’ and to know “how to wield it to further [one’s] cognitive (and perhaps practical) ends.”⁴⁷

I want to show that we do not need to accept Elgin’s central arguments against **Factivity**, before suggesting that we can accommodate the insights behind her argument within the framework of **URT**.⁴⁸

⁴⁴ Elgin (2017: 37).

⁴⁵ See Elgin (2017: chapter 1).

⁴⁶ *Ibid.* 9.

⁴⁷ *Ibid.* 46.

⁴⁸ A distinct approach to these objections would appeal to an idea familiar to philosophers of science, namely verisimilitude. This idea, roughly, is that theories or propositions can become ‘more truthlike’ even though, strictly speaking, they are false. As has been noted elsewhere (e.g. de Regt 2018) there are unanswered questions as to how Elgin’s own view intersects with extant defences—such as Popper’s—of verisimilitude. Indeed, at some points, Elgin’s view (appealing to ‘true enough’ falsehoods) might naturally be categorised as a

Elgin pushes two distinct lines of thought in order to undermine **Factivity**. The first of these—regarding idealisation in science—has been dealt with extensively elsewhere so I only flag it briefly here. The second argument—regarding improved understanding afforded by falsehoods—has received less attention and requires a different treatment than the first of Elgin’s arguments.

The idealisation argument

One of Elgin’s arguments against **Factivity** appeals to the widespread use of idealisations in science. Idealisation occurs when scientists devise models which simplify and abstract away from what they describe in order to help us theorise about and understand a target phenomena. A standard example of idealisation is the fittingly named *ideal* gas law; it describes the behaviour of gases composed of molecules that lack extension and do not exhibit intermolecular attraction. Such gases cannot exist. So, the thought goes, the ideal gas law is actually false. Nonetheless, as Elgin points out, the ideal gas law is at the heart of contemporary thermodynamics—it is not an aim of science to replace this law because of its idealised nature, despite the fact that, strictly speaking, it fails to accurately describe gases. Hence, she concludes, “It is simply not the case that the bodies of information that constitute scientific understanding are, or that their ultimate successors can be expected to be, composed of truths, with any residual falsehoods occurring only at the periphery.”⁴⁹ In other words, the central and ineliminable role of idealisations in science supposedly undermines **Factivity**.

This type of argument has received attention elsewhere, in literatures not primarily concerned with understanding. I suggest that these literatures afford ‘factivists’ about understanding satisfactory ways to respond to the apparent problem of idealisations. For instance, idealisations have been used as a central motivation for antirealism in the philosophy of science. However, a standard realist response is to suggest that while the ideal gas law might be strictly speaking false it is nonetheless ‘approximately true’.⁵⁰ So, although $\langle p \rangle$ may be a false proposition, $\langle \text{approximately } p \rangle$ is a true proposition. Hence, by taking this line, a friend of **Factivity** can suggest that any understanding of thermodynamics conferred by the ideal gas law consists in full-blooded truths of the form $\langle \text{approximately } p^1; p^2; \dots p^n \rangle$ rather than in falsehoods. This thought can be bolstered by the fact that every scientist is aware that the ideal gas law only approximates the behaviour of real gases, and hence, is only approximately true.

verisimilitudinarian view. I will not be concerned with verisimilitude here, for taking this approach does not allow a genuine defence of **Factivity**. As stated earlier, **Factivity** requires that the commitments comprising one’s understanding be true not merely increasingly truthlike. Truth, at least when applied to propositions, is not a property that comes in degrees. Hence, to defend **Factivity** we must take on the more ambitious project of showing that the propositions comprising an agent’s scientific understanding in the relevant cases are true rather than merely increasingly truthlike.

⁴⁹ Elgin (2017: 62).

⁵⁰ For discussion, see, e.g., Psillos (1999); Bird (2007a), and Sullivan and Khalifa (*forthcoming*) provide further discussion of arguments from idealisation in general.

There is certainly more work that could be done in evaluating this response, but my concern here is simply to note that it has long been considered a viable response to Elgin's first argument against factive views of understanding. My real interest is in outlining a new response to Elgin's second argument against **Factivity**.

The upwards trajectory argument

Elgin's second argument against **Factivity** starts from an idea that we have already noted—that understanding is a gradable notion, it comes in degrees rather than all at once. Elgin uses this platitude about understanding in order to generate trouble for the thought that only true propositions can comprise one's understanding. Specifically, she claims that it leaves us unable to account for cases in which we endorse falsehoods but thereby move upwards along the trajectory from ignorance towards a sophisticated understanding. This thought is illustrated with the following simple example:

An eight-year-old's understanding of human evolution might include as a central strand the proposition that human beings descended from apes. A more sophisticated understanding has it that human beings and the other great apes descended from a common hominid ancestor who was not, strictly speaking, an ape. The child's opinion displays some grasp of evolution. It is clearly cognitively better than the belief that humans evolved from butterflies. But it is not strictly true. [...] [T]he pattern exhibited in this case is endemic to scientific education. We typically begin with rough characterizations that properly orient us toward the phenomena, and then refine the characterizations as our understanding of the science advances. [Elgin 2017: 59]

It is, as Elgin suggests, very natural to suppose that the child acquires a better understanding when they come to believe the false proposition that humans descended from apes. Although learning such a proposition might not be as good as *knowing* that we descended from a common hominid ancestor, the falsehood nonetheless contributes to our understanding of evolution. Hence, it seems, factive accounts of understanding are unable to explain this improvement.

6.2 Against Elgin against Factivity

We should note that the defender of **Factivity** cannot simply appeal to the same 'approximation' idea used to respond to the argument from idealisation. Or, at very least, that type of response will not work in all cases. I want to bypass any local debate as to whether it is approximately true that humans are descended from apes—it seems arguable either way—because other examples more clearly elude the approximation strategy. To see this, I begin by flagging a different example discussed by Elgin namely that scientists improved their understanding of the cosmos by adopting a series of false yet comparatively better cosmological theories: from Copernicus' heliocentric model on which the earth has a circular orbit around the sun; to Kepler's model on which the orbit is elliptical; to Newton's model that accounted for gravitational attraction and posited an irregular elliptical orbit; and so on to the present day. Now, one might still suppose that each of these theories better approximates the truth than the previous. However, this response cannot plausibly be maintained for all predecessor theories. Consider a pre-Copernican transition; the adoption of a geocentric model compared to (let's say)

mythical cosmologies that explained the movement of the celestial bodies in terms of divine activity. The adoption of a geocentric model might reasonably be thought to have constituted progress on the trajectory from ignorance towards a sophisticated understanding of the cosmos. Yet, I take it, we cannot plausibly say that it is even approximately true that the sun revolves around a static earth. So far, this is good news for Elgin's argument.

Granting this, I now propose a different way to defuse Elgin's argument and preserve **Factivity**. We can start with her example regarding our understanding of evolution. The false proposition that Elgin says brings about the improvement in understanding is:

(**F**) Human beings are descended from apes.

Let's immediately grant that child believes proposition (**F**) and concede that they have indeed increased their understanding of evolution in virtue of the learning experience that brought about their belief in (**F**). In order to respond to Elgin's argument, I need to make some points about the nature of belief.

Although it may be possible to identify individual propositions which are believed by an agent, it is not natural to think that an agent can have an individual belief in the sense of having a *solitary* belief. A helpful quote from Braddon-Mitchell and Chalmers illustrates what I mean:

What would it be like to believe that there's milk in the refrigerator, and nothing else? It seems as impossible as having money without the social and economic circumstances that give sense to something being money. To believe that there is milk in the refrigerator, you have to have enough by way of belief to count as understanding what milk is, what a refrigerator is, and what it is for one thing to be inside another. It takes a lot of belief to be any amount of belief. [Braddon-Mitchell & Chalmers 2007: 196]

So, the fact that we believe a given proposition (at least usually, but probably always) requires of us that we also believe a whole range of other things. Indeed, if you lacked any beliefs about the nature of milk, refrigerators, or the relation 'being inside', then we probably could not attribute you with the belief <there is milk in the refrigerator>. Of course, it is difficult (if not impossible) to present a list of contents that we *must* believe in order to believe some given proposition. There are different explanations for this difficulty. One follows from growing scepticism amongst philosophers and psychologists about the 'classical' theory of concepts on which concept-possession requires categorising things in accordance with some discoverable set of necessary and sufficient conditions. A second would stem from the truth of even a modest version of semantic externalism. As Burge (1979) argues, it seems possible to have beliefs about some phenomenon—in his example arthritis—while lacking (or even having misconceptions about) certain core aspects of that phenomenon. Despite these knotty issues, the following idea remains extremely plausible: when somebody believes (or comes to believe) a proposition such as <human beings are descended from apes>, they will also believe (or come to believe) a range of related contents. This much, I hope, is uncontroversial. With this in mind, consider that prime examples of the other contents typically implicated by believing (**F**) are the following:

- (T1) Humans have non-human ancestors.
- (T2) Humans descended from hairy creatures that were sometimes bipedal.
- (T3) Humans and apes have a lot in common.
- (T4) Humans were not put on earth in their current form by a higher power.

That anyone who genuinely believes (F) will also typically believe most, if not all, of (T1) – (T4) is straightforward enough. For instance, it will naturally follow from popular theories of belief that derive beliefs from an agent's dispositions, or use belief-attributions to interpret how an agent behaves. Why is this important? Well, the thing to notice about this list of beliefs is that they are all true. Or, at least, they are true as suitably qualified generalisations and approximations—unlike simply false propositions such as 'human beings are descended from apes'. Of course, it may be possible for someone to lack some of these true beliefs and still count as believing (F). For instance, we might imagine a child who fails to believe (T1) because they are somehow confused about what it is for something to count as a different species. This child might, let us suppose, think that *all* of the bipeds in our ancestral tree were a type of human. Thus, they might believe that humans are descended from apes but not that we are a different species. But, for the most part, people who believe that humans are descended from apes will also believe things like humans are descended from non-human hairy creatures that sometimes walk on two legs.

What does this mean for our analysis of Elgin's rejection of **Factivity**? Well, it seems to provide a way for the factivist about understanding to respond to her argument about improved understanding in virtue of learning a falsehood. Specifically, it allows us to say that what attributions of understanding are tracking in the case she describes are actually the true contents of our belief; namely, our belief in things such as (T1) – (T4). Not only is this a potential response, some of our discussion actually suggests that it is very plausible. Consider again the observation that it might be possible for someone to believe that humans were descended from apes but not believe that humans have non-human ancestors because they have the misconception that humans and apes are the same species. I suggest that we would attribute this person a *worse* understanding of evolution than someone else who believed all of (T1) – (T4), despite the fact that both would count as believing (F).⁵¹ This supports the factivist thought that understanding, to the extent that we attribute it, is sensitive to what we truly believe. It is also worth noting that this style of response seems to generalise to capture cases where the approximation argument fails, such as the example of adopting a geocentric model of the cosmos to replace a mythological

⁵¹ Moreover, this thought explains why the truth of a *radical* semantic externalism would be no difficulty for my view. Suppose, to use a Burgean example, that we can have thoughts or make assertions using the concept ORANGUTAN while wrongly thinking that orangutans are a type of fruit drink. Clearly, in such a case, we would attribute the agent with an impoverished understanding of what orangutans are compared to a regular person who knows that they are a type of ape. As my view predicts: attributions of understanding track the true contents of our beliefs, even if the truth of a radical version of semantic externalism would allow us to possess concepts we have systematic misconceptions about.

conception. For instance, in coming to believe the falsehood that the sun revolves around the earth, early astronomers also became to believe the following sorts of contents:

(T5) Many celestial bodies have regular orbits.

(T6) Many celestial bodies move according to mechanistic laws.

(T7) Many celestial bodies are not moved by divine willing.

Again, the factivist about understanding can claim that, to the extent we credit (say) Ptolemaic astronomers with an improved understanding of the cosmos, it was in virtue of coming to learn these truths. Moreover, I suggest, to the extent that we discovered that such astronomers failed to believe things such as (T5) – (T7), we would accordingly attribute them a worse understanding than we might otherwise.

6.3 Judicious falsehoods and education

Although the upwards trajectory argument does not, I have suggested, support a non-factive conception of understanding we should still acknowledge an important insight from Elgin—namely that falsehoods often play a valuable role in scientific and indeed other types of education. I want to close by suggesting that we can do justice to this insight within a fundamentally veritistic conception of epistemology.

A key reason that falsehoods can be useful in the context of education is that directing someone's attention to a judiciously selected falsehood can often be a simple way to thereby convey a body of true information. This educational strategy draws on a very general feature of human interaction. For instance, imagine that someone is making small talk with a stranger at a bar in the United Kingdom. The stranger asks the person 'Where you are from?' and they reply that they are from Helsinki. Let's suppose that this is false and that they are actually from Espoo, a nearby city in southern Finland. Nonetheless, there may be good epistemic and pro-social reasons for them to state the false proposition that they are from Helsinki—they know that it will convey a variety of germane 'headline' facts about themselves, such as: that they are Finnish; that they are from Northern Europe; that they are not from Estonia; that they are used to cold winters; and so forth. Moreover, were they to truly assert that they are from Espoo, their interlocutor would likely not know where this was and would therefore acquire less in the way of true mental content. Hence, not only might asserting the truth invite unnecessary confusion, it would also be *less informative* than asserting a judiciously chosen falsehood.

This general thought applies equally in educational contexts—sometimes in order to teach a body of headline truths that are easy to digest, it might be most effective to do so by means of focusing your students' attention on a false proposition that conveys these truths. Focusing right away on the truth (e.g. that humans are descended from early hominids rather than apes) might invite questions and confusion that is epistemically deleterious given the introductory purposes of one's lesson. However, this is all consistent with the veritistic principle that Elgin denied: namely that the transmission of truth

is still the central (and only) epistemic aim. In education, we only focus on falsehoods as a means to conveying true information and our students only increase their understanding in virtue of learning things that are true.

7. Concluding remarks

This essay outlined a promising framework for investigating the nature of understanding, namely reductionist theories which claim that understanding is a composite state that is metaphysically reducible to other more familiar epistemic states. I then situated reductionism in relation to other prominent views, and developed a way to test different variants of reductionism. We found that a recently influential version of reductionism, on which understanding reduces to a body of knowledge, was subject to a routine counterexample. However, the knowledge based view was but one version of reductionism. I believe that the basic reductionist insight pressed by proponents of the knowledge based view is a good one. In this vein, I argued that a more defensible version of reductionism focuses on the reduction of knowledge to a body of true beliefs. My view claimed that what matters for ascriptions of understanding is the true content of an agent's body of belief—so long as we hold fixed the true content of what an agent believes, I suggest, they will possess an equal degree of understanding regardless of whether any individual proposition in their belief-set amounts to knowledge or not. As such this view accepts that while much of what we truly believe will invariably count as knowledge, we should avoid attributing a lesser degree of understanding to an agent just because certain propositions in their body of belief happen not to be known. Moreover, I argued that the truth based view can plausibly account for the gradability of understanding and I defended it against well-known objections to the idea that understanding is a factive epistemic state. In short, I hope to have motivated the idea that a truth based reductionism of understanding is a view well worth taking seriously in future work. An interesting project in this vein would be to now apply this truth based view to specific debates in which the state of understanding features prominently—such as the relationship between understanding on the one hand and moral worth and testimony on the other.

Essay Two

Collective Inquiry & Intellectual Progress

0. Introduction

In recent decades, an area of particular growth in epistemic theorising has been the rise of *social* epistemology. This research programme bifurcates into different strains. One research programme highlights the relevance of social factors for the epistemic lives of *individual agents*; prominently, by considering the role of testimony in underpinning much of what we know and believe, and by considering the importance of interpersonal disagreement for what it is rational for individuals to believe.⁵² Although both testimony and the evidence we receive from the disagreement of our peers comprise our social environment, their significance in this research programme is tied only to their influence on the epistemic perspective of individuals. A second research programme goes beyond examining the epistemic lives of individuals and instead theorises about the epistemic lives of *collectives*. This project seeks to understand what underpins the fact that we routinely talk as if groups and collectives are themselves the subjects of epistemic states such as knowledge or belief in their own right. It is within the latter of these research programmes that this essay falls.

Despite the fact that social epistemologists have focused heavily on the conditions required for groups to possess epistemic states such as belief and knowledge—and how ascriptions of these states relate to the epistemic states possessed by the group’s members—there has been much less said by social epistemologists about the processes that lead to groups possessing these well-theorised epistemic states. For just as individuals form beliefs and acquire knowledge as a result engaging in a process of inquiry, it is plausible that by the same token groups often come to possess beliefs and knowledge as the result of engaging in a process of *collective inquiry*. This essay attempts to explore this conceptual space by demonstrating that thinking about collective inquiry is a philosophically fruitful endeavour. I will do so by using ideas associated with inquiry to provide a new perspective on a particular debate: whether and how there can be intellectual progress in domains such as science and philosophy. In brief, I will argue that simply appealing to the presence or absence of epistemic states such as knowledge fails to capture the different facets of intellectual progress that can be made by an intellectual community. Rather, so I will claim, a much more nuanced approach is facilitated by viewing progress as relative to the success of collective inquiry. In developing this account, we will also come to better appreciate some broad continuities that hold between the epistemology of how individuals and groups engage in inquiry.

⁵² See Lackey and Sosa (2006) for an influential collection on the former; see Feldman and Warfield (2010) for essays concerning the latter.

1. Progress: Foundational Issues

1.1 Extant Work

Many of us care not only about the success of our own intellectual endeavours, but also those of the discipline to which we belong. This concern is the correct attitude to have. After all, public discourse about the value and significance of academic pursuits are often connected to the macro-level success or otherwise of these pursuits, typically referring to collectives such as groups of researchers or to institutions such as universities. Given this, it is discomfiting to note that a degree of pessimism about progress within the philosophical community has gained traction lately (see van Inwagen 2004; Dietrich 2011; Chalmers 2015 for representative broadly pessimistic views and Stoljar 2017 for recent optimism).⁵³ In contrast, even though philosophers of science disagree on how to characterise it, there is a widely shared view both within and outwith academic circles that our scientific endeavours are progressing handsomely.

While there has been surprisingly little overlap between discussions of philosophical and scientific progress, a leading strategy in each debate takes acquiring *factive epistemic states* to be constitutive of progress. Recent work in philosophy of science put this commitment front and centre, simply arguing over which epistemic state best explicates progress—the epistemic view focuses on knowledge (see Bird 2007a), the noetic view focuses on a factive conception of understanding (see Dellsén 2016), and the semantic view focuses on truth or verisimilitude (see Rowbottom 2010 or Niiniluoto 2014). A similar focus on factive epistemic states is found within discussions of philosophical progress: for example, Stoljar (2017: 22) focuses on producing knowledge answering philosophical questions, and Chalmers (2015) looks for “large collective convergence” on true philosophical positions.⁵⁴ In what follows I work within this general factive framework, even though it is certainly (and rightly) controversial.⁵⁵ For instance, most parties to the present debate take understanding to be a broadly factive epistemic state but, as we have discussed in the previous essay, some have questioned this assumption (e.g. see Elgin 2004; 2017). More radically, there is a rich tradition of thinking about progress in both philosophy and science which is “anti-realist”, which on some views will entail pushing against explicating progress in terms of truth.⁵⁶ Although this essay will not engage with the project of

⁵³ Comments in Horwich (2012) also express sympathy with pessimism; see Slezak (2018) for recent criticism of Stoljar; Gutting (2016) argues to a more ambivalent conclusion, but nonetheless draws a contrast between philosophical and scientific progress.

⁵⁴ Also see comments in Horwich (2012: 24) and Dietrich (2011: 335-6).

⁵⁵ One interesting issue that I will bypass here is what differences there might be in the standards required for achieving certain epistemic states in different disciplines. Interested readers might consult Jennifer Nado (2017) who provides discussion of the idea that intellectual disciplines such as philosophy and science have their own *proprietary standards* for epistemic states such as knowledge.

⁵⁶ For instance, this will encompass views that construe progress in terms of solving (or dissolving) problems. Within philosophy of science we might look to Kuhn (1962) or Laudan (1977) as examples. Within philosophy,

attempting to refute those who approach progress from within a non-factive framework⁵⁷, much of this essay will be relevant to such theorists so long as the acquisition of *beliefs* figures in one's conception of progress.⁵⁸

A focus on acquiring factive epistemic states is deeply woven into the literature on progress, but prevailing approaches leave two key questions unanswered. Firstly, we need to know what entity is the proper *subject* of the epistemic states purportedly constitutive of progress.⁵⁹ After all, epistemic states such as knowledge can be possessed by both individuals and by groups—so should we hold that progress is a matter of individuals or collectives gaining knowledge? This question has ramifications for substantive verdicts about progress: the conditions for an individual versus a group possessing knowledge are by no means identical. And secondly, the focus on states such as knowledge leaves us unable to say much of substance about *intermediate cases*: cases in which an epistemic subject is on an upwards trajectory away from ignorance, and as such has made some partial progress, but has not yet fulfilled the demanding conditions required to possess knowledge. Such intermediate cases occur routinely. For example, they occur when some true theory has a modicum of support within an intellectual community but is subject to considerable disagreement. A lack of clarity about intermediate cases in the context of thinking about philosophical progress is particularly unfortunate—on one plausible view, an intermediate state between ignorance and knowledge is precisely where the philosophical community finds itself regarding many philosophical questions.

This essay develops a framework that resolves both of these issues by examining progress through the lens of *collective inquiry*. This framework has two broad components. Firstly, after some necessary groundwork, I argue that our best theory of progress should focus on how *collectives* can settle inquiry by acquiring, or by putting themselves in a position to acquire, factive epistemic states (§2). This approach, I suggest, better accounts for judgements about when an intellectual community has answered a question than alternative frameworks focusing on individual inquiry. Moreover, this discussion will also support the utility of an inflationary approach to social epistemology; an approach allowing that

certain conceptions of Wittgenstein's metaphilosophical project are focused on the dissolution of philosophical puzzles through linguistic analysis.

⁵⁷ Some well-known objections to the factive approach to progress—e.g. that many scientific theories are idealisations, or that progress occurs through adopting false theories—have been addressed both by other proponents of the factive approach (see Bird 2007a for an example), and in Essay One of this thesis.

⁵⁸ Another worry (presented to me informally by various philosophers) is that progress might consist in inventing certain instruments or methods; these cannot obviously be construed in terms of propositional content and, so the objection goes, are not truth-evaluable. As I see it, the invention of tools such as telescopes or techniques such as chromatography were progressive for epistemic reasons: viz. because we gained knowledge of how to use them, of the fact that using them might help answer certain theoretical questions, and because using them in fact did lead us to acquire more knowledge.

⁵⁹ For example, Dellsén's noetic theory focuses on individual scientists; Rowbottom is sympathetic to focusing on output such as textbooks and lectures; while Bird looks to what scientists know collectively. These different starting-points are largely assumed, yet each party takes themselves to be describing the same phenomenon.

collectives can possess epistemic statuses which float free of those possessed by their members. Secondly I show how, by focusing on the process of *inquiry* rather than simply on the presence or absence of factive epistemic states, we can appreciate how successfully answering a question often involves a number of overlapping *stages* (§3). As will become evident, thinking about the different stages of inquiry allows us to account for various intermediate cases of partial progress that traditional accounts are silent on. Finally, in §4, I close by considering the implications of my approach for recent angst about progress in philosophy—I will suggest that focusing on collective inquiry enables us to reasonably reject outright despair, while also providing the methodologically circumspect with the resources with which to articulate different varieties of optimism and pessimism about philosophical progress.

1.2 Assumptions & Desiderata

Before getting into the details of the argument, we need to defend some basic assumptions and desiderata to corral our discussion.

Firstly, judgements about progress are relative to some *end*. Often the relevant end is a self-consciously adopted *aim*: e.g. throughout 1944, the Manhattan project made considerable progress towards building an atomic bomb. But it can also be relative to an external *standard*: e.g. someone might judge that their daughter is making good progress on her times-tables. Such a way of thinking, on which progress is end-relative, implies the value of progress depends on the value of the relevant end. This seems correct. After all, someone might observe that vaccination-sceptics have made considerable progress in recent years—such progress is lamentable because the end at which their endeavours are directed is harmful. We'll have more to say about this as we go on, but we are interested in the distinctively intellectual ends relevant for academic inquiry.⁶⁰ Here, we will work within the dominant tradition that supposes the ultimate end for intellectual disciplines to be something like acquiring broadly factive epistemic state such as true belief, knowledge or understanding.

Secondly, progress is relative to some *subject* being judged against some aim or standard. This is exemplified by the truism that different subjects can progress to different extents towards the same end: I can aim to climb the mountain, the Joneses can aim to climb the same mountain, and we can progress to different extents. As the introduction suggested, specifying the subject of progress is something about which there has not been explicit consensus. This was a gloss on the fact that there are actually two distinct questions that we will want to get clear about: (i) with respect to our precise *explanandum* when talking about progress, and (ii) having clarified the explanandum, with respect to what *epistemic subject*

⁶⁰ Phillip Kitcher's (1993) discussion of the impersonal epistemic ends of science is instructive. On his view, the scientific discipline has aims which are *impersonal* insofar as they do not depend on the goals of any individual researcher, and *epistemic* insofar as they have an intellectual character rather than being directed at practical goals (e.g. advancing science for economic reasons).

provides the best focus for theorising about that explanandum. I'll chiefly be concerned with (ii), but need to first clarify (i). To illustrate, consider that the following is a viable research programme: focus on cases in which an *individual* makes solitary progress on a question and then attempt to pinpoint what this progress consists in. While a worthwhile project, this doesn't seem to be what much of the recent work concerns. For one thing, within debates on philosophical progress, a dominant thought takes *absence of consensus* to be evidence for, or constitutive of, a lack of progress. This naturally suggests that the relevant explanandum is some multi-agent endeavour. Moreover, the literature on scientific progress largely appeals to cases involving the dissemination of theories amongst multiple agents. So, in order to best track current debates, I propose to focus on progress within the *intellectual community* engaged in the relevant academic inquiry. The types of intellectual communities we will be interested in are exemplified by the scientific and philosophical communities, collectives which are individuated with reference to the subject-matter that they are concerned with.

There is a common idea that subject-matters can be identified with sets of questions.⁶¹ Combining this with the idea that there are intellectual communities individuated by subject-matters fits nicely with a focus on collective inquiry, because as we discussed at the start of this thesis, inquiry is directed at questions. Hence, intellectual communities can be characterised by their commitment to certain sorts of inquiry. There are certainly further questions that might be posed regarding this approach—e.g. whether we would be better-off individuating intellectual communities with a finer grain; how exactly we carve up questions into different categories—but I think focusing on intellectual communities in this way is an informative presification of what extant discussions about progress have been concerned with.⁶² With this in mind, we get a very rough gloss on what maximal progress looks like: maximal progress occurs when an intellectual community has settled on the answers to the questions comprising its distinctive subject-matter.⁶³ This may well be an impossible goal, but it does give us a cue about how to think of the normative judgements we make when we talk about measuring intellectual progress. These normative judgements—whether negative or positive—concern how well a given intellectual community is doing against the aim of settling inquiry into the questions comprising its subject-matter.

Having clarified that progress is relative to ends and to subjects, we need some desiderata for a satisfactory account. The first is that scientific and philosophical *regress* is a live possibility. This follows naturally from the fact that intellectual communities can move farther away from achieving

⁶¹ See Friedman (2013a) for some discussion.

⁶² I won't offer an analysis of 'intellectual community', because the terms 'scientific community' and 'philosophical community' are common enough. I presuppose that we roughly grasp their referent, though this is not to say that attempting more precision would not be an interesting project.

⁶³ I am inclined to suppose that questions need not have been entertained by the relevant community to count as within the ambit of that community. In this vein, I think that the philosophical community aims to answer *all* philosophical questions—even if there are some philosophical questions we have not yet formulated. Indeed, as I suggest later in the essay, formulating questions might itself be a form of progress.

their aims. But regress is hardly mentioned in the literature. Therefore, one methodological innovation of this essay is using judgements about regress as an explanatory test for a good theory of progress.⁶⁴ Next, one difference in emphasis between debates about scientific and philosophical progress is the importance accorded to *novelty* and *convergence*. Philosophy of science shows a preoccupation with discovery. Conversely, regarding philosophy, much discussion focus on consensus on the correct views. A comprehensive theory of progress will do justice to the importance of both novelty and of convergence. And finally, progress seems to be a *gradable* notion—intellectual communities can make partial progress (and regress) relative to some intellectual end. More precisely, progress is gradable both with respect to novelty and convergence. Some discoveries contribute more than others, and an intellectual community can converge on the truth to greater or lesser extents. All of these thoughts should be accounted for in the framework we use to think about progress.

2. Progress and collective inquiry

My framework for evaluating the progress of an intellectual community is centred around *collective inquiry*. This idea, as I develop it, will involve taking seriously the thought that intellectual communities can settle inquiry as a collective entity. In this section I will firstly (§2.1) stake out three different approaches that one can take to issues in social epistemology—eliminativist, deflationary, and inflationary approaches—before then (§2.2) arguing that an inflationary approach, focusing on when the intellectual community itself has settled inquiry, will better account for judgements about progress than alternative approaches which merely focus on facts about individual inquirers within that community.

2.1 Three Approaches to Social Epistemology

It has been a matter of considerable controversy whether groups qualify as agents which can possess epistemic states in their own right, what types of groups might qualify if they can, and what sort of relationships would have to hold between the mental states of a group's members for the group itself to possess something like knowledge. Certainly we often speak as if small groups constituted for a specific purpose can have beliefs or knowledge: “the court believes that the witness is a qualified forensic expert” or “the ship's crew know how to navigate this boat safely into port” are common enough locutions. In such cases, it is possible (though perhaps not exactly straightforward) to trace out clear connections between the knowledge or belief we attribute to the group and what is known or believed by individual group members. However, when it comes to making sense of the epistemic ascriptions involved in our discussion of larger and more diffuse groups it is a rather more difficult enterprise to attempt to reduce these ascriptions to what is believed or known by individual members. Nonetheless,

⁶⁴ A straightforward way to incorporate such judgments is to assume symmetry between progress and regress—hence, if an intellectual community progresses towards achieving its epistemic ends in virtue of acquiring x (e.g. knowledge) it will regress in virtue of the loss of x .

as has been widely noted, we do often talk as if large and diffuse groups such as academic communities or even entire societies possess beliefs and knowledge: e.g., consider assertions such as ‘Sweden believes that global warming is a threat’ or ‘geologists know what the centre of the earth is composed of’.

When philosophically engaging with the fact that we speak as if large and diffuse groups can be epistemic subjects in their own right, there are three broad positions that one might take. This way of carving up the territory goes beyond the typical binary distinction between summativist and non-summativist views in an attempt to capture the nuances of recent work in social epistemology, but should be broadly familiar from the recent literature.⁶⁵

Eliminativist Views: The first type of views are *eliminativist* responses that suggest that collectives do not genuinely possess epistemic states at all. Under these views, when we ascribe collective epistemic states to collectives, we are merely engaging in a sort of metaphorical or loose-talk.⁶⁶ Hence, social epistemology need only attend to facts regarding individual epistemic subjects.

Deflationary Views: The second type of views are *deflationary* theories on which groups can genuinely possess epistemic states, but holding that these statuses are entirely grounded in facts about the group’s members. The most straightforward deflationary views say that a group has an epistemic state of a given type only if (a subset of) its members possesses that epistemic state.⁶⁷ However, on other deflationary views, a group can have an epistemic status of one type in virtue of its individual members having attitudes of a different type.⁶⁸ Finally, other deflationary views could appeal not only to facts about what epistemic states are possessed by its members but also counterfactual facts about what states the members of a collective would possess under various

⁶⁵ One clue that the summative versus non-summative distinction is not in great shape is that prominent social epistemologists disagree on how to categorise certain theories—for instance, Bird (2010) and Lackey (2016) each take opposing positions (without intending to make controversial claims) as to whether ‘joint acceptance’ accounts should fall into the summativist or non-summativist category. Summative views roughly correspond to deflationary views in my terminology insofar they treat group epistemic states as grounded in individual group-member states. Something to note is that judgement-aggregation views, in my terms, are deflationary positions even though the existence of an aggregation function does to some extent allow what the group believes to ‘float free’ of what the members believe.

⁶⁶ One author who is sympathetic to such a view with respect to certain issues seems to be Ludwig (2016).

⁶⁷ The simplest examples of such views would be positions such as ‘a group believes that p iff most of its members believe that p ’ (sometimes attributed to Quinton 1976: 19). This also captures various types of judgement-aggregation view that have some function for delivering a collective belief if a given number or privileged set of members hold a certain belief. See List and Pettit (2011) for extensive discussion.

⁶⁸ This captures ‘joint acceptance’ views on which a group can have a belief in virtue of its members accepting a proposition (where acceptance doesn’t entail belief). The following is an influential version of such a theory due to Gilbert (1989: 306–307) and endorsed by Schmitt (1994): “A group G believes that p if and only if the members of G jointly accept that p . The members of G jointly accept that p if and only if it is common knowledge in G that the members of G individually have intentionally and openly ... expressed their willingness jointly to accept that p with the other members of G .”

circumstances.⁶⁹ The common-factor shared by all deflationary views is that the epistemic states possessed by the collective must be grounded in epistemic facts about the members of the collective.

Inflationary Views: The third type of views are *inflationary* views. These accept that groups can possess genuine epistemic states but reject the guiding thought behind deflationary views, namely that collective epistemic states must be grounded in epistemic facts about the collectives' members.⁷⁰ In this sense, inflationary views suppose that groups are epistemic subjects in their own right which can possess epistemic states that float free of the epistemic states possessed by their members.⁷¹

These positions are general approaches that can be used to analyse various issues in social epistemology such as the nature and possibility collective belief, collective knowledge, collective justification, and so forth. My intention is to show that an inflationary approach to social epistemology is best suited to theorising about progress. This will involve denying the thought that the intellectual progress made by an intellectual community is just a metaphorical way of referring to progress made by individual members, and the thought that collectives make intellectual progress solely in virtue of their individual members making progress. Rather, I think that collectives can make intellectual progress in their own right.

2.2 Progress & Collectively Settling Inquiry

As we outlined earlier, the dominant approach to progress has focused on factive epistemic states without reaching consensus on the appropriate epistemic subject of progress. Here, I will argue that our considered judgements about intellectual progress are best captured by focusing on the relevant *intellectual community as an epistemic subject* in its own right, rather than merely looking to facts about its members. Moreover, I will demonstrate that the best way to integrate the widespread recognition of the importance of factive epistemic states in accounting for progress is by noting that *acquiring factive states is a way of settling inquiry* into a given question. In this sense, we can settle inquiry into some Q

⁶⁹ For instance, Lackey (2016) on group justification seems to be an example: she suggests that a group can only be justified in believing that *p* only if a proportion of relevant members have a justified belief that *p*, but her account also includes further conditions about counterfactual scenarios involving the disclosure of additional evidence.

⁷⁰ For instance, see Bird (2010) on group knowledge. When mapping different positions of social epistemology, he appeals to the idea of supervenience—in his terms, inflationary views would say that the epistemic states possessed by a group do not necessarily *supervene* on those possessed by the individual members of a group. I've pitched the distinction in terms of grounding but it is not my intention to enter into the difficulties in distinguishing these metaphysical approaches.

⁷¹ A quote from Bird (2010: 24-25) illustrates the spirit of inflationary views: "Although we cannot dispense with individuals altogether, in the sense that societies cannot exist without them, they and their mental states play a minimal role in my account of social epistemic subjects, according to which we should focus on certain social structures as performing functions analogous to the functions of individual cognitive faculties."

by believing, knowing, or understanding an answer to Q.⁷² So, the overall view of progress will focus on how intellectual communities qua group settle inquiry into the questions comprising their distinct subject-matter.

Focusing on settling inquiry will have three advantages. Firstly, it allows us to stay neutral on whether the epistemic, semantic or noetic account of progress is best—adjudicating between these different theories is a delicate topic that properly lies beyond the scope of this essay. Secondly, focusing on collective inquiry will neatly integrate points I make later on (in §3) about how moving through the different stages of inquiry is also progressive—it is progressive because it brings us closer towards settling inquiry. Thirdly and finally, focusing on settling inquiry will enable us to accommodate diverging views of a controversial type of case recently discussed by Alexander Bird and Jennifer Lackey. For ease of discussion I'll largely focus on settling inquiry by acquiring knowledge in what follows, but I intend to encompass other factive epistemic states too.⁷³

Let's start first by providing reason to reject an eliminativist approach on which we can capture judgements about progress simply by focusing on facts about individuals without any need to consider the existence of a broader epistemic group.

The most important motivation for rejecting eliminativist approaches and instead focusing on collective inquiry is that judgements about intellectual progress in a given academic community are ineliminably sensitive to the epistemic position of the community as a whole. Taking progress to be relative to the success of collective inquiry, such as by the acquisition of group knowledge, allows us to explain why this is so. To illustrate, we can warm up with some instructive judgements. Consider a straightforward case: it wouldn't amount to scientific progress if a scientist, today, produced a new set of observational data establishing that mercury boils at around 356°C, and nor would pursuing such observational data be a suitable research programme for a professional scientist. Moreover, the failure of such research to be progressive doesn't change if there happened to be a very small number of

⁷² Of course, there is a sense in which we can 'settle' inquiry with a false belief. However, here we are concerned with progress towards *correctly* answering questions, so our use of 'settling' inquiry will be normative insofar as it focuses on true answers.

⁷³ By this, I mean that my defence of a focus on collective inquiry will apply equally to theories of progress that tie it to the acquisition of true belief (the semantic view) or understanding (the noetic view). In this vein, if one denies that collective knowledge is the ultimate aim, it would be easy to reframe my comments in terms of the presence or absence of collective true belief or collective understanding in settling or failing to settle collective inquiry. It is, however, worth noting that very little has been said about the possibility or nature of collective understanding. The reductionist framework about understanding defended in Essay One suggests a natural approach: collective understanding reduces to (sufficiently well-connected) true beliefs and knowledge that is possessed by the relevant group. If understanding is reducible in this way, there is no principled reason why a group cannot possess understanding in the same way as an individual agent. And indeed, locutions involving the attribution of collective understanding are common enough. However, it is not my intention here to offer a theory of collective understanding beyond indicating that might can appeal to a similar reductionist approach as when theorising about individual understanding.

scientists with false beliefs about the boiling-point of mercury. And finally, I suggest, it wouldn't amount to scientific *progress* if a colleague walked into the office of the ignorant scientists and informed them of this fact, thereby transmitting knowledge, and nor would it amount to scientific *regress* if a single scientist permanently forgot the boiling-point of mercury, thus losing their knowledge. These simple cases suggest a broader lesson: our general judgements about progress cannot be easily explained just by looking at whether progress has been made by a given member of an intellectual community. Rather, we must look at the broader position of the community itself.

It is obviously a non-starter to suppose that *every* member of the group must have some belief for an intellectual community to progress. Nor does it require *most* members to have certain beliefs, as many academic advances are niche—they are only known to a small subset of the relevant community with the inclination and expertise to know them. For instance, in recent decades modern biology has afforded us with a much greater understanding of the highland midge; but very few scientists, even biologists, know much about midges. In this sense, progress is compatible with most of the relevant community lacking knowledge. It is also telling that judgements about progress don't necessarily march in lockstep with the *proportion* of group members with certain knowledge. If a biologist who studies termites expands her range and gets up to date on the latest work on midges, there doesn't seem to have been scientific progress just in virtue of her learning the pre-existing consensus amongst midge researchers. Nor does it seem to matter much, from the perspective of the community, if a research group acquiring knowledge of some new phenomenon contains n or $n+1$ members. Rather, it seems like the scientific community can definitively settle question notwithstanding the presence or absence of additional individual knowers.

All of these judgements are neatly explained by focusing on whether or not the community has settled inquiry by acquiring group knowledge. For instance, regarding our initial simple case, the scientific community has already settled inquiry by knowing the correct and complete answer to the question <at what temperature does mercury boil?>⁷⁴, explaining why further research into this question would be inappropriate and non-progressive.

Moreover, on many views, a group can know that p even if a minority of group members are ignorant about p ; this explains the range of cases in which we do not judge there to have been progress (or regress) just if one individual learns (or forgets) that p , where this change is irrelevant to the overall epistemic standing of the group. Indeed, there is a compelling line of thought on which a group can know or believe something even without most of its members possessing the item of knowledge. For instance, we might suppose that when there is division of labour amongst a group's members, a group

⁷⁴ Some other questions admit of *partial* answers. For instance, the question <what weather is typical in Stornoway?> admits distinct propositions as answers: 'wet' and 'windy' both being correct partial answers. We'll discuss examples of such questions later on.

can sometimes know that p just in virtue of those engaged in the relevant task knowing that p . The intuition behind this position is purely general, not just restricted to intellectual communities. For instance, we might suppose that the Police Dept. knows who committed the crime—indeed, we would say that they had *solved the crime*—even if only the detectives working on the case (rather than the hundreds of other employees) know who committed the crime. This supposition, conjoined with a focus on collective inquiry, provides an explanation for why we do not require any particular *number* or *proportion* of a group's members to have certain beliefs or knowledge in order for the larger group to make progress. Of course, accepting the claim that our focus should be on collective inquiry is not to suggest in any way what individual group-members know is somehow unimportant. Rather, the suggestion is that we can vindicate the importance of success for individual academic inquirers (e.g. when a scientist acquires knowledge) by thinking about how this contributes to the success of the collective inquiry of their wider intellectual community.

So far, the preceding insights about collective inquiry could be accommodated by purely deflationary approaches to theorising about intellectual progress. On deflationary views, facts about the epistemic lives of groups are explained by looking at facts about the epistemic lives of individual group-members. In many ways this is an attractive family of views—it is natural to suppose that, for example, what a group believes or knows will be determined by what its members believe or know. There are different ways to attempt to pursue a deflationary approach in our context. The simplest approaches would hold that a group has collectively settled inquiry into some question Q (e.g. by acquiring knowledge) only if all or most of its members have successfully inquired into Q . This approach would clearly be inadequate for our purposes—we have just observed the fact that groups seem to be able to successfully inquire on the basis of delegating the task of inquiring to particular members. Hence, a more promising approach will allow that a group can settle inquiry in virtue if, for example, its privileged members come to know the answer to particular questions.

I will now suggest that such views will not suffice to explain all of our judgements about progress. In rejecting these views, I will motivate the advantages of taking an inflationary approach to social epistemology when theorising about progress.

An initial argument against deflationary views, particularly those that appeal to what is believed by privileged members of the group, is that we can readily devise scenarios in which there is a fundamental disconnect between what is believed by individual group members and the wider group of which they are a part. Here's a simple case, showing that successful inquiry on the part of individual members cannot itself be *sufficient* for the wider group to make intellectual progress.

TRAGIC RESEARCH: Three logicians openly begin work on a research project. The project comes to fruition and they prove some novel theorems. Before sending their work for review, they decide to take a hot-air balloon trip over Tuscany in order to celebrate their

success. Tragically, the balloon explodes and no-one survives. Worse still, in thrall to a toxic work culture, each of the logicians had taken their laptops containing their research along with them. Their research is destroyed, and nobody ever learns anything about it.

Firstly, it seems that the philosophical community does not know the proofs of the novel theorems despite the fact that these proofs are known by a privileged class of members who inquire on behalf of the group. Secondly, it seems like the philosophical community has not made maximal progress towards proving the relevant proofs despite the success of the individual group members. Successfully completing some project of inquiry even while being a privileged member of some community is not sufficient for the community of which you are a member to settle inquiry if what you have learned is not somehow in a position to be transmitted to, or at least be endorsed by, the wider community.

At this juncture, one might ask whether there is a ready fix for deflationary views in order to deal with cases where a group member's private beliefs make no contribution to collective inquiry. For instance, one might restrict attributions of collective knowledge to cases in which the privileged individual group member's belief is common knowledge—or, at least, to cases in which it is common knowledge that the privileged member *has* a belief that pertains to the relevant question. Although this local fix or others like it may deliver intuitive verdicts in a broader range of cases, I do not think that such strategies involving a restriction on the class of individual beliefs that are relevant for collective epistemic attributions can succeed in general. This is because there are further cases that provide reason to think that the success of collective inquiry is not dependent on there currently being any individual member who has successfully inquired into a given question. We can illustrate this with a case inspired by one discussed, to a slightly different purpose, by Alexander Bird (2010a: 32):

ARCHIVED RESEARCH: Professor Plum is a respected and conscientious researcher currently working on the biology of *Culicoides impunctatus*—the humble Highland Midge. As a topic on which there has been only a little scientific work, there are many rudimentary questions that remain unanswered. Plum is attempting to work out how far on average the male midge will range in search of food. She performs a series of studies that unequivocally support a certain answer: n kilometres. The resulting paper is sent to a respected and well-indexed journal and published after the normal process of peer-review at $t1$. A few years later, at $t2$, Plum has died and everyone who read the paper when it appeared has either died or forgotten about it. Then, another biologist needs to find out about the behaviour of the male midge. They find Plum's paper after a quick search and quickly digest its contents, before citing it in their own paper at $t3$.

Although *Archived Research* is a simplified case, its structural features are not uncommon. The massive distribution of labour in contemporary academic communities, along with the sheer volume of

results being published, make it is natural to suppose that many findings might not currently be known by any particular researcher. So, what should we say about such cases?

Firstly, by stipulation, Plum made a novel discovery and made it accessible to the scientific community to be used in future research.⁷⁵ Without yet taking a stance on the epistemic state of the group, I suggest that Plum's discovery constituted progress for the scientific community *at the time of publication*. The contentious question is: what happened at t_2 ? I suggest that whatever progress the community made *persisted* throughout the period of Plum's death until the work was taken up again at t_3 . To bolster this thought, consider a competing framework looking only at individual epistemic states. If progress were only sensitive to individual epistemic states, then we ought to judge there to have been *regress* when Plum died. However, this doesn't seem right—even though no individual believed the content at t_2 , the wider community was in just as good a place as it was beforehand, still able to locate and use the research when needed. All of these data-points are captured by a focus on collective inquiry, just as long as we suppose that the community has successfully settled inquiry into the Q <what is the range of the male midge?> throughout the entire period. An inflationary approach is vindicated by the fact that epistemic position of the intellectual community qua group has been improved, even though no individual is doing any better epistemically speaking than before the publication of the research.⁷⁶

Some theorists, notably Alexander Bird (2010a: 32), controversially vindicate the epistemic improvement in the group's epistemic position by supposing that an intellectual community can retain *knowledge* of results that have been peer-reviewed and appropriately indexed, despite no longer being known or even believed by any individual group-member.⁷⁷ In this sense, Bird suggests that the way in which intellectual communities endorse and promulgate information amounts to a functional analogue of individual knowledge, making it the case that the group itself is the proper subject of knowledge in its own right. Others disagree. Lackey (2014) has argued against ascribing collective knowledge in such cases by drawing on plausible norms of licensing assertion and action that knowledge is often thought to fulfil. For instance, suppose that the question being inquired into was not one of midge biology but rather concerned the safety of some drug. Would it be rational to prescribe

⁷⁵ However, we should avoid the error of thinking that research must figure in future research in order for it to count as progressive—this leaves us unable to account for progress at the end of inquiry in some area (e.g. midge biology) when the final piece of research has been completed.

⁷⁶ Here's one objection: *Archived Research* depends on the knowledge of the original researcher and so doesn't show that individual knowledge isn't necessary for progress, only that it needn't *remain* known. This objection, tacitly relying upon an asymmetrical account of progress and regress, cannot be sustained. For, there might be cases where a trusted and reliable researcher lacks knowledge themselves but still produces high-quality output. Such examples might include cases where someone *accepts* rather than believes the results of her research, or where—unbeknownst to the wider community—they have lost knowledge due to the presence of some epistemic *defeater*.

⁷⁷ For other potential sympathisers, see Rolin (2008) or de Ridder (2014). Wray (2007) defends a more modest variant that only allows for such knowledge in smaller research groups.

the drug simply on the basis that the research vindicating its safety was in the archives although not actually believed or known by any individual group-member? We might suppose not. Rather, Lackey suggests that cases such as *Archived Research* only put the group in a position to know.

There is an important sense in which my focus on collective inquiry accommodates both positions regarding these type of case. This is because there is a perfectly ordinary sense in which inquiry can be successfully completed just by being in a position to know something. For example, suppose that you are cooking an ambitious dinner for a pair of critical in-laws and want to have a reserve option in case things go wrong. You decide you will need to find out the answer to the Q <what is the phone-number of a very good Chinese restaurant?> for emergency takeaway. You do some research, find out which restaurant has the most promising reviews, and you write down the phone-number on a slip of paper. Then, naturally enough, you forget the number—you no longer know what the number is. Nonetheless, I suggest, you have successfully completed inquiry because you have put yourself in a position to know the answer to Q should you need to know the answer in the future. I suggest that an intellectual community can collectively settle inquiry in the same way—by acquiring compelling sources of evidence that provide the answer to a question within the ambit of that community’s intellectual aims, and enabling the answer to that question to feed back into future inquiry whenever it is needed.⁷⁸

However, even if we decide to take Lackey’s approach here and focus on the epistemic status of being in a position to know rather than attributing knowledge, an inflationary approach is still indispensable. This is because an intellectual community has settled inquiry only when it puts itself in a position to know, in virtue of group-level activity. It is not enough that the members, even the privileged members, are individually in a position to know. For it may be the case that these members have put themselves in a position to know the answer to some question without recourse to resources possessed by the wider group. For instance, individual members of a group may be in a position to know something because of private research they have conducted. In such a case, we would not suppose that the group had (yet) settled inquiry—and nor would we suppose that the group had maximally progressed on answering the relevant question. In order to see whether we have managed to *settle* inquiry into a question, we must look at the epistemic position of the group itself—in particular, we must ask what information has the intellectual community endorsed and promulgated, in addition to asking about the members of the relevant community. Moreover, nor can cases such as *Archived Research* be easily be dealt with by deflationary views involving counterfactual facts about the members of a collective. For, in that case, not only are the actual members of the collective unaware of Plum’s conclusions, they are also unaware of the evidence for her conclusions—so, as a result, the actual members of the collective

⁷⁸ It is worth noting that even if one supposes that cases such as *Archived Research* do not *settle* inquiry for the group, but rather just improve the epistemic position of the group somehow, the framework I develop in the next section will still account for this. For, I will show how collective evidence-acquisition is an important stage on the trajectory from ignorance towards settling collective inquiry.

would not have the disposition to accept the content of her conclusions were they presented to them.⁷⁹ In this sense, collectives can settle inquiry into some Q without their doing so being entirely reducible to facts about the epistemic position of their members.

Before moving on, it is worth briefly noting that cases such as *Archived Research* are not by necessity confined to the fringes of academic research. To see this, we can consider recent discussions of the possibility—and tentative evidence supporting the actuality—of *automated* scientific research. Partly motivated by the fact that the volume of scientific research being published is quickly eclipsing our ability to digest it effectively, there has been growing discussion in philosophy of science about automating aspects of scientific inquiry—from data collection, hypothesis-generation, even to making novel discoveries previously unknown to the scientific community.⁸⁰ Indeed, Spangler et al. (2014) provide empirically supported evidence of data-mining artificial intelligence having discovered previously unknown enzymes that can potentiate an important tumour-suppressing protein. I suggest that the successful operation of automated systems could, in principle, constitute progressive episodes of scientific research.⁸¹ While we are certainly some way from this stage at present, we can entertain an eventuality where the outputs of such AI research comes to have an excellent track-record of reliability and enjoys a high degree of confidence from the community. In such scenarios, it seems wrongheaded to rule out a priori that the operation of such systems could constitute collective progress despite the fact that no factive epistemic state is currently instantiated by any particular group member. This further motivates the thought that how far a collective has advanced towards settling inquiry need not depend solely on what is in the heads of its members at a given time.

Let's take stock. I have begun to motivate the idea that theorising about progress should focus on the success or otherwise of collective inquiry. In this section we focused on cases at the end of inquiry into some question, and argued that many plausible judgements about progress are relative to whether or not the relevant intellectual community has collectively settled inquiry. In so doing, we found that

⁷⁹ Of course, perhaps suitably idealised and well-informed individuals may have such a disposition; but then we would not be grounding the epistemic states of the collective in epistemic facts about the *actual* members of said collective. It is also not plausible to suppose that academic communities naturally accept that *whatever* is published in a journal will stand as their collective view—rather, it depends on the content, to what extent that content hews to existing consensus, and various other factors. (Of course, the plausibility of accepting content merely on the basis of it going through certain procedures such as peer-review for certain journals might differ with the intellectual community in question. It seems much less plausible in philosophy than in some areas of science, for example.)

⁸⁰ This type of research is also discussed in a social epistemology context by Bird (2010: 35). Other examples of—empirically supported—discussion of automated science include King et al. (2004; 2009).

⁸¹ You might wonder whether we could really acquire understanding from such a mechanism. However, if we follow Dellsén (2016) in identifying the core of understanding relevant for scientific progress as ‘correctly explain[ing] or predict[ing] more aspects of the natural world’, then it seems that we could acquire understanding from such mechanisms. After all, predicting which enzymes potentiate which proteins seems as valuable as many routine scientific discoveries.

an inflationary approach to social epistemology—on which collectives can genuinely possess epistemic states and where the possession-conditions for these cannot exclusively be reduced to the epistemic states possessed by collectives’ members—was best suited for our approach.

3. Between ignorance and knowledge

While the focus on settling collective inquiry via the acquisition of group knowledge is plausible at the upper echelons of intellectual progress, such an approach is impoverished in a very significant sense. Namely, it leaves us unable to say much about *intermediate cases* that intellectual communities invariably find themselves in. Intermediate cases occur when a group has intuitively made progress towards answering some question, yet lacks any factive epistemic state towards the answer. I’ll firstly illustrate such cases, pinpointing why they are problematic for extant approaches to progress that focus merely on factive epistemic states. Then, adapting insights from work on the nature of inquiry and the attitudes associated with it, I show how we can vindicate the progressive nature of intermediate cases while still focusing on the success of collective inquiry.

3.1. Intermediate Cases

To make our discussion more vivid, consider the following historical example:

CONTROVERSY. At $t1$, the consensus view is that celestial objects other than earth are perfectly spherical crystalline bodies. Galileo then publishes *Siderius Nuncius* in 1610 detailing telescopic evidence suggesting that the lunar terrain is mountainous and crater-pocked. This evidence is deeply controversial at $t2$, splitting the intellectual community and prompting vigorous debate. Numerous intellectuals reject Galileo’s claims, arguing that the apparent lunar imperfections were either projected there by his telescopic equipment, or that they were actually encased *within* a perfect exterior. Centuries later, at $t3$, Galileo’s view has triumphed and the Aristotelian paradigm in astronomy is universally rejected.

Firstly, we can ask: does the intellectual community in question *believe* that the lunar terrain is mountainous and crater-pocked at $t2$?⁸² Clearly, it does not. On deflationary views, a group belief is a function of individual belief—for example, a group believes that p if most of its members believe that p , if the group has privileged members (e.g. those delegated to inquire) who believe that p , or we can apply some plausible aggregation function on what individual members believe that yields an overall group judgement that p (e.g. List and Pettit 2011). None of these are the case here; by stipulation there is no majority for the Galilean view, Galileo and his followers held no privileged position, and no plausible function for judgement-aggregation (in a non-dictatorial group) would spit out a group belief

⁸² Drake (1978) discusses the reception of *Siderius Nuncius* in the weeks and months following publication. I use a historical example for illustration: if one has reservations about whether there really was a cohesive intellectual community in 1610, an analogous modern example works just as well.

in p in this case. Of course, on inflationary views, groups can have beliefs without these being grounded in what the individual group members believe. However, it is still nonetheless implausible for an inflationary epistemology to attribute p as a collective belief when (at least) half of a group's members actively disbelieves p and are presently engaged in debate as to whether p is a credible view or not. Moreover, as we mentioned earlier (footnote 79), it is not plausible to suppose that an intellectual community accepts as a collective belief *any* content that is produced by its members—this is especially apparent in the foregoing case given that the *Siderius* was not the product of any truth-filtering mechanism (e.g. peer-review) before publication.

Secondly, despite the dissensus and the collective lack of any factive epistemic state, scientific progress had been made by the relevant community at t_2 . As the publication of the *Siderius* is one of the most celebrated episodes of scientific history, disputing this would be deeply revisionary. And such revisionism is difficult to motivate; it doesn't seem plausible to insist that we must wait until there is *outright* consensus on a theory before we judge that *some* progress has been made by the relevant community. For example, there are currently theories in theoretical physics that are subject of deep disagreement. While reaching full consensus on whichever of these theories turns out to be the most accurate would doubtless constitute progress, it is deeply counterintuitive to suggest we have made no progress whatsoever on answering the relevant questions. We can support the suggestion that progress is compatible with substantial disagreement by thinking about *regress*. Imagine that some correct scientific orthodoxy falls into controversy such that it can no longer be ascribed as a group belief. While it seems like there has been *some* regress, it doesn't seem right to say that there has been *complete* regress; that is, regress comparable to point at which the theory was not even under consideration by the relevant intellectual community. One lesson, then, is the following: consensus *is* important for judgements about progress and regress, but it is more plausible to take the influence of agreement and disagreement to come in degrees rather than as a binary enabler or disabler of progress (for instance, being determinative only when a community acquires or loses a collective belief).⁸³ This also coheres with the platitude we observed at the outset of the essay; namely that progress is gradable.

These observations create a tension. We'd established that focusing on when an intellectual community had collectively settled inquiry constituted a powerful framework for theorising about progress, both tracking and precisifying extant work that focused on factive epistemic states, and avoiding the pitfalls of focusing only on individual epistemic states. However, if *Controversy* is a progressive episode in scientific history, it is not one that can be explicated by any of the following:

⁸³ There is another relevant type of gradability here. Knowing that the lunar surface is imperfect is only a *partial* answer to the question <what is the lunar surface like?>. It was further progress when we settled on a more complete answer such as, upon sampling the lunar maria (so-called because some mistook them for seas), finding out that they are primarily composed of basalt. This respects the fact that progress is gradable not merely on the axis of consensus, but also in the *scope* of discoveries. Knowing a more complete answer to a question is more progressive than knowing a less complete answer.

group true belief, group knowledge or group understanding. For, it is platitudinous that: for a group to truly believe p , it must believe p . And, given minimal and eminently plausible assumptions: (i) for a group to know that p , it has to believe that p , and (ii) for a group to understand why p , it has to believe that p or at least believe some q that explains p .⁸⁴ In *Controversy*, at best only some individuals within the group, rather than the group itself, possesses these epistemic states regarding Galilean astronomy.⁸⁵ And more generally, if the influence of agreement and disagreement comes in degrees, it is not easy to capture this gradability by focusing only on binary, belief-entailing epistemic states that are not apt to be taken as degreed notions. This is the problem of intermediate cases: where an intellectual community has progressed away from ignorance, yet has not fulfilled the conditions to possess a group epistemic state like belief, knowledge or understanding.⁸⁶

3.2. Progress without belief

Simply looking at whether or not a group possesses knowledge turned out to be a blunt instrument for measuring progress, eliding incremental advances on the trajectory from being ignorant to knowing. I propose that we can remedy this lack of nuance by looking at *other inquiry-related processes, states and attitudes* which are distinct from outright belief. In particular, we can extend work on individual inquiry in order to provide a richer framework focusing on collective inquiry that can help us theorise about different types of collective progress.

Firstly though, we should pause to consider a natural thought: can we deal with cases such as *Controversy* simply by appealing to the familiar notion of *justification*? In this vein, one might suppose that progress occurred here in virtue of the group acquiring justification (whether in an outright or gradable sense) to believe in Galileo's astrology. My primary reservation with supposing that an appeal to justification can do all the work we need is that it will lack of the nuances of the account I will develop below. In this sense, the proof will be in the explanatory pudding only once I outline my account. Nonetheless, we can explicitly raise one worry for a justification based approach now. Clearly, as the group fails to believe in Galilean astronomy, any appeal to justification to explain progress would

⁸⁴ To the extent that one might doubt these assumptions (e.g. by thinking that one can know on the basis of acceptance, or understand without an outright belief) these worries are irrelevant unless they can be converted into a positive argument for attributing collective knowledge or understanding in the *Controversy* case.

⁸⁵ Here's an objection defused: intermediate cases can be explained by attributing some *different* group belief other than in the content C of the correct theory. For example, one might appeal to the belief/knowledge that <the evidence for C undermines the orthodoxy>, or <the evidence for C needs explained>, etc. However, this response does not capture all relevant cases. For, in the Galilean case, the community was bitterly divided even in their *estimation* of the evidence: many didn't even believe that C should be taken seriously. If such cases are conceivable, then the progress cannot be explained by any belief or knowledge; if there's a split on whether research should be taken seriously, there isn't collective belief that it should be taken seriously.

⁸⁶ Although I didn't provide an analysis of epistemic statuses such as 'being in a position to know' earlier on, it should be clear that *Controversy* is nothing like the *Archived Research* case about which I claimed that the group was in a position to know. In the latter, the relevant information (i) had been endorsed and promulgated by the group, and (ii) was apt to be accepted by the members of the group. Neither of these is the case in *Controversy*.

have to concern propositional rather than doxastic justification. However, it seems that this focus is unable to capture the way in which *convergence* amongst group members of an intellectual community is a form of progress in itself. For, take a group that already has propositional justification to believe p . It seems progressive if this group becomes more confident in p over time, even if this happens without that group acquiring *more* justification for p . This intuition does justice to the preoccupation with convergence in the literature. But it cannot be captured by a view that just focuses on whether a group has or lacks propositional justification for some position; groups can become more or less confident (thus making progress) without gaining or losing propositional justification for the view in question. With this in mind, we can now turn to the more nuanced inquiry-based approach.

Recall the basic features of Friedman’s account of inquiry. Under her account, “a subject inquiring at t has an Interrogative Attitude at t .”⁸⁷ (Note that Friedman uses ‘attitude’ as a loose catch-all term; nothing really hangs on it if—as I sometimes will below—you prefer to talk in terms of states or processes rather than attitudes). *Interrogative* attitudes are a range of different attitudes that we can hold toward some question, Q . These IAs include states such as wondering, curiosity, contemplation, deliberation, suspension of belief, and so forth. The terminus of inquiry is typically (but not always) a settled belief in an answer to Q . In many instances of inquiry, we increase our credence in a given proposition until we are suitably confident as to move into the ‘settled’ state of having an outright belief. However, it is also possible to continue increasing one’s confidence beyond that required for belief, and to stop inquiring before achieving an outright belief. Adopting these different interrogative attitudes—becoming curious about some phenomenon, wondering about a question, suspending belief until more evidence is acquired, contemplating the answer, deliberating over alternatives, and increasing one’s credence on answers to the question that is the object of inquiry—are essential components of inquiring.

While Friedman’s work mainly focuses on individual agents, it can naturally be extended to *collective* inquiry. For, it is natural to suppose that groups can be in a state of: wondering about some Q , deliberating about answers to Q , and becoming increasingly confident in the answer to Q .⁸⁸ Certainly, it is a common enough to hear people ascribe these states to groups—it is perfectly felicitous to say (for example) that the scientific community is inquiring into some issue, wondering about the answer to some question, deliberating over a body of evidence, and becoming more confident in some theory. My interest will be in showing that focusing on these states and processes enables us to do a better job of explicating progress than thinking only in terms of epistemic states like knowledge, justification, or outright belief.

⁸⁷ Friedman (2019: 4).

⁸⁸ It might be the case that certain of the states Friedman discusses, particularly curiosity, have a phenomenological component not easily realised by a group. As such, I won’t be concerned with these states here.

Consider *Controversy* again. Instructively, it is precisely transitioning into and between different interrogative states that occurs in this case. The question we are judging progress relative to is: <what is the lunar surface like?>. While there was no collective belief in a correct partial answer to that question, viz. that the lunar surface is imperfect and pock-marked, the intellectual community seemed to have progressed towards settling on the right answer just in virtue of considering Galileo’s new view. Thus, vexingly for our promising framework for thinking about progress, settling inquiry by acquiring group knowledge was not necessary for partial progress. It is here that we can appeal to interrogative states and transitions for a more nuanced account of how intellectual communities make progress. More pedantically, considering such attitudes suggests sufficient conditions for partial progress towards answering a relevant question. The progress found in *Controversy* at *t2* compared to *t1*, so I argue, was at least threefold: (i) the intellectual community transitioned from ignorance to having interrogative attitudes towards a significant question it had not yet correctly answered, (ii) the intellectual community acquired evidence supporting the answer to this question, and (iii) the intellectual community assigned higher credence to the correct answer to this question, an answer which equated to a fundamental fact about the nature of the cosmos. I suggest that each of these achievements are individually sufficient for an intellectual community to make some progress towards settling inquiry into a question. Moreover, focusing on these inquiry-related processes respects the importance of novelty and convergence in a comprehensive theory of progress. Adopting an interrogative attitude towards a question is a prerequisite for discovering *novel* truths, while increasing collective credence represents progress towards *consensus* within the intellectual community; acquiring evidence is important for both of these processes.

With this sketch in mind, I’ll now separate out and briefly characterise three different processes that roughly correspond to three different stages of collective inquiry; transitioning into these stages represents three different facets of collective progress that does not require the acquisition of any factive epistemic state.

Wondering: A group moves into the state of wondering about Q by formulating a question and adopting some commitment to inquiring into Q. While wondering inevitably occurs alongside other interrogative attitudes, it typically begins prior to them—agents wonder about a question before (for example) deliberating over answers. Indeed, it is the fact that we are wondering about Q that usually explains why we are deliberating over answers to Q. Moving from a state of ignorance about some significant Q—whether it be ignorance in the form of the absence of an answer, or a belief in a false answer⁸⁹—to a state of wondering about Q constitutes partial progress towards answering Q. One reason for this is a fact that we observed earlier—the ultimate aim of an intellectual community is to answer all of the

⁸⁹ I qualify that we must move from *ignorance* to wondering because it is not (always) progressive to *know* the answer to Q and then start wondering about it all over again. However there may be cases in which resuming wondering despite getting it right is progressive, e.g. when counterevidence against a view is discovered.

questions properly within the ambit of their distinctive subject-matter, even when the group has not yet conceived of these questions. Identifying and then wondering about these questions is the first progressive stage on the trajectory to settling inquiry into these questions. Not only is identifying and formulating Q the first step towards answering Q, wondering about new and more fine-grained questions is an iterative process that is crucial for successful inquiry.⁹⁰ In particular, devising insightful research questions is particularly important for the acquisition of *novel* knowledge. For instance, once we affirmatively answered the question <did Neanderthals control fire?>, it was further progress to then wonder <whether Neanderthals created fire or merely harnessed naturally occurring sources of fire?>. A group, like our own scientific community, that was able to articulate this question, was doing better than the hypothetical group that was unable to form another relevant question. This is because wondering was the first step towards answering this question, beginning a process terminating in valuable new knowledge about our ancestors. Wondering about new questions is also important for answering *pre-existing* questions a group is already inquiring into: for instance, there are many cases where a group realises that, in order to answer some Q, it will be necessary to firstly answer Q*. For example, it might be necessary to know whether Neanderthals migrated to some area A (where evidence of purposive fire-making has been found) in order to know whether Neanderthals had the capacity to start fires. Until an intellectual community has *fully* answered all of the significant questions pertaining to its subject-matter, it does well by being in a state of ever more advanced wondering about them. Moreover, to press the point about the comparative explanatory power of a justification based account, note that wondering does not consist in the acquisition of propositional justification for some answer. Nor can it be explicated by appealing to the notion of evidence. Wondering about a new and important question is a distinct form of progress in its own right. (Note that these forms of progress are thoroughly collective in nature—while I won't belabour this further, brief reflection shows that it would not be sufficient for progress simply for an individual member of a collective to move between these inquiry-related processes. For instance, an individual scientist starting to wonder about an important new question is not progressive if their having this interrogative attitude in no way feeds into the larger collective epistemic standing of the group.)

Investigating and Evidence-gathering: One interrogative process mentioned by Friedman is that of contemplation. A second, not mentioned by Friedman, but naturally taken to be a cognate, is investigation. One plausible way to think about contemplating Q is something like entertaining Q in thought, while investigating Q, roughly, can be thought of as seeking information relevant to answering Q. As I am conceiving of it here, contemplation can be a way of investigating a question, and is often how we investigate abstract philosophical questions. But many questions (including some philosophers seek answers to) demand purposive activity such as experimentation and building on prior research.

⁹⁰ The extent of progress might be sensitive to the type of question being asked; some questions are plausibly more worthwhile than others. Indeed, some questions might not be worth answering at all.

However, *merely* investigating Q (whether via contemplation or otherwise) is not necessarily progressive; an individual or group might investigate some Q ineptly, learning nothing. Rather, investigation is important because (if things go well) it can yield *evidence* that facilitates answering Q. Although the exact conditions for evidence to come within the epistemic ken of a group has been little discussed it is nonetheless perfectly commonplace to suppose that (e.g.) the scientific community can acquire evidence. It isn't possible to settle knotty debates about the nature of evidence here, but we can provide a very rough gloss on the conditions for collective evidence-acquisition: an intellectual community acquires evidence when propositional content providing evidential support for some theory is made accessible to the relevant members of the community, such that they can use this content in reasoning and inference. This, paradigmatically, occurs when books and articles are published and disseminated within the relevant community.⁹¹ The collective acquisition of evidence that supports the correct answer to Q, I claim, is progressive. Taking evidence-acquisition to be a loci of progress coheres nicely with our focus on group epistemic states like collective knowledge as the terminus of inquiry; gathering evidence that supports a correct answer to Q is progressive because it facilitates the group settling inquiry by *knowing* the answer to Q. This is true regardless of whether (e.g. like Lackey) one prefers a traditional approach or (e.g. like Bird) one prefers a knowledge-first account of collective knowledge. On a traditional approach, a group's evidence provides *justification* that enables their collective belief in an answer to Q to be knowledgeable; on a knowledge-first account, evidence is itself a form of knowledge that supports knowing the answer Q by entailing it. We celebrate historical episodes such as the publication of Galileo's *Siderius*, because Galileo promulgated powerful evidence for a ground-breaking intellectual discovery. Engaging in investigation that bears the fruit of evidence is one way for an intellectual community to progress, without acquiring a settled belief on the answer to whatever Q it is wondering about.

Deliberating and Credence-revision: Another important part of answering a question is deliberating, viz. weighing the evidence for different answers on the way to settling on a belief. Groups like intellectual communities deliberate by directing the relevant group-members to consider the evidence for and against these views. Deliberating itself is not necessarily progressive; deliberation can tend towards the false just as well as the truth. However, when deliberating over answers to Q, groups become more or less confident in different positions.⁹² When a group becomes increasingly confident

⁹¹ Of course, we must be careful to draw a distinction between cases like *Controversy* and cases like *Archived Research*. In both types of case there was evidence acquired by the group. However, in the latter case, the compelling nature of the evidence acquired by the group and the fact that it was not apt to elicit widespread disagreement meant that it was sufficient to *settle* inquiry into the relevant question. This is not the case in *Controversy*; this is an intermediate case of partial progress, because it is implausible to suppose that the relevant community had settled inquiry at *t2*.

⁹² I lack a theory of collective credence. There is a small technical literature attempting to devise formal rules—and encountering problems—for aggregating group credences from individual beliefs and credences (see, e.g. Russell et. al. 2015; Dietrich *forthcoming*). Certainly, how confident individual group members are will play an

in the right answer, this constitutes progress; an essential part of an intellectual community transitioning from having an interrogative attitude towards Q towards settling on (and knowing) the correct answer is increasing the credence assigned to that answer. Progress in intermediate cases is not exhausted by increasing credence in the correct position. This is because wondering about some significant Q and gathering evidence for the correct answer can be progressive even if the group remains agnostic throughout this process. However, due to the fact that persistent disagreement is often a primary roadblock preventing a group transitioning from inquiring to having a settled belief or knowledge, a group becoming more confident in a correct answer to Q represents one of the most important facets of intellectual progress. Appreciating the importance of credence-revision captures the sense in which, as we noted at the beginning of the essay, *convergence* is a crucial component of progress for intellectual communities; groups make progress towards converging on the right answer by becoming more confident in that answer.⁹³ Focusing on collective credence also takes into account the lesson of our earlier discussion; progress relative to consensus does not come all at once only when collective belief is gained—rather, it is sensitive to how confident the intellectual community is in the right answer. And finally, there is a further advantage of taking increased credence in the right answer to be progressive. This advantage is that it allows us to judge as progressive those cases in which a group is confident enough to possess a factive epistemic state such as true belief, but continues inquiring in order to become even more confident.⁹⁴ It is plausible, I think, to suppose that a group can make progress by becoming increasingly confident in the correct answer to a question even when it already confident enough to correctly believe the right answer.

The various interrogative attitudes and processes we have outlined do not involve group belief or knowledge. However, thinking about these other attitudes and processes affords us with a richer appreciation of the different types of progress a group makes on the way to settling a question. This allows us to explain why an intellectual community is doing well by transitioning into an intermediate case between ignorance and knowledge: short of knowing the right answer, a community can make

important role in determining how confident the group is. However, I am also sympathetic to the thought that, just as with collectively settling inquiry, there will be situations favouring an inflationary approach to collective credence. For instance, we might imagine cases in which an intellectual community has a number of published articles on some question Q, most of which support one answer <A> but some of which dissent and support <¬A>. It may, moreover, be the case that few group members have any attitude towards Q. In such cases, it strikes me as plausible to suppose that the group has a confidence in <A> higher than .5 but lacks outright belief. (Of course, this case is under-described—however, it illustrates the sorts of scenarios that might speak in favour of an inflationary approach to collective credence.)

⁹³ I am sympathetic to supposing that groups must increase their confidence in the correct view on the basis of *evidence*. Those who think of progress in terms of mere true belief might not require such a condition; this may be a mark against that view.

⁹⁴ This coheres with an emendation made to Jane Friedman's framework in the Introduction—namely, that it is both possible and sometimes normatively appropriate for an agent to inquire beyond the point at which it has formed belief in an answer to the Q that is the object of inquiry.

progress by formulating the right sorts of question, by gathering evidence, and by weighing up this evidence while becoming increasingly confident in the right answer.⁹⁵

Finally, appealing to interrogative states doesn't only explain *progressive* intermediate cases, it also provides a helpful perspective on different scenarios involving *regress*. For example, focusing on collective credence allows us to explain why it is sufficient for partial regress if a group (at t_2) becomes less confident in a true theory even if, prior to this at t_1 , the theory did not command quite enough support within the group to be attributed as a group belief. This is regress, even though no group knowledge has been lost, and the framework defended here explains why this is so: because the group has lost confidence in the right view. However, not all cases of regress can be given a credence-theoretic treatment; some require that we look to the absence of specific interrogative attitudes. For instance, suppose that a group has formulated an interesting question, started the process of investigating it, but all of the group's members remain avowedly neutral on the answer. It would seem to be regressive if (for, say, sociological reasons) this question became entirely neglected at some later time. But this won't be explicable due to any changing credence in the answer. Rather, regress has occurred because the group has stopped wondering about an important question. Looking at interrogative attitudes not only explains why there can be progress without new belief, but also why there can be regress without the loss of belief.

4. Coda on Philosophical Progress

I would like to close by considering how this general framework impinges upon our thinking about philosophical progress. A number of philosophers have endorsed pessimistic evaluations of philosophical progress, derived from focusing on an apparent lack of consensus on the big philosophical questions within the philosophical community. One upshot of focusing on progressive attitudes that fall short of belief is that it enables us to reject outright pessimism about philosophical progress, but without committing to the strategy (e.g. taken by Stoljar 2017) of claiming that the philosophical community has actually answered many fundamental questions. Instead, we can press the thought that philosophical progress consists in formulating the right sorts of questions, in gathering evidence for correct theories, and in the increasing popularity of correct theories (if there are such in the corpus) within the philosophical community. Each of these forms of progress is consistent with the absence of collective belief in particular theories. Of course, it is not entirely transparent to us whether we have identified the correct theories. Hence, the extent of our progress cannot be entirely transparent to us either. But a lack of transparency is an entirely typical feature of progress; it flows naturally from the difficulty of certain

⁹⁵ A final question: is it progressive for a group to simply jettison, or become less confident, in a false belief? This is unclear. Rejecting false views is typically associated with acquiring evidence, and resuming inquiry into some question. My intuitions are muddy upon considering whether a group makes progress in virtue of replacing a false belief with no belief at all, without acquiring evidence, and without reopening inquiry.

aims and the limitations of the subjects pursuing those aims. For example, if a group is traversing a ridge in foggy conditions it may well be unclear to them—despite their best efforts!—whether or not they are progressing by moving in the right direction.

In a broader sense, the focus on the different aspects of collective inquiry allows us to provide fine-grained loci for different varieties of optimism and pessimism about philosophical progress. For instance, one might endorse optimism about the philosophical community's ability to ask important questions and to gather evidence for theories, but pessimism about our ability to collectively settle on particular views. Such a view, while acknowledging some ways in which our current philosophical practices are vulnerable to criticism, avoids the lack of nuance exhibited by those who loudly assert that there has been no philosophical progress whatsoever. A valuable project for the future will be to try and understand what explains the strengths and weaknesses in how the philosophical community attempts to settle inquiry. One important way to approach this project will be to look at other intellectual communities as comparators; looking at how they transition through the different stages of inquiry, how they gather evidence, and how they use this evidence to collectively settle on theories. The framework outlined in this essay has provided a way to approach these comparisons—this, I hope, constitutes progress of a sort.

5. Conclusion

This essay has looked at intellectual progress through the lens of inquiry, attempting to both get clear on what epistemic entity should be the subject of our judgements about progress, and to find a framework that provides an informative evaluation of intermediate cases that lie on the trajectory from ignorance to knowledge. After incorporating the widespread focus on factive epistemic states in extant literature by focusing on cases in which inquiry into some question is settled, I argued that the best framework for theorising about progress tracks the epistemic standing of intellectual communities *qua group*, rather than on their individual members. Reflecting on celebrated examples of intellectual progress revealed that the dominant approach—taking epistemic states such as knowledge or belief to be exhaustive of progress—failed to do justice to ubiquitous *intermediate cases* of partial progress, where an intellectual community has not yet fulfilled the conditions required to achieve knowledge. A more flexible account of progress that could capture these intermediate cases, I suggested, could be facilitated by paying attention to less heavily theorised attitudes and processes associated with inquiry. Using this approach, we were able to capture some different varieties of progress that intellectual communities can make: formulating significant research questions; gathering evidence; and becoming collectively more confident in the correct answer to these questions, none of which necessarily involve the acquisition of collective belief or knowledge. This allows us to provide a fine-grained account of different varieties of intellectual progress. One upshot of using this more nuanced framework is that we can find a middle path between extreme optimism and extreme pessimism about progress in philosophy.

Essay Three

The Virtue of Curiosity

0. Introduction

Curiosity plays a central role in our intellectual lives. It focuses our attention on the unexplained, sustains our intellectual activity, and its satisfaction rewards us when our epistemic efforts go well. The disposition to be curious also plays an important role in various types of explanation. For example: curiosity is ineliminable from psychological explanations for why agents act in certain ways; curiosity features in evolutionary explanations for the adaptive success of various creatures; and curiosity—especially when excessive or deficient—can feature in normative explanations for why we praise or criticise the conduct of epistemic agents. In spite of these roles, curiosity has been less extensively studied within philosophy, psychology and neuroscience than one might expect. The contemporary philosophical literature is growing but still rather modest, with a great deal of extant work simply mentioning it in passing or discussing curiosity alongside a family of other attitudes and mental states.⁹⁶ On the empirical side research into curiosity has started to undergo a very modest renaissance, but the phenomenon is still rather poorly understood.⁹⁷ This essay attempts to develop our understanding of curiosity by providing an account of why curiosity is important for epistemic theorising. In short, I will argue that curiosity has an important normative role because the trait of curiosity can amount to an intellectual virtue—a virtue which conduces to various types of effective and praiseworthy inquiry.

In more detail, the plan is as follows: firstly, after disambiguating the difference between state and trait curiosity, I outline a working theory of the nature of curiosity (§1); before clarifying different ways in which curiosity can cease and introducing some extant work on whether there are norms governing state curiosity (§2). Then, I begin to outline my positive proposal regarding the normative importance of curiosity by discussing the relationship between the trait of curiosity and virtue theory (§3). §4 fleshes out my approach by arguing that a virtuous disposition to curiosity must at least be discerning in its objects, exacting in its persistence, and timely in its occurrence. I close by discussing how my theory impinges upon two broader debates in virtue epistemology—firstly, regarding what sort of motivational structures are compatible with being intellectually virtuous (§5), and secondly, regarding whether intellectual traits need to reliably secure epistemic goods in order to qualify as virtues (§6).

1. The state of curiosity: a working theory

To immediately head off any confusion, we should begin by observing that attributions of curiosity are ambiguous between two distinct readings. For instance, when we say that ‘*S* is very curious’ we could be referring to either of two facts about *S*. Firstly, we could be attributing to them some sort of curiosity

⁹⁶ Notable exceptions who provide focused a detailed discussion of curiosity include Whitcomb (2010); Carruthers (2018); and Kvanvig (2013). Below we will also discuss Friedman, who includes curiosity in her list of interrogative attitudes. Even more recently, essays in Inan et al (2019) further develop contemporary discussion.

⁹⁷ See Loewenstein (1994) and, more recently, Silvia (2006) for thorough overviews. Carruthers (2018) provides philosophical treatment of recent empirical evidence, and also see Kidd & Hayden (2015) for useful discussion.

trait. This attribution is what is intended when we say things such as ‘George is a curious person’. When we make this attribution we do not mean to say that George is curious about any particular thing but instead mean that he has a liberal capacity or disposition to exhibit curiosity. In this vein we can offer more fine-grained attributions of trait curiosity by for example noting that George tends to be curious about certain subject-matters. Very little has been said by philosophers about the trait of curiosity; it is more often discussed in the psychological literature. Secondly, attributions of curiosity can refer to occurrent curiosity—the *state* we are in when presently experiencing curiosity about something in particular. Thus, we might say ‘Rita is curious about the football score’ when right now she is curious about who won the match. This essay will discuss both state and trait curiosity, but we should begin by sketching a serviceable theory of what the state of curiosity is. The rationale for starting with state-curiosity is that, in the first place, getting clear on what trait-curiosity amounts to will require referring to the state—having a curious character must after all have something to do with the tendency to experience curiosity.

In the existing work on curiosity, some points of agreement have emerged that I will also adopt throughout this essay as follows.

Firstly, it is widely agreed that curiosity is a cognitively basic state found in (at least some) non-human animals. Existing empirical research suggests that it is an *affective* state—borne out by the natural thought that curiosity has desire-like and emotion-like elements—that motivates inquisitive behaviour in all sorts of creatures.⁹⁸ As such, it has been taken as a desiderata that any theory of curiosity is consistent with the fact that it can be experienced by creatures across a spectrum of cognitive sophistication rather than being unique to mature human inquirers.

Considerations of having a cognitively undemanding account of curiosity have led to a second point of consensus. Although there is a common and superficially tempting gloss on which curiosity is a desire for knowledge or truth⁹⁹, the idea that curiosity is a desire with any *de dicto epistemic contents* has been widely rejected in recent work. This is because there is agreement that it is possible to feel curious without possessing any epistemic concepts.¹⁰⁰ For example, even very young children can experience curiosity—indeed young children are often paradigm exemplars of curious inquirers—yet they may lack the cognitive sophistication required to possess epistemic concepts such as KNOWLEDGE. More broadly, appeals to curiosity often feature in explanations for why various animals act in certain ways and animal curiosity is its own area of independent study. Hence, the *de dicto* desire-

⁹⁸ See, in particular, Carruthers (2018) for discussion.

⁹⁹ Often in relation to a traditional rendering of Aristotle on which “all men by nature desire to know.” Kvanvig (2013: 159) discusses interpretative issues with this rendering. Williamson (2000) and Connee (2017) endorse, in passing, similar views.

¹⁰⁰ See Whitcomb (2010), Friedman (2013a) and Carruthers (2018). Carruthers provides particular philosophical defence of the suggestion that curiosity is a basic affective state that is not restricted to mature human inquirers.

for-knowledge view has typically been excluded because it is difficult to attribute epistemic contents to the desires of cognitively unsophisticated creatures who lack epistemic concepts.

Thirdly, there has been increasing consensus that curiosity—in some constitutive sense—aims at questions. Under this general view, whenever we are curious about something there are some question(s) that are the object of our curiosity. Earlier on, we introduced the work on Jane Friedman who has been an influential proponent of this approach by including curiosity among her list of interrogative attitudes which have questions as their contents. Such a view is very flexible. For it is not limited only to straightforward cases in which one is explicitly interested in some question, but it can also be used to model curiosity about objects, subject-matters, events, and so forth. For example, we can say that curiosity about some phenomenon Φ that is not itself a question (e.g. curiosity about the moon) is implicitly directed at very general questions—e.g. <what is the moon like?>—concerning that phenomenon. By appealing to the idea of implicitly being curious about a question, focusing on questions is rendered consistent with attributing curiosity to cognitively unsophisticated creatures. Just as a creature’s belief can have a proposition as its object without the creature having the concept PROPOSITION so curiosity can have questions as its object without the creature having the conceptual ability to represent questions in thought. There are different ways that one might precisify this idea, but here I only rely on the general theoretical framework that has been discussed more comprehensively elsewhere.¹⁰¹ When I use the phrase ‘object of curiosity’ henceforth, I have in mind the questions constituting the instance of curiosity.

2. Ceasing to be curious

Before turning to outline my virtue-based account of why curiosity is important for normative epistemic theorising, I will firstly disambiguate two different ways in we can cease to be curious. This will enable us to refer to these phenomena when developing my own account, and also allow us to distinguish some extant discussions about whether there might be norms regarding when the state of curiosity should cease. I will end this section by taking a close look at one of these discussions—on the idea that the *function* of curiosity is knowledge-acquisition—and instead argue that the function of curiosity is better understood as motivating inquiry. Then, with this in mind, I turn to outline my account of when this motivational trait qualifies as an intellectual virtue.

2.1 Curiosity Extinction

Most discussion of when we stop being curious focuses on those cases in which the acquisition of some epistemic state causes curiosity to cease. However, it is a little-discussed truism that we often stop being

¹⁰¹ In particular, see Friedman (2013a) and Carruthers (2018) for both conceptual and empirical arguments for this view. Whitcomb (2010) also endorses this focus on questions. Inan (2012) gets into the weeds of this commitment, raising some complications for such views.

curious without acquiring *any* epistemic state. Simply put, creatures sometimes stop being curious without learning anything new. Thus, from the perspective of psychological observation, there are different ways in which curiosity can cease. We can call the phenomenon of ceasing to be curious without acquiring any new epistemic state *curiosity-extinction*. Curiosity-extinction is a very common phenomenon. Here are two examples:

Phone: A woman is curious why the neighbour's children are causing such a ruckus. Her phone rings and she starts talking to a colleague. She never learns what the commotion was about, forgetting about it entirely.

Loggers: A monkey is curious about a loud noise and goes to investigate. It spots a team of loggers moving through the forest and bolts to the safety of its shelter. The monkey falls asleep. Upon waking, it no longer shows any inclination to investigate.

In these routine cases, the subject stops being curious without changing their epistemic position towards the question constituting their curiosity.¹⁰² We can offer the following preliminary gloss on curiosity-extinction:

Curiosity-extinction: One's curiosity about Φ is extinguished when it terminates without any change in one's epistemic position towards Φ .

As the cases above illustrate, curiosity can be extinguished by various non-epistemic factors and without leading to any new belief or knowledge about the object of one's curiosity.

Curiosity-extinction is a common psychological phenomenon. But what should we say, normatively, about such cases? Is there something inappropriate about the extinction of curiosity? This question has not been discussed in extant literature. However, I suggest that there is nothing *necessarily* inappropriate about curiosity extinction. Rather I suggest that, taken in isolation, individual episodes of curiosity function much like a wide-scope view of how the requirements of rationality function—you behave in accordance with your normative obligations *either* by ceasing to be curious *or* by improving your epistemic standing towards the object of curiosity.¹⁰³ Of course, curiosity is involuntary insofar as one does not consciously choose to be or cease to be curious. Instructively, this is also the case with other basic affective states such as hunger or feeling cold. These affective states, hunger and coldness, can cease without you 'satisfying' them by acquiring food or increasing your bodily temperature. This is perfectly acceptable, all else being equal. Although criticism may well be due if, say, failing to eat conflicts with *prudence* (e.g. you are starving), *rationality* (e.g. you intend to win the heavyweight

¹⁰² When I use terms such as 'epistemic position' and 'epistemic states' I stipulatively refer to both doxastic facts such as the set of beliefs (and credences) one has bearing on Φ and their epistemic properties including justification and knowledge. And if Φ is a placeholder for some question, then we are concerned with cases in which there is no change in the set of beliefs one possesses relevant to answering the question constitutive of one's curiosity.

¹⁰³ See Broome (2007) for succinct discussion of the wide-scope view of rationality.

championship), or *morality* (e.g. if you don't eat the cake, grandmother will cry) it is not necessarily incumbent upon you to eat every time you feel a pang of hunger. Moreover, when prudence, rationality or morality make such demands, one has reason to eat *irrespective* of whether one feels hunger or not. While it might be unfortunate if the endogenous system regulating one's hunger is not in harmony with these reasons, this is not by itself cause for criticism. Similarly with curiosity. All else being equal, you are not normatively obligated to actively inquire every time you experience curiosity. This is the case even when it comes to curiosity about the most epistemically valuable projects—extinction of curiosity about some questions only becomes normatively problematic, I suggest, when failing to inquire into such questions amounts to an intellectually vicious character trait. Criticism for failing to improve your epistemic standing following an *individual episode* of curiosity is only due when this conflicts with some other normative obligation you are under.¹⁰⁴

2.2 Curiosity Satiation

Let us now turn to cases in which one's curiosity ceases as a result of changing our epistemic position towards the object of curiosity. We can call when curiosity ends as a result of some epistemic change *satiation*:

Curiosity-satiation: One's curiosity about Φ is sated when it ends due to a change in one's epistemic position towards Φ .

An interesting question concerns whether we might offer a more precise description of this phenomenon, perhaps holding that curiosity-satiation itself is a distinctive attitude, feeling, or mental state? This question has been little discussed. However, we already mentioned two guiding thoughts that might prove instructive: (i) avoid over-intellectualising curiosity, and (ii) provide an account of curiosity applicable across the animal kingdom. Since cognitively unsophisticated creatures can sate their curiosity, these criteria are well-placed in this context as well. These desiderata seem to rule out any higher-order thought being requisite for curiosity-satiation, such as the reflective process of 'taking one's curiosity to be sated'.¹⁰⁵ The lightweight account of satiation above does not rule out the fact that mature humans often do reflect upon whether their curiosity has been sated, but avoids imposing it as a necessary precondition. A more plausible attempt to precisify our characterisation might appeal to some basic affective 'satiation experience' or the functional operation of some regulatory mechanism. However, as the neuropsychology of curiosity is very much in its infancy¹⁰⁶, we do not at present have the knowledge required to further precisify these ideas. As such, we can work with the definition above

¹⁰⁴ A common example might be following: when curiosity causes you to form an *intention* to inquire, then the enkratic principle gives you rational reason to fulfil your intention.

¹⁰⁵ Friedman (2013a) discusses issues surrounding appealing to higher-order thought with respect to the nature of curiosity at length, although her discussion doesn't much deal with curiosity *satiation*.

¹⁰⁶ See Kidd and Hayden (2015) for a selective overview.

while leaving open the possibility that empirical research may allow for a more precise account in future.

So, what sates curiosity? Given that curiosity is an interrogative attitude with an intimate relationship with inquiry, a natural place to start is with the thought that our curiosity into Q will be sated with whatever epistemic change means that inquiry into Q has been settled. So the idea goes, once inquiry into Q has been settled then there is no longer any need for curiosity about Q—there is no need to have an attitude that leads to inquiry when inquiry is already settled. In light of our earlier discussion, this would suggest that curiosity can be sated by an epistemic change amounting to a new belief or increased confidence in an answer to the question one’s curiosity is directed at.

Generally speaking, it does to be true from the psychological perspective that curiosity can be sated by the acquisition of belief. A simple example: you are curious whether there is milk in the fridge; you ask your partner; you believe their answer; then you stop being curious. Moreover, a belief can sate curiosity even if it is false and/or unjustified. For instance, I might be curious about some question, consult an unreliable friend, and receive a false answer.¹⁰⁷ So long as I believe their testimony, my curiosity will be sated. In this sense (assuming knowledge entails truth), neither knowledge, truth, nor justification are necessary for curiosity-satiation.¹⁰⁸ Rather, just simple belief generally does the trick. In other cases, perhaps less frequent, curiosity might be sated with mere increased confidence. As with inquiry, this seems possible both when one already has a belief in the relevant proposition and when one doesn’t. On the former, we can again appeal to entirely felicitous ascriptions such as “I’m curious as to whether she ever received my letter—I believe she did, but I’m not entirely sure.” As with inquiry, it seems that continued curiosity is compatible with any attitude towards a proposition short of certainty.¹⁰⁹ On the latter, we can imagine cases in which one is only mildly curious about some question that might be difficult to answer unequivocally; in such cases increased confidence might suffice to sate one’s curiosity. In summary, curiosity can be sated—i.e. cease following an epistemic change, without anything being amiss from the perspective of the agent—just by acquiring a new (false) belief or by increasing one’s confidence in some proposition.

¹⁰⁷ I take it that such a possibility can be discussed so as to yield a lack of justification regardless of whether one treats justification in an internalist or externalist way. For example, consulting an unreliable friend might amount to an unreliable process on an externalist-reliabilist view of justification. It might also be unjustified on an internalist-evidentialist conception—for instance, if brief reflection would have brought to mind one’s belief that the friend isn’t to be trusted.

¹⁰⁸ As curiosity can be sated by entirely false belief, this also rules out understanding being necessary for satiation under any account on which understanding cannot consist in entirely false beliefs—this includes the factive conception of understanding I have already defended in Essay One, on which understanding is comprised of a body of true beliefs. See Kvanvig (2013) and Brady (2013) who are sympathetic to a connection between curiosity and understanding (but do not claim that it is necessary to sate curiosity in the sense I describe here).

¹⁰⁹ The idea that one can be curious about any fact about which one is not utterly certain is suggested in a little-discussed but important article by Inan (2014).

Turning to the normative dimension of curiosity satiation, some have suggested that curiosity has a special connection with the acquisition of particular epistemic states over others. Knowledge has frequently been identified as playing this role.¹¹⁰ For example: Williamson (2000) states that curiosity brings about a desire for knowledge, while Conee (2017) asserts that curiosity sets an epistemic goal of knowledge. Although these comments appear in fleeting discussions of curiosity, other philosophers provide more detailed accounts. Whitcomb (2010) provides one of the best developed discussions. Observing that, for instance, merely sating hunger with a diet pill does not satisfy one's hunger in the appropriate sense, Whitcomb suggests that nor is merely sating curiosity always sufficient to comply with the norms on ceasing to be curious. Rather, he introduces the normatively-loaded idea of curiosity 'satisfaction' which uniquely demands the acquisition of knowledge:

Curiosity is a desire for knowledge, not in that its contents always involve some concept of knowledge, but instead in that it comes to be *satisfied iff you come to know the answer to the question that is its content*. Curiosity is thus satisfied by knowledge alone, in the same way hunger is satisfied by nourishment alone. In each case there is a state and a desire, such that the desire comes to be satisfied iff you come to be in the state. [Whitcomb 2010: 673]

Whitcomb's is not the only discussion of the norms on curiosity cessation. For instance, Kvanvig (2003) argues that curiosity is satisfied by epistemic justification ('the kind that might be defended as a necessary condition on knowledge') and that this feeds into a broader aim that curiosity end in an increased understanding of the object of one's curiosity. Here, I do not intend to rehearse all of the moves in this debate. Many of the arguments they discuss—invoking lottery-cases, Moorean sentences, and blameworthy patterns of inquiry—have been extensively analysed in other literatures, and the dialectic regarding these is advanced and complex. Rather, I instead discuss and reject a unique argument suggested by Whitcomb that appeals to the *function* of curiosity in order to argue that curiosity uniquely aims at knowledge. I will argue that considerations surrounding the function of curiosity actually strike against *any* epistemic state playing a privileged role. Rather, I argue, the role of curiosity is to motivate inquiry. Then, with this in mind, I turn to outline my own positive account that focuses on the trait of curiosity.

¹¹⁰ In addition, see comments in Carruthers (2018). This claim also has historical pedigree, with versions of it being attributed to Aristotle, Freud, and Augustine. An essay by Engel (2019) contains an interesting historical discussion of philosophical views on curiosity.

2.3 The function of curiosity

A number of philosophers suggest that the function of something provides us with a grip on the norms governing the operation of that thing—whether it be an artefact, trait, faculty, or behaviour. Here’s Peter Graham, an influential exponent of such a ‘teleofunctional’ approach to epistemology:

[F]unctions are norms. Norms are standards. Fulfilling a function meets a standard; it satisfies a norm. Fulfilling a function, and so measuring up to a norm, is a success, fulfilment, or achievement. The heart is supposed to pump blood; that is what it is for; that is its “aim.” When it pumps blood, it does what it is supposed to do; it measures up to the standard set by its function. [Graham 2012: 461]

In this vein, one might suppose that reflecting on the function of curiosity will tell in favour of certain epistemic norms governing when we ought to stop being curious. This question has also been touched upon in the empirical sciences, so is also of independent interest.¹¹¹ Above, we highlighted comments in Whitcomb (2010) that suggested a version of a function argument for the normative priority of knowledge in sating curiosity by drawing an analogy to the function of hunger. Whitcomb claims that the function of hunger is driving an agent to acquire *nutrition*, because only acquiring nutrition—rather than merely sating hunger without nutrition—can have a ‘positive influence on hunger’s broader role in keeping you alive and healthy’. Analogically, he suggests the function of curiosity is to drive us to acquire *knowledge* rather than any lesser epistemic state.¹¹² The best version of this argument does not look at the contribution that some *instance* of knowledge-acquisition makes to wellbeing. For, in predictable cases, acquiring knowledge or nutrition will not contribute to wellbeing; consider the curiosity of the jealous spouse, or the hunger of a compulsive eater. Rather, the argument should be taken akin to ‘teleofunctional’ approaches which derive the function of an item from its explanatory history. For instance, with Graham, we can observe that hearts are *for* pumping blood by observing that the causal explanation for hearts being the way they are is so that they performed the function of pumping blood and this function explains the existence of creatures with hearts now. In this sense, a heart that fails to pump blood is a bad heart with respect to the norms given by its teleofunctional role. So, is curiosity *for* knowledge-acquisition?

A first issue with this claim is, *prima facie*, that it will be difficult to establish that the unique function of curiosity is knowledge rather than another epistemic state such as true belief. After all, as is familiar from recent literature on epistemic value, it is difficult to show why knowledge of the answer to some

¹¹¹ Such an argument is briefly discussed by both Kvanvig (2013) and Whitcomb (2010).

¹¹² Whitcomb (2010: 668).

question is any more *advantageous* for an agent than true belief.¹¹³ I merely note this worry to set it aside.

A second important observation, one that has been pointed out elsewhere (e.g. by Kvanvig 2013), is that psychologists who study curiosity have not coalesced around its function being uniquely bound up with acquiring knowledge.¹¹⁴ Rather, the literature contains various characterisations, often emphasising motivating information-seeking activity. I will argue that we can capture this characterisation by a focus on *inquiry*. With this in mind, my view about the function of curiosity is that the function of curiosity is best understood in terms of motivating *inquiry* rather than acquiring any particular epistemic state such as knowledge.

Here is a first line of defence for this claim. The idea that the function of curiosity is to motivate inquiry attempts to capture the following thought: curiosity fulfils its role by motivating an agent to inquire, but whether or not the agent inquires successfully (e.g. by ending up with knowledge) or unsuccessfully (e.g. by ending up with false belief) is a separate matter. Of course, this is consonant with curiosity having an intimate relationship with the outputs of good inquiry (e.g. knowledge)—but it doesn't mean that the *function* of curiosity is to acquire the outputs of good inquiry. Consider an analogy with a team, where its members have different roles. The role of the goalkeeper in a football match is to prevent the opposition from scoring. Indirectly, this helps the team to advance their design of winning the match. But the goalkeeper can fulfil her function perfectly even if the team fails to win—if, say, the strikers are hopeless at putting the ball in the back of the net. The goalkeeper should only be evaluated against the standard of preventing the opposition from scoring—this standard, rather than that of winning the game, provides the norm of good goalkeeping. In this vein, I claim, curiosity should be judged against the standard of motivating inquiry; whether inquiry goes well or badly is a matter for the agent's other capacities, such as how they gather and evaluate evidence. This dovetails nicely with our reactive attitudes. When an agent has their curiosity sated by an unsupported belief, we do not—I suggest—consider them to be *doubly criticisable* both in virtue of forming an unjustified belief *and* in virtue of sating their curiosity without knowledge. Rather, there is just one error; the agent has not regulated their beliefs in an appropriate way.

Here is the second argument. Whitcomb rightly considers the adaptive role of curiosity in theorising about its function. Supporting its adaptive role, evidence suggests a species' disposition to curiosity is linked to how advantageous inquiry is likely to be for it. For example, studies have suggested that predators, well-armoured animals, and 'generalists' who exploit a range of environments rather than a specific niche, tend to be more curious about novel objects than animals who are comparatively

¹¹³ See Kvanvig (2003) for extended discussion. This problem, of course, stems back to Plato's *Meno* where it was observed that truly believing you are going the right way seems, to all intents and purposes, to be just as good knowing you are going in the right direction.

¹¹⁴ And see references throughout, esp. fn 97 for relevant overviews.

vulnerable or flourish within a very specific environment.¹¹⁵ Curiosity has an essential adaptive role in motivating these creatures to engage in inquiry which will be valuable for them insofar as it predictably leads to the acquisition of adaptively beneficial goods. Most significant of these will be the acquisition of food, shelter, and mating opportunities—but other potentially adaptive outcomes such as enjoyment and socialisation may figure in a complete account of why curiosity would be selected for in the animal kingdom. With these adaptive benefits in mind, we find a further consideration against concluding that the function of curiosity is knowledge-acquisition. Call this the *argument from serendipity*. The basic thought is the following: one adaptive advantage of experiencing curiosity is that animals inquiring from curiosity will *serendipitously* acquire beneficial goods, irrespective of whether they acquire knowledge about whatever their curiosity was directed at. This will occur when curiosity about some Φ motivates inquiry that leads to a windfall acquisition of some good unrelated to Φ . Consider an example. An animal is moving from one location to another and hears an unfamiliar noise at the mouth of a cave. Whether or not the animal is sufficiently curious will determine whether it inquires into the source of the noise. An animal disposed to inquire will, thereby, explore the contents of the cave. Engaging in such exploration, regardless of whether the source of the noise is discovered or not, opens the possibility of acquiring other goods—such as sources of food, shelter, or water.¹¹⁶ So long as such inquiry is unlikely to lead to adaptive penalties (e.g. predation) it is easy to see why curiosity is selected for in such situations; it is better than doing nothing. In this sense, curiosity is valuable because it motivates inquiry in lieu of occurrent hunger or thirst—it facilitates the acquisition of information that might be useful in the future rather than to satisfy an immediate need. All this is so, even if nothing (or nothing useful) is learned about the actual object of curiosity. Of course, this phenomenon occurs at a more complex level in humans too—we need only look to serendipitous discoveries in science that have resulted from inquiring into separate questions. The adaptive value of curiosity is broader than yielding knowledge answering the questions constitutive of one’s curiosity; motivating inquiry alone can suffice to yield the relevant adaptive goods.

From these arguments, we should prefer an account on which the role of curiosity is just to motivate inquiry. While simple, this account suffers no explanatory loss compared to other views involving epistemic states. We can still explain the fact that something goes awry when one sates one’s curiosity with an ill-supported belief by pointing out that such cases involve failing to conform to familiar epistemic norms governing belief. This gives us a unified explanation for our criticism of curious and

¹¹⁵ Glickman and Sroges (1966) provide the classic study of animal curiosity. Byrne (2013) provides some discussion. However, this is an area where there is currently a paucity of research.

¹¹⁶ One might object that finding these things amounts to acquiring knowledge. But recall that curiosity aims at answering questions. Hence, we are interested in knowledge *answering* the questions constitutive of the episode of curiosity. All my argument requires is that any knowledge involved in acquiring goods serendipitously is not an answer to whatever question constitutes the curiosity. For example, knowing <here is water> does not answer <what was that noise?>.

non-curious agents who exhibit similar failings. Furthermore, our account accommodates two additional data-points. Firstly, it allows us to see how serendipitous discovery coheres with the functional role of curiosity. Such discoveries are enabled by the fact that curiosity motivates inquiry, not the fact that it yields any determinate output such as knowledge. And secondly, it allows us to explain a fact we observed earlier—the extinction of curiosity without knowledge is not necessarily inappropriate. It is not obvious how to explain this fact, that one is not automatically due criticism for simply ceasing to be curious, on a view on which the function of curiosity is knowledge-acquisition. However, on the present account, we can simply point out that curiosity can fulfil its role of motivating inquiry without an agent actually engaging in inquiry. Curiosity fulfils its function so long it provides the agent with some motivation to inquire, even if they suppress this urge. The analogy with hunger is instructive here. Hunger fulfils its adaptive function by eliciting a *pro tanto* motivation to find food, even if you suppress this motivation (e.g. because you are dieting). In this sense, affective states can fulfil their motivational function even if an agent decides against performing the actions that such states paradigmatically motivate.

3. Curiosity: state, trait, and virtue?

In this remainder of this essay, I will provide an account of why curiosity is relevant for epistemic evaluation—in brief, I argue that it is relevant for assessing the epistemic character of inquirers because the trait of curiosity can constitute an *intellectual virtue* with an intimate connection to epistemically good inquiry.

Before we begin, some brief theoretical orientation is required. Within virtue epistemology, there are two broad approaches.¹¹⁷ One approach—often called ‘virtue reliabilism’ and exemplified by Greco (2010) or Sosa (2007)—discusses *faculty*-virtues such as the faculty of memory or eyesight. This approach often seeks to understand these faculty-virtues primarily in order to argue that certain epistemic states, paradigmatically knowledge, are a result of appropriately employing our faculties. A second approach—often called ‘virtue responsibilism’ and exemplified by Zagzebski (1996) or Baehr (2011)—seeks to identify and understand *character*-virtues that are the intellectual analogues of familiar virtues (like benevolence or courage) discussed in moral philosophy. Broadly speaking, the latter approach has a greater concern with evaluating agents rather than the aetiology of individual epistemic states. My approach here falls under the responsibilist banner and, furthermore, is not at all concerned with providing a reductive account of knowledge—or any other epistemic state—in terms of intellectual virtue.

¹¹⁷ The extent to which these approaches are complementary or in tension is a matter of debate. See Fleisher (2017) for discussion.

A thriving project in contemporary epistemology concerns identifying and explicating the epistemic virtues within the responsibilist tradition.¹¹⁸ Although there is little sustained argument for this claim, some prominent sources have suggested that curiosity is one such virtue.^{119,120} Here, I outline a theory of the virtue of curiosity—outlining and hazarding an explanation for the ways in which curiosity can exhibit excess and deficiency (§4), investigating what sort of motivations are compatible with virtuous curiosity (§5), and finally by rejecting the need for any reliability condition on the virtue of curiosity (§6).

Earlier, we noted that attributions of curiosity are ambiguous between state and trait curiosity. Put roughly, the former is the occurrent affective experience that one has when presently curious about something, while the latter is a standing disposition or propensity for experiencing curiosity. Although there is not a great deal of consensus on the schematic criteria that must be satisfied for something to count as an epistemic virtue, the responsibilist orthodoxy—drawing on the veritable Aristotelean tradition—agrees that virtues are a class of intellectual character traits: they are *enduring dispositions* (virtues are stable; they cannot be a one-off), they are *cultivated* (in that their possession depends, loosely, on one’s effort and choices over the course of a life), and they are *excellences* (insofar as they are amenable to, and receive, positive normative evaluation). Therefore, we should be focusing on the *trait* of curiosity—individual episodes of experienced state-curiosity cannot qualify as cultivated, dispositional excellences.

So, what sort of standing disposition might the virtue of curiosity amount to? I will be operating with the natural assumption that the virtue of curiosity will, at least, involve excellence with respect to one’s disposition to experience curiosity. That is, I will assume that: (i) the trait of curiosity is possessed by agents with some sort of disposition to be in the state of curiosity, and (ii) if we can identify an *intellectually excellent* disposition to be curious—one that captures and explains our normative judgements about certain good and bad forms of intellectual conduct—then we will have identified a plausible candidate epistemic virtue. Of course, one might adopt an alternative approach that seeks to trace normative judgements about curiosity back to the exercise or absence of other independent epistemic virtues. On such an alternative view, there would be no distinct ‘virtue of curiosity’; rather, our ordinary practice of praising (and dispraising) agents for their (lack of) curiosity could be explained

¹¹⁸ Throughout, I use ‘intellectual’ and ‘epistemic’ virtue interchangeably.

¹¹⁹ For example, some of the most influential monographs on virtue responsibilism suggest that curiosity is a virtue such as Baehr (2011: 19, 21) and Zagzebski (1996: *passim*).

¹²⁰ Baumgarten (2001) treats the *moral* dimensions of curiosity, while Manson (2012) also discusses the moral contours of curiosity in outlining a virtue he calls ‘epistemic restraint’. See Inan et al. (2019) for an edited collection on the moral psychology of curiosity. Other purported virtues that are plausibly related to curiosity, such as ‘inquisitiveness’, have also been subject of some discussion: e.g. Watson (2015). Yigit (2019) also provides some discussion of virtue epistemology and curiosity. See Engel (2019) for a very recent paper on the vice of curiosity.

by citing other intellectual excellences and deficiencies. This is not the approach I take here.¹²¹ Rather, I provide a way to vindicate the idea that curiosity is a virtue by investigating the idea that it involves an excellent disposition to be curious.

Earlier we provided a rough characterisation of occurrent curiosity on which it is directed at questions, motivates inquiry into these questions, and is sated by acquiring putative answers to these questions in the form of belief or increased confidence. With this rough characterisation in mind, it is easy to see why it is tempting to suppose that curiosity has an intimate relationship with epistemic virtue. Curiosity, one might think, can be virtuous when it tends to identify important intellectual questions and motivates us to engage in and complete epistemically valuable projects of inquiry. We often assign value to epistemic projects that do not merely service practical or prudential goals—it is eminently plausible to think that curiosity plays an essential role in prompting these sorts of inquiries that would not otherwise be motivated by different affective states such as hunger or cold. Insofar as the intellectually virtuous agent is someone who is motivated to engage in epistemically valuable projects of inquiry, it is natural to suppose that such a character will, by the same token, be a curious agent. Moreover, in the previous essay, we discussed the way in which wondering about questions is an important progressive step on the way to answering them. For all of these reasons, it is natural to suppose that curiosity is capable of qualifying as an intellectual virtue.

However, this promissory vindication relies on the vague idea that epistemically virtuous curiosity is some sort of disposition to experience curiosity. To make further progress we must clarify what sort of curious disposition should be merited as a virtue.

4. Excessive, deficient & skilful curiosity

A very simple attempt to identify the relationship between the state and virtuous trait of curiosity might start with the following proposal, the failure of which will inform our later discussion:

? Liberal Disposition: The virtue of curiosity is a disposition to frequently experience, and attempt to satisfy, curiosity.

¹²¹ However, I do not reject this approach out of hand. I am starting from plausible assumptions, rather than applying any comprehensive theory of virtue-individuation. It remains open that someone with such a comprehensive theory might reject the suggestion that curiosity is a distinct virtue. Nonetheless, I take it that the burden of proof is with the proponent of such a view to motivate it. Given that we routinely praise agents for their curiosity, several influential theorists already endorse the claim that curiosity is an intellectual virtue, and I am not aware of any already-theorised virtues that can do the fine-grained work of capturing our normative judgements about curiosity better than the account I outline in this essay, there are strong—albeit defeasible—reasons for holding that curiosity is a distinct virtue.

In favour of this proposal is the observation that it accords with a plausible rendering of the relationship between certain other affective states and traits. For instance, it is natural to say that an angry person is just someone with a disposition to frequently experience and be moved to exhibit anger.

However, **Liberal Disposition** clearly fails to constitute an epistemic virtue. Simply having the tendency to experience curiosity is sufficient to exhibit such a disposition. But, other canonical epistemic virtues cannot be attributed *merely* on the basis of a subject having certain experiences. Rather, a subject exhibits virtue only when their response to a situation stems from an ability to appropriately regulate their intellectual conduct and affective responses. Unlike the descriptive observation that a subject is disposed to have certain experiences, virtue-attributions are inescapably *normative*. Consider: on the traditional account derived from Aristotle, virtue lies at a mean between vicious excess and vicious deficiency. Conscientiousness is a paradigmatic epistemic virtue. One cannot, for instance, exhibit *too much* conscientiousness because, as a virtue, conscientiousness lies at the mean between the vicious excess of fussiness and the vicious deficiency of carelessness. This highlights the fatal flaw with **Liberal Disposition**. From the psychological perspective, it seems that curiosity can be directed at an incredibly vast range of questions. In this sense, curiosity is not restricted to questions of any particular type, subject-matter, or importance. Indeed, one can become curious about esoteric questions, banal questions, questions taking different interrogative complements (who, what, where, why, and so forth), and even questions to which we have no conception of what the possible answers might be.¹²² Moreover, it seems that our curiosity about different questions often waxes and wanes in a more or less idiosyncratic way according to the vagaries of our character and mood. It cannot simply be the case that *any* disposition to curiosity qualifies as virtuous. It is easy to conceive of ways in which one might have a liberal disposition to curiosity that: (i) betrays vicious excess or vicious deficiency to the detriment of your intellectual character, or (ii) while not obviously vicious, does not plausibly count as intellectually excellent either.

In order to refine our discussion, we can use common-sense judgements to identify prominent types of intellectual failure that can beset the disposition to be curious, before looking for a unifying story below.

Discernment. Curiosity can aim at different questions. In light of this, consider the following normative observation: it is intellectually better to be more disposed to curiosity about certain questions and less about others.¹²³ So, one can exhibit excess by being disposed to curiosity about

¹²² See Friedman (2013: section 4.2). She suggests that one can be partially ignorant by being unaware of some possible answers to a question but also radically ignorant by lacking a conception of any possible answers. One reason for this is that asking a question requires much less cognitive sophistication than does representing possible answers—for an illustration, consider abstract questions at the higher reaches of science and philosophy.

¹²³ This should not be read as a claim about *degrees* of occurrent curiosity, but rather about one's disposition to be curious about certain things *tout court*. While there may be something to be said for this idea, I do not wish to

epistemically trivial topics: overweening obsession with football transfers does not exhibit epistemic virtue, while the same degree of curiosity regarding a fundamental question of science is grounds for intellectual admiration. Our normative evaluations, in this sense, do not just respond to our disposition to a certain quantity or degree of curiosity—they are also qualitatively sensitive to what we are curious about. And one can exhibit deficiency through being insufficiently disposed to curiosity about significant epistemic topics. While we don't expect that every agent will be curious about every significant area of inquiry, it is implausible to think that an agent can exhibit intellectual excellence if they routinely fail to exhibit any curiosity about nature, the sciences, philosophy, the arts, history, and so forth. Here, such mild epistemic elitism is rather plausible: we typically take an intellectual life filled with wonder about art and science to be more intellectually valuable than one preoccupied with sports and gossip. Of course, this isn't to say that it is vicious to be periodically curious about sports, only that one's overall disposition towards curiosity must typically be oriented towards the right sort of objects. Furthermore, we shouldn't suppose that the appropriate objects of curiosity can be fully settled with reference to topics or subject-matters. One can exhibit a failure by only ever becoming curious about rather basic questions, regardless of what subject-matter one is curious about. For instance, a student who is preoccupied with the dates of significant historical events but never their causes seems to exhibit an epistemic deficiency. Therefore, the virtuous agent (if we assume the view that curiosity aims at questions) will possess the disposition to be curious about certain significant questions rather than others—we can call this the requirement of *discernment*.

Exactness. Even if one reliably shows discernment in becoming curious about the right questions, one can exhibit an intellectual failing by possessing the disposition to cease curiosity prematurely, or to continue to experience curiosity to the point of pathology. One important variety of this failing concerns when our curiosity is sated. Consider that—on any plausible semantics—why-questions generally admit of more or less comprehensive true answers. For instance, a student of 20th century European history might be curious: why did the Weimar Republic undergo hyperinflation in 1918-1924? They might receive the true answer, via reliable testimony, that the Republic adopted imprudent economic policies. However, while true, this answer doesn't afford any insight into the nuances or distinguishing features of this historical episode. Routinely having one's curiosity sated by such information would not yield any deep knowledge or insight into any topic—one's intellectual life would be well suited for trivia, but little else. The corresponding excess here is perhaps less common, occurring when one's curiosity leads them to be preoccupied about small and insignificant details. For instance, there will be some items in a complete causal explanation for some

claim that virtue requires any particular degree of curiosity. For instance, if two agents differ in their degree of curiosity, but whatever degree of curiosity they experience suffices to motivate them to inquire, this seems on a par from the perspective of normative evaluation.

phenomenon that are simply too trivial—in the sense of being uninformative—to spend time being motivated to inquire into. Another related type of failure concerns inappropriate cessation and continuation of curiosity in response to difficulties and obstacles. This most often manifests itself as a deficiency in the form of persistently having one’s curiosity extinguished as the result of being exposed to minor setbacks. The intellectually virtuous person has tenacity and determination, and their curiosity will not evaporate at the slightest indication of difficulty in completing their inquiries. Less commonly, it is possible to exhibit a different type of failing by persisting in the face of truly insurmountable adversity—for instance, one should not be consumed with curiosity about 'whether I can build a perpetual motion machine' over the course of many years, refusing to acknowledge the overwhelming evidence that it will be impossible.¹²⁴ About such questions, it is more appropriate for one’s curiosity to become extinguished. In these senses, virtuous curiosity must be *exacting*: it should motivate the inquirer to overcome reasonable difficulties and attempt to acquire comprehensive answers to significant questions.

Timeliness. Thirdly, there also seems to be a form of excess in being motivated by curiosity in the wrong circumstances. Curiosity exerts a powerful motivating influence, and makes demands upon our cognitive resources when it goes unsated—as such, the virtuous agent will regulate their curiosity accordingly. One controversial form of this excess is curiosity impinging upon non-intellectual matters—for example, encroaching on your valuable personal relationships. Whether this sort of consideration precludes an instance of curiosity from exhibiting *epistemic* virtue depends on whether such excesses can be explained away as *moral* vice, and subsequently, on whether it is legitimate to cleave apart the intellectual and moral virtues into distinct evaluative spheres.¹²⁵ Without settling this controversy, there are clearly cases where curiosity can exhibit vice by motivating inquiry in circumstances that are inappropriate from a purely intellectual perspective. This occurs in the following common scenario: when you are engaged in some valuable intellectual inquiry into *x* but your curiosity about *y* interferes with your inquiry.¹²⁶ Oftentimes this takes the form of banal, recalcitrant curiosity into trivial matters—the latest gossip, football scores, and so forth. But you can also be curious about the right thing in the wrong circumstances. For instance, it is not a virtue if incessant curiosity about novel ideas prevents the academic from ever completing any rigorous research project. Such curiosity would be better being extinguished, or at least felt only to a lesser degree, until a more opportune moment arises. Curiosity in inappropriate circumstances

¹²⁴ Strangely, this example is not an entirely idle worry. Such is the volume of applications, some international patent offices have adopted exceptions (that do not apply to any other type of patent) to discourage attempts to register perpetual motion machines.

¹²⁵ While Aristotle suggests that the intellectual and moral virtues are different in kind, some contemporary philosophers have doubted this claim. For instance, some of Aristotle’s motivations for his view arguably stem from the dubious division of the soul into distinct thinking and feeling parts. See Zagzebski (1996: 137-158) for a thorough discussion and critique of Aristotle’s arguments on this front.

¹²⁶ Some psychologists, such as Berlyne, call the drive for novel stimuli ‘diversive curiosity’.

amounts to a distinct regulative failure—an intellectually virtuous trait will involve the disposition to *timely* curiosity.

In short, we have found that intellectual virtues cannot *merely* be dispositions to have certain experiences. Rather, virtues must be determined by what Aristotle calls an *orthos logos*—a sort of skill in regulating one’s feelings and issuing responses that are appropriate to the circumstances. Just as experienced anger should be regulated by a virtue of the temperament, so we have found that experienced curiosity must also be regulated by some virtuous intellectual disposition. In order to get a better grip on the virtue of curiosity, it is worth asking whether the considerations identified above have provided us with any purchase on whether there is any broader account unifying the regulatory failures identified above?

Here is the sketch of an explanation: it is the role of curiosity to motivate inquiry. A common theme running through the failures identified above is the observation that certain approaches to inquiring are ineffective ways securing of *valuable epistemic goods*. This coheres with the rough and commonplace thought that certain forms of intellectual activity are more epistemically fertile than others. One way to put this is that you can come to understand *more* about the world in virtue of engaging with certain questions and topics over others, by engaging in sustained rather than superficial inquiry, and by exhibiting reasonable tenacity rather than dilettantism or flaccidity. We have already discussed the difficulty of quantifying epistemic states such as knowledge or understanding in Essay One. After all, you might come to know a vast *number* of propositions about football transfers, while still naturally being thought to have learned less than had you spent the same time studying chemistry. The sketch we provided there drew on a model due to Treanor (2013) that involved supposing that understanding more or less could be approached by looking at the sorts of explanations privileged (and ruled out) by an agent’s cognitive commitments—someone with greater understanding will privilege explanations of a finer grain than someone who understands less. With this in mind, we can suppose that virtuous curiosity will be regulated by a sensitivity to the fact that certain inquiries are apt to yield greater understanding than others—given our limited intellectual energy and cognitive ability, the virtuous agent will tend to experience curiosity in such a way as to motivate them to substantially improve their epistemic position, rather than be moved to acquire merely inconsequential and fragmented pieces of information.

Our discussion, then, has suggested the following dispositional excellence regarding curiosity:

? Skilful Disposition: The virtue of curiosity is a disposition to experience, and attempt to satisfy, curiosity that is appropriately discerning, exacting and timely.

By exhibiting this disposition, an epistemic agent will be characteristically motivated to engage in projects of inquiry that will be epistemically enriching—at first pass, this seems like a plausible presification of the virtue of curiosity.

Having made some progress in refining the virtue of curiosity, we can now turn to consider it in light of two theoretical issues that have been discussed by virtue-responsibilists. Firstly, in §5 I will discuss the relationship between curiosity and motivation, before turning to the relationship between curiosity and reliability in §6.

5. Curiosity & Motivations

Within responsibilist virtue epistemology, one popular idea is that motivations play an important role in determining the status of an intellectual trait as a virtue. For instance, Zagzebski claims that is *constitutive* of intellectual virtues that they “are all based in the motivation for knowledge”, where ‘knowledge’ should be read as encompassing epistemic goods in general.¹²⁷ At first pass one might think this criterion unequivocally supports the contention that the disposition to curiosity can count as intellectually virtuous, given that by its very nature curiosity is a motivational state aimed at improving one’s epistemic position—it seems to involve motivation for epistemic goods *par excellence*.

However, while agreeing with the general thought that facts about motivations are important for delineating virtues, I suggest that reflection on the different motivational structures an agent might possess yields a rather more complex picture than we might initially have expected. Indeed, I will suggest that motivational facts can actually *debar* certain dispositions to curiosity from manifesting intellectual virtue (§5.1) before attempting to sketch a more precise account of the required motivational structure for virtuous curiosity (§5.2).

5.1 Vicious Motivations

An agent can experience motivating affective states in virtue of underlying facts about what they value and care about. For instance, someone might feel disgust directed at a piece of meat as a result of their long-standing belief that carnivorous diets are immoral. Interestingly, such underlying explanatory facts can be rather idiosyncratic and even non-transparent to us—e.g. one might be angry at the smell of a cigar because it unwittingly evokes an association with a much-disliked family member. This is just to say that our tendency to experience many basic emotions admits of explanation. Furthermore, these further explanatory features of our psychology—the vegetarianism, the dislike of the family-member—can themselves be evaluated as to whether they can undergird positive evaluation in some domain or other. For instance, as we will discuss below, there is a long-standing discussion in moral philosophy as to what the basis of our motivations must be in order to warrant praise for our conduct.

¹²⁷ See Zagzebski (1996: 167). Similar thoughts are to be found throughout Baehr (2011). And Montmarquet (1993: 21) emphasises similar motivations regarding the acquisition of truth and avoidance of error. Others—such as Roberts and Woods (2007)—disagree with this line of thought, suggesting that there is no single motivation unifying the intellectual virtues.

Although curiosity is a paradigmatic motivation for inquiry, curiosity itself can be motivated by a diverse range of idiosyncratic concerns and interests. In light of this, consider the following typically Hobbesian passage on how we might explain differences in intellectual character:

The passions that most of all cause the differences of wit are principally more or less desire of power, of riches, of knowledge and of honour. All of which may be reduced to the first—that is, desire of power. For riches, knowledge and honour are but several sorts of power. [Hobbes 1998: 48]

Zagzebski (1996: 169-170) claims that the truth of such a view would not impact any determination of epistemic virtue, suggesting that it is irrelevant if the motivation for knowledge is *derivative* of the desire for power. However, I disagree—attributing the disposition to curiosity as an intellectual virtue cannot be reconciled with certain motivational structures, such as being entirely derivative of lust for power. This can be appreciated by considering the following scenario:

Desire of Power. Clement is a scientist who reliably becomes curious about the latest scientific controversies and inquires into them in order to acquire knowledge. However, this disposition has been cultivated at length by Clement as it enables him to realise his intense desire for intellectual recognition by his colleagues, to trump the achievements of his rivals, and to gain professional influence. His underlying motivations are apparent when, if a puzzling scientific phenomena is explained by an ‘opponent’, Clement experiences bitter resentment rather than taking any pleasure in the satiation of his curiosity.

I suggest that in such cases—when an agent’s motivation for epistemic goods is counterfactually dependent on their desire for acclaim and social standing—their disposition towards curiosity does not manifest epistemic virtue, even though it might be skilful in the sense of being appropriately discerning, exacting and timely.¹²⁸ A rough way to explain this thought is with the observation that such cases feature a motivation to acquire epistemic goods, but without showing any appreciation of them. Consider this by way of the analogy with anger suggested above. If someone feels anger and contributes money to anti-smoking causes solely in virtue of the smell of cigar smoke reminding them of a much-disliked family member, then on many accounts they lack moral virtue—they are motivated to act morally without any appreciation for the relevant moral factors making the tobacco industry immoral. Similarly, when you feel curious and conduct scientific research from a deep-seated desire to dominate your opponents, you lack epistemic virtue because you are motivated to acquire knowledge without

¹²⁸ I leave the question of how such motivations affect the attribution of other intellectual virtues as an interesting question for future research. For instance, while it seems that one can possess the *trait* of conscientiousness even if this is derivative from power-lust, it is an open question whether the possession of this trait ought to count as a *virtue* in such cases.

displaying any appreciation for the relevant epistemic factors making scientific research valuable. Below, in the next section, I will substantiate this thought in more detail. But first, a potential objection.

A potential—although I think misguided—objection to my line of thought denies that such discreditable motivations can engender genuine curiosity. This is an interesting worry because little philosophical attention has been devoted to what separates curiosity from merely desiring to answer a question.¹²⁹ Brief reflection suggests there *is* a clear difference between curiosity and desiring to answer a question. While we can form epistemic desires directly from external influence, curiosity seems to paradigmatically arise endogenously. Consider: you might desire to find out the answer to a question—<how many reams of paper are in the cupboard?>—simply because your boss requested you to find out the answer. It seems obvious that such a desire does not *entail* curiosity about how many reams of paper are in the cupboard. Given that there is such a distinction, we need to be sure that curiosity really can be borne from non-virtuous intellectual motives.

In fact, curiosity can stem from a variety of motivational structures, not all of which warrant virtue-ascription. Inquirers with a diverse variety of peculiar motivations exist, and it is rather implausible to deny that they can experience curiosity about the progress of their endeavours. For instance, if a Bond villain wakes up and, without any external pressure, is consumed with the desire to find out whether 007 has fallen into his trap, it is difficult to motivate the view that this attitude cannot possibly count as curiosity just because it stems from a motivation for world-domination. Notably, experiencing curiosity about some question *does not require considering, or inferring from, one's background motivations*. Regardless of whether one's preoccupations are with advancing scientific knowledge or with world-domination, long-running interests and goals simply tend to generate—and form the explanatory basis for—curiosity relevant to these goals and interests. Indeed, this holds for entirely mundane cases of curiosity that are neither virtuous nor vicious. For instance, the explanation for my occurrent curiosity about the Blues football score is the fact that I support the Blues and desire my team to perform well—yet, I don't entertain these motivations prior to experiencing curiosity on Saturday afternoon about the score. Rather, my curiosity just arises as a predictable result of these concerns. Such examples show that intellectually pure motivations are not constitutive of, nor necessary for, experiencing curiosity. This intuition is vindicated by the consensus in the psychological, neuroscientific, and philosophical literature that curiosity can be experienced by cognitively unsophisticated animals. It is implausible to claim that the curiosity of the fieldmouse is explained by any appreciation of intellectual goods.

¹²⁹ For example, each of Whitcomb (2010), Friedman (2013a) and Carruthers (2018) attempt to pin down the nature of curiosity by identifying it with an attitude or desire that aims at answering questions.

In summary: there is no reason to think that curiosity cannot be rooted in inappropriate motivations, such a lust for power. When curiosity is dependent on such motivations, it does not manifest intellectual virtue—even if it otherwise exhibits a skilful disposition.

5.2 Virtuous Motivations

In this section I will try to refine our understanding of what the right sort of motivational structure is for virtuous curiosity, and distinguish it from a different way in which an agent can exhibit epistemic virtue.

One thought that appears throughout the responsibilist literature is the idea that intellectual virtues should be grounded in something like a ‘love’ of epistemic goods, where this is taken to be a sort of appreciation of the value of knowledge, understanding, and so forth.¹³⁰ In light of this, we might wonder whether curiosity is virtuous when it is motivated by this sort of epistemic love? While this is a suggestive idea, it is not yet sufficiently precise to clarify the conditions for virtuous curiosity. For, it seems that one can be curious—in an appropriately discerning, timely and exacting way—without ever entertaining any thoughts about epistemic value. This is because, as we have discussed, curiosity often arises spontaneously rather than from any sort of inferential reasoning. Therefore, we need to clarify the sense in which curiosity can rightly be said to be motivated by an appreciation of epistemic value.

To sharpen our investigation, it will be helpful to further consider a rough parallel with a motivational debate in moral philosophy alluded to earlier. An ongoing controversy concerns what sort of motivations are required in order to deserve moral praise for performing an action—what motivations have moral worth?¹³¹ Philosophers have divided into broadly two camps on this issue: the *de dicto* and the *de re*. The *de dicto* side claims that morally right actions are praiseworthy when they are motivated specifically *by the concern to do the morally right thing*.¹³² As such, this position holds that morally worthy motivations are motivations that refer to morality. This position, in addition to being rather intuitive, has the advantage of capturing our deliberation in cases of moral dilemma; we are uncertain what to do, because we are uncertain what the *right thing* to do is.¹³³ However, a number of criticisms have been raised against this view. Some have argued that *de dicto* motivations are inappropriately ‘fetishistic’.¹³⁴ For example, it has been suggested that you have “one thought too many” if you bring moral concepts into your deliberation about whether to alleviate someone’s immediate suffering—shouldn’t one just be motivated to act directly by the suffering?¹³⁵ Another related criticism is that being

¹³⁰ E.g. see comments throughout Baehr (2011) and Roberts and Woods (2007), especially chapter 6.

¹³¹ The modern precursor to this debate is often taken to be Smith (1994), though this is of course a question that has a long pedigree—for instance, found in the Kantian claim that one should act from moral duty.

¹³² See Isserow (2018) and King (*forthcoming*) for very recent discussions of this literature.

¹³³ Sliwa (2016) presses this point.

¹³⁴ See Drier (2000) for discussion.

¹³⁵ This general thought is due to Williams (1981), though the precise interpretation is controversial.

motivated by the thought of doing the right thing doesn't display any responsiveness to the salient features of the situation, leaving open the possibility of only accidentally reaching the morally correct conclusion.¹³⁶ These worries suggest an alternative *de re* view of appropriate moral motivations: roughly, that one ought to be motivated by the morally salient features of the situation—e.g. the suffering of the person—rather than by entertaining any moral concepts.

I suggest that a roughly analogous distinction can be drawn regarding one's motivations to inquire.

Firstly, one can engage in a valuable epistemic activity from the very thought that so doing is an intellectually good thing to do. Consider two examples:

Logic: Miles is practicing some logic exercises. He isn't a natural and finds the practice frustrating, but recognises that going through a large number of exercises is the best way to become proficient and that it will be intellectually good for him to overcome this challenge.

Project: Julie has been working on an article about Roman history and been made aware of some evidence that might be difficult to explain on her view. Hearing about this new evidence was disheartening, but Julie recognises that it is important to be rigorous when engaging in academic inquiry.

Here, we have examples of the motivation to inquire from the very thought that certain projects have epistemic value for ourselves or for others. These are not cases of curiosity. Sometimes when we are motivated by the very thought of doing an intellectually beneficial thing—when we take our epistemic medicine, so to speak—we are not curious at all. However I think that such motivations are an important part of being a virtuous inquirer, because sometimes grit and determination are required for intellectual flourishing. It is unreasonable to expect that every epistemically important project of inquiry will be one we happen to have sustained curiosity about. We need not insist here, as has been implied in the ethical debate, that only one type of motivation for inquiring can be praiseworthy.¹³⁷ One can inquire from a virtuous appreciation of epistemic goods, despite lacking curiosity. This occurs when one reflects on what is in the epistemic interests of yourself or your community, and acts accordingly.

In contrast, curiosity often arises without reflection on any distinctively epistemic or intellectual concepts. Indeed, not only does experiencing curiosity not require one to consciously infer from any

¹³⁶ E.g. see Arpaly (2002) or Markovits (2010).

¹³⁷ Although, to the extent that virtue-based approaches emphasise the importance of harmony between one's desires, emotions, intentions and actions, it will certainly not be *optimal* for an agent to be motivated, across the board, by de dicto epistemic considerations. One is more likely to have a healthy appetite for inquiry when one inquires from curiosity and takes pleasure in learning, compared to the agent who is motivated by the thought of intellectual duty.

consideration of epistemic value, we often become curious without having any particular justifier for our curiosity in mind. Consider the following routine and appropriate case of curiosity:

Hiking: On a hiking holiday, Hamish becomes curious about a number of geological questions regarding the landscape: “what causes Earthquakes?”; “is Ben Nevis growing or shrinking—and why?”; “how do geologists distinguish different types of rock?”

While such curiosity will not—at least not typically—result from any inference, it is still possible to cite certain reasons as explanations for one’s curiosity. Furthermore, it is perfectly intelligible to ask *why* someone is curious about something, and be impressed or unimpressed with their response. This further supports our thought that curiosity can be evaluated against the deeper values, interests and concerns an inquirer has. Often, the reasons we cite for curiosity will be rather general and basic. For instance, Hamish might simply say: ‘I like to learn about nature’, ‘I enjoy knowing why things are the way they are’, or ‘it’s fun to find out how geologists think about the world’. Despite their simplicity, I suggest that these commonplace responses are precisely the sort of considerations that bespeak a virtuous appreciation of inquiry and the epistemic goods it yields. That is to say, these answers show a sensitivity to the reasons why such inquiry into geology is epistemically fertile: it will provide you with explanatory knowledge about the natural world, a deeper appreciation of the processes shaping our environment, and acquaintance with new ways of categorising and investigating different phenomena. Plausibly, it is the promise of these epistemic goods that explains why inquiring into geology—rather than into the latest football gossip—holds considerable epistemic value. So, one’s curiosity can exhibit an appreciation of epistemic goods just through showing a sensitivity to the very general reasons for why some project of inquiry is likely to be intellectually valuable. This doesn’t require any sophisticated consideration of the epistemic good; we can exhibit this sensitivity just by accurately making use of very general abilities to identify fruitful areas of inquiry. A welcome upshot of this view will be that it avoids requiring an agent to hold any high-level theoretical beliefs about epistemic value in order to qualify as holding a virtuous motivation for inquiry.

An agent’s curiosity will be virtuous, I suggest, to the extent that it is *non-instrumentally* motivated by their appreciation of epistemic goods.¹³⁸ This aims to capture the intuitive thought that an agent who is motivated to understand the natural world for its own sake displays intellectual virtue, but not the agent who views understanding the natural world as a means to acquiring lucrative employment (such as the geologist motivated by the prospect of winning a job with an oil prospecting company). I do not

¹³⁸ In reality, agents will likely often have mixed motivations—this is why I suggested that one is virtuous *to the extent* that one is motivated by epistemic goods. Very few agents will entirely exclude self-interested considerations—such as how they are perceived by others—from the development of their intellectual character. We shouldn’t require perfection in order to ascribe intellectual virtue. But, given the difficulty of measuring the influence of various motivations on our intellectual character, I express no view on whether and how we should determine a *threshold* for epistemic virtue.

think there is any single test that we can apply in order to discern whether an agent's underlying motivations are really rooted in their appreciation for epistemic goods, rather than in their concern for something else. We are left with the procedures that we use in ordinary life to try and work out what an agent really values. That is, we look at what they say (can they give an account of why they care about a given topic?), at how they act (do they inquire when there is no prospect of any non-intellectual benefit?), we reflect upon the sort of things that seem to provide them with pleasure (do they relish personal glory or contributing to the success of a project?), and we speculate as to how they would behave in certain counterfactual scenarios (would they still be so interested in that project if it was not currently academically fashionable?). In this, we determine whether curiosity stems from virtuous motivations just like we consider whether an agent has morally virtuous motivations—we examine the agent's self-reports in light of their behaviour and other values, before consulting our considered intuitions about the normative status of their character.

In summary, curiosity needs to be situated in the right sort of motivational orientation in order to be virtuous. I've suggested that this motivation is a non-instrumental appreciation of intellectually valuable inquiry and the epistemic goods it promises—and one can show this appreciation just by exhibiting a sensitivity for the very general reasons for why certain forms of inquiry are intellectually good to engage in.

6. Curiosity without Reliability

Our discussion has yielded the following account of the virtue of curiosity:

? Virtuous Disposition: The virtue of curiosity is the disposition to experience appropriately discerning, exacting and timely curiosity, motivated by a non-instrumental appreciation of epistemic goods.

Must we further augment this account? One topic of debate amongst virtue responsibilists concerns whether dispositions and traits need to be 'reliable' in order to qualify as intellectual virtues. For instance, Driver (2001) and Zagzebski (1996)—contra Montmarquet (1993) and Baehr (2011)—have suggested that virtues need to be reliably successful in achieving the epistemic ends they are motivations for. This position aims to capture the thought that epistemic virtues only deserve to be identified as such due to their connection with facilitating us achieving our epistemic goals, such as acquiring knowledge or true belief.

How does this commitment sit with curiosity? As we have mentioned, curiosity is quite clearly a motivation for epistemic goods. Dominant approaches will unpack this as a motivation to answer particular questions. And, as was clarified above, virtuous curiosity will stem from being non-instrumentally motivated to acquire these epistemic goods. Furthermore, we have already posited a number of excellences that will be involved in having a disposition to skilful curiosity. The question

now is whether virtuous curiosity requires the further excellence of being reliable in securing epistemic goods? While there is no straightforward way to settle this issue, I will now argue that this is not an attractive requirement on the virtue of curiosity.

An initial argument against the reliability requirement concerns whether it can explain our normative judgements. We can, and do, evaluate an agent's curiosity (or lack of it) *before* and *irrespective of* whether they have engaged in any inquiry. Indeed, this essay has been replete with such judgements. Persistent and overweening curiosity about football transfers is less admirable than discerning curiosity about the sciences. These normative judgements are available to us without considering whether the agent will be *reliable* in acquiring knowledge or not. Such normative judgements cannot be explained if we impose a reliability requirement on the virtue of curiosity, because there has not yet been any attempt to acquire epistemic goods whether reliably or unreliably. This is *prima facie* evidence that the question of whether a given instance of curiosity exhibits intellectual excellence is distinct from the question of whether the agent will then be reliable in securing certain epistemic goods.

A second observation is that insofar as it plausible to think that epistemic virtues have certain *roles*—for instance, in governing certain parts of our intellectual conduct, or protecting us from certain forms of intellectual deficiency—it does not seem to fall under the ambit of a virtue of curiosity to ensure that our inquiries are successful. Put simply, while the possession of a curious disposition is crucial in *motivating* inquiry, it seems to have little to do with *regulating* it. In this sense, it is quite consistent to imagine someone who is both appropriately curious yet inquires in a haphazard and unskilful way. Furthermore, there are already a number of virtues in the literature that are aimed at ensuring our inquiries are successful—conscientiousness, impartiality, tenacity, and so forth—so it would be seem theoretically redundant to insist that a virtue of curiosity must range over the same territory. The natural thought here is that curiosity is important in spurring inquiry, and then it is the exercise of other epistemic virtues that determine whether that inquiry goes well or badly. Indeed, *regardless* of whether one's inquiry is motivated by curiosity or by something else, it is still the exercise of other epistemic virtues that will determine whether or not our inquiry is successful. And if an agent fails to be reliable, this will be in virtue of intellectual failings independent of their curiosity. So, given that it is the presence or absence of other epistemic virtues that determine whether or not we are reliable in securing epistemic goods, this raises the question as to why reliability should be taken as a requirement for virtuous curiosity at all—after all, the issue of reliability must be settled by looking at facts about the agent that have little connection with the fact that their inquiry was motivated by curiosity.

Thirdly, it seems that curiosity can be unreliable in virtue of it stemming from *admirable* epistemic tendencies. For instance, an inquirer might tend to be curious about novel questions at the forefront of scientific research that are extremely challenging. This seems to be an intellectually praiseworthy disposition. However, even if they are an entirely responsible inquirer, the difficulty of these questions

might make it rather unlikely that such an agent will uncover the correct answer in any given inquiry. While this sort of research might incidentally yield epistemic goods ‘along the way’, this sort of agent will not be reliable in achieving the true answers to the question that their curiosity is directed at. Regardless, it seems compelling to think that their curiosity is epistemically virtuous. This can be further supported by considering intellectual exemplars who happened to get things wrong.¹³⁹ An example: it seems right to praise Thales for his curiosity about metaphysics despite his unreliability in reaching the correct conclusions. Having lofty intellectual ambitions that negatively influences your overall reliability does not preclude a high estimation of your intellectual character. This is further evidence that curiosity can amount to an intellectual excellence irrespective of whether it is reliable. Imposing a reliability requirements seems to involve endorsing a tension between what counts as a genuine display of virtue on the one hand, and what conduct attracts positive intellectual evaluation on the other.

Of course, none of these are knock-down objections to the reliability requirement—such a requirement could be maintained on broader theoretical grounds, regardless of how it pans out for any particular epistemic virtue. However I do think that, by showing the reliability requirement to be rather implausible when applied to individual excellences that seem to be compelling candidates for qualifying as epistemic virtues, we strike at the overall plausibility of such a requirement. While we might balk at attributing any sort of virtue (such as curiosity) to a hopelessly poor inquirer, we should remember that such agents will not be deemed intellectually virtuous *overall*. All we need to concede is that such a figure is closer to being considered epistemically virtuous than someone who is both a hopeless inquirer *and* exhibits no tendency to be curious in an appropriate manner. Absent any persuasive argument to the effect that we should take the intellectual virtues to be unified in such a way as that you either possess or lack them all, as a package-deal, it seems more reasonable to accept a position whereby one can possess the virtue of curiosity without being reliable in securing epistemic goods.

7. Conclusion

This essay has provided the beginnings of an account of the normative role of curiosity for epistemic theorising. Starting with the observation that curiosity takes both a state and a trait reading, we found that state curiosity has an important role in motivating inquiry but told us little about what broader norms should regulate an agent’s intellectual endeavours. Moving on to focus on the trait of curiosity, I argued that curiosity can qualify as an intellectual virtue when it amounts to an excellent disposition to experience curiosity. Taking our cue from various deficiencies and excesses that a curious inquirer can display, we found that the virtuous agent will exhibit a *skilful* disposition to be curious: they will experience curiosity that is appropriately *discerning* (in being directed at the right object), *timely* (in arising in appropriate circumstances), and *exacting* (in not being satisfied too easily). In this way, a primary normative role for curiosity is as an intellectual character trait that motivates agents to engage

¹³⁹ This type of argument appears in various discussions stemming back to, at least, Montmarquet (1993: 21).

in epistemically valuable inquiry. I then considered how my account impinges upon broader issues in virtue epistemology. Drawing on debates within the responsibilist literature, I identified two potential further challenges to the sufficiency of my account: (i) concerning what sort of deeper motivational orientation is required for virtuous curiosity, and (ii) concerning whether one's curiosity needed to be reliable in achieving particular epistemic goods. Contrary to what some have held, I argued that the virtue of curiosity is incompatible with a range of motivations such as an underlying desire for power. This provided us with some insight into how different motivational configurations—even when they involve a motivation to acquire epistemic goods—can debar a trait from qualifying as a virtue. Rather, I suggested that curiosity must be rooted in a *non-instrumental* appreciation of epistemic goods in order to be virtuous. And finally, I provided a number of reasons for thinking that curiosity does not require reliability in order to count as an intellectual excellence—thus putting some pressure on those have claimed, schematically, that all epistemic virtues require reliability.

Essay Four

Legal Inquiry & Statistical Evidence

0. Overview

Earlier in the thesis, in Essay two, I outlined a perspective on collective inquiry which emphasised some broad continuities between collective and individual inquiry. Namely, we focused on similarities in the various attitudes and processes involved when individuals and groups move along the trajectory from ignorance to knowing the answer to some question. Although these similarities are instructive, it remains to be seen how extensive the commonalities are between different entities that engage in inquiry. Particularly, in that earlier essay, there was little discussion of the sorts of *epistemic norms* that apply across different cases of collective inquiry. With this in mind, an interesting global question that we will begin to address in this final essay is the following: do the same epistemic norms governing individual inquirers apply to other entities that engage in inquiry (courts, corporations, democracies, etc.)? This is a very broad question, and I will here approach it via an instructive proxy debate that has garnered significant attention in recent years; namely, how should the legal system treat purely statistical evidence?

The legal system is an institution that—via the practice of instituting courts—inquires into various questions of pressing public importance. The paradigm examples are questions about the criminal guilt or civil liability of particular parties. Just as occurs in individual inquiry, courts consider evidence for or against various conclusions and issue verdicts endorsing one conclusion or the other. Given these similarities, it is not surprising that epistemologists have recently become interested in legal philosophy. Furthermore, it is also perhaps not surprising that many epistemologists who have become interested in the law often apply their own favourite epistemic theories to make prescriptive claims about legal inquiry. In this vein, a number of philosophers suggest that epistemic norms developed for the purposes of capturing intuitions about—and providing normative guidance to—individual inquirers should be transposed to the legal context.

This aim of the following essay, via considering the use of statistical evidence in different contexts, will be to identify some key *differences* between how individuals and courts ought to engage in inquiry. If we can show that epistemic norms from individualistic theorising are in fact not well-suited to the role of regulating legal inquiry, then we will have some evidence in favour of a negative answer to the global question about the normative symmetry between the norms governing individuals and other entities that engage in inquiry. In broad strokes I will argue that courts are subject to a variety of different goals and imperatives that do not apply to individual agents and that this constrains legal inquiry in important ways. The upshot is that we should not accept the idea that both individuals and institutions such as courts must be governed by the same epistemic norms when they engage in inquiry. In this sense, the conclusion will be that in order to formulate acceptable epistemic norms for different institutions, we need to closely attend to the nature and function of their inquiry in the specific context in which it occurs.

1. The puzzle of purely statistical evidence

Recently, the practice of deciding legal cases on purely statistical evidence has been widely criticised.¹⁴⁰ Many feel uncomfortable with finding someone guilty on the basis of bare probabilities, even though the chance of error might be stupendously small. This is an important issue: with the rise of DNA profiling, courts are increasingly faced with purely statistical evidence. In a series of prominent papers, various philosophers appeal to epistemic norms governing individual inquirers to argue that legal verdicts should never be based on statistical evidence alone. My project is to show that not only have recent discussions mischaracterised how the law actually treats purely statistical evidence, but also argue that there are cases in which relying on such evidence is justifiable. In developing these claims, I demonstrate why we cannot solve jurisprudential questions by simply appealing to the norms found in theorising about individual epistemic agents. Rather, settling issues of legal theory requires us to look further afield, taking into the account the unique tasks and challenges that face legal systems.¹⁴¹

This argument will be developed over five sections. In this section I present the intuitive case against purely statistical evidence, explain why it creates a jurisprudential puzzle, and separate out different responses to this puzzle. In §2, I introduce key concepts of evidence law necessary to fully appreciate legal treatment of statistical evidence. In §3, I demonstrate why DNA profiling is properly characterised as statistical evidence before showing that—contrary to recent assertions—legal systems do not generally prohibit basing verdicts on bare statistics. In §4, I introduce and reject a range of arguments *against* the use of purely statistical evidence which appeal to epistemic norms that govern individual inquirers. These arguments overlook the fact that individuals, unlike courts, can *hedge* in response to statistics. In §5, I suggest that our qualms about purely statistical evidence stem from the fact that it makes the possibility of error salient. However, because managing the possibility of error in different contexts is a delicate yet necessary task that every fallible legal system must fulfil, I close by arguing that the best solution to the puzzle of statistical evidence takes each case on its individual merits, rather than entertaining any general prohibition against such evidence. One upshot is that we might reasonably convict in cases involving DNA evidence but not in some of the other cases that have particularly discomfited critics of statistical evidence. Finally in §6 I conclude and offer some suggestions about how the issues discussed generalise beyond legal inquiry and have ramifications for how we theorise about the nature of blame and responsibility.

¹⁴⁰ For an overview see: Gardiner (2018). For specific arguments, Smith (2018); Littlejohn (2018); Blome-Tillman (2015; 2017); Moss (2016); Di Bello (*forthcoming*); Thomson (1986); Enoch et al (2012) provide representative examples. Hedden and Colyvan (*forthcoming*) provide perhaps the most sympathetic treatment in the literature, in their defence of construing standards of proof probabilistically.

¹⁴¹ This essay focuses on the common law and will not discuss civil law traditions. However much of what I say is relevant to any legal system.

Below are two cases—taken from the criminal and civil domain respectively—standardly used to motivate the idea that there is something amiss about legal reliance on purely statistical evidence. (I’ll sometimes call it ‘PSE’ for short).¹⁴²

PRISONERS: 100 prisoners are exercising in the prison yard. 99 of them attack the guard, putting into action a plan that the 100th prisoner knew nothing about. The 100th prisoner played no role in the assault and could have done nothing to stop it. There is no further information that we can use to settle the question of any particular prisoner’s involvement.¹⁴³

BLUE BUS: A bus causes injury to a pedestrian, but it is not known which company the bus belongs to. On the route where the accident occurred, the Blue Bus Company runs 75% of the buses. There is no further information that we can use to settle the question of which company the bus belongs to.

Many have the intuition that it would not be permissible to sanction in these cases. In the psychological literature, this phenomenon is called the ‘Wells’ effect’ after an influential study of juror reluctance to assign liability on the basis of bare statistics.¹⁴⁴ These patterns of intuition give rise to the *puzzle* of purely statistical evidence because, given what we know about the fallibility of other types of evidence such as eye-witness testimony, statistical evidence will often be *more truth-conducive* than other sorts of evidence on which we routinely base legal verdicts. This comparative discomfort towards statistical evidence has been found so puzzling that it has been referred to as the ‘proof paradox’, when combined with the thought that it strikes against our very understanding of satisfactory legal proof itself. The issue, then, is the following: how concerned should we be about the intuitive reluctance many have about basing legal verdicts on PSE? Should we accept these intuitions and construct theories that vindicate them, or should we ignore them and focus on the fact that such evidence can be very reliable? Schematically, there are three different approaches to PSE. (In the following, what counts as ‘strong’ probabilistic support will vary depending on the standard of proof being used).

Always: It is always acceptable to base a legal decision on purely statistical evidence, so long as it provides strong probabilistic support to that verdict.

Sometimes: Sometimes it is acceptable to base a legal decision on purely statistical evidence when it provides strong probabilistic support to that verdict.

Never: It is always wrong to base a legal decision on purely statistical evidence even when it provides strong probabilistic support to that verdict.

¹⁴² Here, and throughout, evidence is used in the sense familiar to the legal system rather than as a philosophers’ term of art. Thus, evidence refers to that which is adduced in the context of a trial to support or rebut some contention. It does not refer to any doxastic or epistemic state.

¹⁴³ Wording taken from Redmayne (2008).

¹⁴⁴ See Wells (1992).

We should hold at the outset that **Always** is incorrect. In some scenarios, there are specific moral reasons to avoid appealing to statistics in legal inquiry. One example might be the use of crime statistics about certain demographics. For instance, even if it were the case that 99.9% of all game console thefts were carried out by women, this shouldn't be enough to convict some individual woman if, for example, a games console is stolen in a situation where only one man or one woman could have been involved. The repeated use of such statistics would, I suggest, amount to an *oppressive* pattern against certain demographic groups.¹⁴⁵ However, even though the use of statistics can constitute oppression in some cases, this worry does not generalise. For instance, in PRISONERS, a group of prisoners who happen to be exercising at the time of collective disorder does not constitute a stable demographic group. This suggests that there are different strategies for vindicating the intuitions we have about PSE cases.

One approach is to identify *specific features* of certain cases to explain why those particular types of case should not be settled by purely statistical evidence. For instance, one worry about PRISONERS might be that, if all 100 prisoners were captured, we would knowingly convict one of them wrongly. A rather different approach is to argue that *generic features* of all PSE cases mean that such evidence is not a fitting basis for a legal decision in general. The latter approach has been the most popular strategy, and has been used to motivate **Never**. This view will be my target—I do not think that the generic features shared in common by PSE cases suffices to justify a general prohibition against relying on bare statistics. In this sense, we should not draw an inference from our aversion to some cases involving PSE to every case involving PSE. Rather, in this essay, I will be advocating for **Sometimes** by arguing that different uses of PSE invoke different issues. On balance, sometimes we should allow that a legal decision can be based on purely statistical evidence. Furthermore, I will claim that these cases are relatively common in the law.

2. Evidence law: essential concepts

Before getting into the weeds, we must first briefly introduce three key concepts of evidence law: the Standard of Proof; the Evidential Burden; and Evidential Admissibility. Extant discussions have tended to be light on legal details, only focusing on the first of these. However, a cursory appreciation of other aspects of evidence law is essential to correctly understand the current place of statistical evidence in the law and ask normative questions about what its proper role should be.

¹⁴⁵ See Mogensen (*forthcoming*) on how racial profiling can constitute oppression. There is a growing literature on what we should think about such demographically targeted statistics. An important debate is whether forming beliefs on the basis of racial profiling is a *moral wrong* (see e.g. Basu *forthcoming*), an *epistemic wrong* (see e.g. Bolinger *forthcoming*) or something else. Here, I will be leaving this debate aside—I will be not be dealing with examples of statistical evidence that targets any particular demographic group.

2.1 The Standard of Proof

A central part of evidence law—familiar to most readers and well-trodden in recent philosophical literature—is the Standard of Proof (‘SoP’).

The SoP is the standard applied to determine whether a body of evidence provides strong enough support to render a specified legal (or quasilegal) judgement appropriate. There are various different SoPs which apply to different types of legal judgement. Everyone knows the ‘beyond a reasonable doubt’ (‘BRD’) standard that applies to verdicts in criminal cases. Most academics will also recognise the ‘balance of probabilities’ (‘BoP’) standard that applies to verdicts in civil cases (e.g. contractual disputes). BRD and BoP are not the only standards found in legal systems. Various other formulations are adopted by different (quasi)legal institutions for various purposes, imposing more and less exacting standards.¹⁴⁶ SoPs often apply outside the context of trial verdicts; for example, prosecutors are governed by a SoP when deciding whether to charge a suspect with a crime. The way in which different standards of proof have been formulated and interpreted has changed significantly over time.¹⁴⁷ For instance, the standard for criminal conviction was once glossed as ‘moral certainty’, a fallibilist contrast to the notion of ‘mathematical’ or ‘metaphysical’ certainty that is immune to sceptical doubt. The way in which current standards have been interpreted has also seen significant variation—for instance, attempts to clarify the BRD by quantification (e.g. a .95 credence) or by conflating it with the standards used in non-legal inquiry (e.g. the standards used to make an important decision in one’s life) have waxed and waned in popularity over time.

To be explicit about how these standards work: the BRD standard, for example, dictates that ‘a court should only issue a finding of criminal guilt iff the entirety of the evidence adduced supports the guilt of the accused beyond a reasonable doubt!’ Judging whether or not a body of evidence meets the applicable SoP is left to the relevant *fact-finder*. In the context of a trial verdict, the fact-finder will be a judge or a jury. Deciding whether evidence meets the SoP is *not* a technical judgement in the sense of requiring the fact-finder to interpret the law. Rather, the role of the fact-finder is to use their everyday facility in discerning the plausible from the implausible to make a judgement about a body of evidence which they have had the opportunity to consider first-hand. Fact-finders, so long as they stay with bounds of reasonableness, have a broad discretion in judging whether or not the evidence meets the relevant SoP.¹⁴⁸

¹⁴⁶ E.g. ‘reasonable suspicion’ (often governing police searches) and ‘reasonable degree of likelihood’ (used to assess UK asylum claims) are standards less demanding than the BoP. Other standards—e.g. ‘clear and convincing evidence’ (sometimes used in civil cases regarding children’s welfare or psychiatric assessment) fall between the civil and criminal standards.

¹⁴⁷ See Laudan (2006) and Roth (2010) for discussion.

¹⁴⁸ This is why appellate courts are reluctant to overturn first-instance judgements about whether the evidence satisfies the SoP solely on the grounds that the appellate court disagrees with the original judgement.

2.2 Sufficient evidence: discharging the ‘evidential burden’

A widely overlooked yet crucial legal concept is that of leading *sufficient evidence* and discharging the evidential burden.

In contrast to the non-technical task of the fact-finder to decide whether a body of evidence meets whatever SoP is in play, the doctrine of sufficient evidence is a matter of *law* regulated by the judge. (Note: judges can be both legal arbiters and fact-finders—these are distinct roles).¹⁴⁹ The doctrine of sufficient evidence places a legal burden—often called the ‘*evidential burden*’—on the party aiming to establish some claim. This burden acts like a filter that weeds out unacceptably weak cases. Specifically, it means that the party aiming to establish some contention must adduce sufficient evidence in favour of that claim before it can legitimately be put to the fact-finder (who will then, taking into account *all* of the evidence, decide whether enough has been done to prove the contention on the relevant standard of proof).¹⁵⁰ For instance, a very minimal construal of sufficient evidence in the context of a criminal case is that *some* evidence must be adduced which: (i) suggests that a crime has been committed, and (ii) suggests that the accused is the perpetrator. If this burden is not discharged, then the court does not move into the deliberative stage of considering the evidence against the relevant SoP. For instance, to continue using criminal procedure as an example, a failure to discharge the evidential burden will force the judge to hold that the accused has ‘no case to answer’ and bring proceedings to a close. Some jurisdictions have more heavyweight conceptions of sufficient evidence than others. One rationale for stricter conceptions is to mitigate the possibility of miscarriages of justice by imposing more judicial oversight on cases before they are left to the jury.¹⁵¹

Crucially, sufficiency of evidence is not about the *quality* of evidence. The conditions under which evidence counts as sufficient is a technical matter of law for the judge to decide, while quality of evidence is a matter for the fact-finder. For example, it might happen that the prosecution in a criminal case adduces evidence that is rightly put to the jury, only for the jury to rationally find the evidence unpersuasive. This is entirely consistent; sufficiency of evidence in a criminal jury trial concerns

¹⁴⁹ To illustrate: in a criminal *jury* trial the judge will be the legal arbiter but not the fact-finder. In a non-jury criminal case, the judge will be the legal arbiter throughout the trial *and then* play the role of the fact-finder in deciding whether or not the accused is guilty.

¹⁵⁰ There are differences in how the burden to prove facts are allocated in criminal and civil trials, but these nuances need not detain us for our purposes.

¹⁵¹ An interesting example is found in the Scots’ law requirement of *corroboration*: in Scottish criminal trials, crucial facts must be corroborated by at least two independent sources of evidence to discharge the evidential burden. This further exemplifies the distinction between sufficiency of evidence and the SoP; e.g. one might suppose that one individual source of evidence (e.g. an eyewitness account) could conceivably establish guilt beyond a reasonable doubt, but due to lack of corroboration this—in a Scottish criminal trial—might fail to constitute sufficient evidence.

whether enough evidence has been adduced to warrant putting the case to the jury in the first place, not whether enough has been done, *ultima facie*, to persuade the jury beyond a reasonable doubt.

2.3 Admissibility of evidence

A third concept of evidence law will be familiar to readers so we need only introduce it very briefly. This is the idea of *evidential admissibility*. Rules governing admissibility determine when certain types of otherwise relevant evidence must, as a technical matter of law, be excluded. There are a variety of different rationales behind different rules: for example, some rules are in place to discourage improprieties at other stages of the judicial process (for example, prohibiting evidence acquired using coercion during police interview) while other rules preclude evidence that is perceived as unreliable or irrelevant (for example, well-known rules about hearsay evidence).

When evidence should be excluded is a matter of law, determined by the judge. Evidence that is ruled inadmissible and excluded does not count towards discharging the evidential burden, and is not considered by the fact-finder in determining whether or not the applicable SoP has been met.

3. Statistical evidence in the law: a closer look

Recent discussions create the impression that legal systems are generally adverse to purely statistical evidence. One influential paper asserts outright that such evidence is *inadmissible*.¹⁵² Other philosophers do not locate the problem in admissibility, but nonetheless suggest that evidence law is unfavourable to bare statistics.¹⁵³ These are empirical assertions about the common law. Not only are they interesting in their own right, but, if they were true, they would lend credence to certain normative views. This is because it would suggest that **Never** is the orthodox position and imply that those hostile to PSE are simply seeking to vindicate rather than revise existing legal practice. (Indeed, Blome-Tillmann 2017 calls wholesale antipathy to PSE in the law *conservative* and opposing views *revisionist*).

Against this characterisation I show that close inspection reveals no general prohibition against PSE in the law. Rather, the rise of cases involving DNA profiling show that PSE is playing an ever more prominent role. Furthermore, brief reflection will show that the most promising way to reject PSE appeals to the idea of *sufficient* evidence rather than, as some have suggested, a rule of admissibility.

¹⁵² See Enoch et al (2012: 198). Moreover, Blome-Tillmann (2017) suggests that such evidence should be inadmissible in defending what he views as a ‘conservative’ position on evidence law.

¹⁵³ For instance, see Smith (2018), Gardiner (2018), Di Bello (*forthcoming*) for a selection of those who claim that using PSE would be legally revisionary.

3.1 DNA evidence is statistical evidence

In recent years the vast majority of reported legal cases involving PSE use DNA profiling. As such, it is necessary to say something about DNA evidence and why it is properly regarded as statistical.

Everyone knows that courts—particularly criminal courts—consider DNA evidence: but what is the nature of this evidence and how is it presented to the court? DNA evidence comes from extracting, analysing, and making statistical inferences from genetic material. A typical case is where such material (e.g. saliva, hair, semen) is found at a crime-scene—for instance, investigators finding a cigarette-butt containing saliva at the site of a break-in. This genetic material is processed in a laboratory to build a profile of which alleles—the genes we inherit from our parents—appear at which locations on a strand of DNA. From this process, DNA evidence is used in different contexts. One is where suspicion has already fallen upon some person and DNA profiling is used to bolster the case against them. This might involve taking further genetic material from a suspect—e.g. with a saliva swab—and testing it against the initial genetic material found at a crime scene. If the samples match, this further supports the pre-existing suspicion. A second use of genetic material, the use that we will primarily be concerned with, is so-called ‘cold-hit’ DNA evidence. In such cases genetic material (e.g. taken from the scene of a crime) is found to bear similarity to a DNA profile already held on file for some other reason. This evidence is then used as the basis of a criminal case against the person to whom the matching DNA profile belongs. In a cold-hit case, this person is *not* someone upon whom suspicion has independently fallen.

DNA evidence is presented to the court in the form of *expert testimony*. An expert witness, often a forensic scientist with a doctoral degree, is called to court to testify to the significance of the DNA evidence. In a cold-hit case, the most important element of the testimony concerns an estimated statistical frequency of the allele-configuration in a given target population: for example, here is Dr Clara O’Sullivan providing the key piece of inculpatory evidence in a recent high profile cold-hit DNA case:

Expert Witness: “We give an estimation of the chance of somebody randomly unrelated in the population, having the same profile, given that [the suspect has] the profile, is one in a thousand million.” [*Wilson v DPP* 2017 IESC 54 at 5.18]

In this sense, DNA evidence is statistical evidence: on the assumption that the genetic material taken from the crime-scene belongs to the person who committed the crime, the *incriminating component* of the DNA evidence is that it is exceptionally unlikely that the genetic material does not belong to the person being accused of the crime.^{154,155} When there is no additional evidence apart from the statistical

¹⁵⁴ It is worth pointing out that even critics of PSE do not tend to argue that DNA evidence is non-statistical.

¹⁵⁵ To be explicit, the probability of error being testified to by the expert witness concerns the possibility of there being a *random match* in the target population who shares the relevant DNA characteristics as the sample, but

estimate based on the DNA profile—as was the case in *Wilson*—we have a case involving *purely* statistical evidence. With this in mind, we can return to consider how the law treats PSE.

3.2 Purely statistical evidence in the law

There has been some confusion in the philosophical literature regarding statistical evidence in the law. One mistaken suggestion is that statistical evidence is generally *inadmissible*. If it were, we wouldn't ever see criminal convictions relying on DNA evidence. Indeed, if statistical evidence were inadmissible, then it would be treated like other types of inadmissible evidence—such as evidence obtained through torture—i.e. it would not be found even in conjunction with non-statistical evidence.¹⁵⁶ This is clearly not the case: courts take statistical evidence into account all the time, on a range of different subject-matters. Of course, there might be particular cases in which certain types of statistical evidence *are* inadmissible—such as evidence pertaining to demographic crime-statistics—but this does not mean that statistical evidence is generally inadmissible.¹⁵⁷ These points are neatly exemplified by a quote from the very same case we drew the example of expert testimony from earlier. The judges in this case, drawing on a range of common-law authorities from different legal systems, stressed the following about Dr O'Sullivan's statistical testimony:

[I]t is appropriate to emphasise that the Court is not here concerned with the admissibility of such evidence. Clearly even evidence of a tangential connection may be relevant in the overall context of a particular case but would be unlikely to provide sufficient evidence, without more, to allow for a safe conviction. [*Ibid.* at 5.5]

As this quote suggests, the salient issue is whether PSE can discharge the evidential burden of providing *legally sufficient evidence*. Evidence can be admissible without being legally sufficient. A prosecutor can lead evidence that the accused owns a set of kitchen knives. While admissible, this evidence would not be sufficient by itself to support a murder conviction.

So, have courts typically held that PSE can discharge the evidential burden? In *Wilson*, the judges answered in the affirmative. Indeed, if we focus on PSE in the form of criminal cases relying on DNA evidence alone, then there are now numerous examples of courts affirming the sufficiency of such

was not involved in the relevant incident. Such estimates do not include other possibilities of error like the DNA sample being corrupted due to cross-contamination or laboratory mistakes.

¹⁵⁶ See *R. v Alan James Doheny; R. v Gary Adams (together)* [1997] 1 Cr. App. R. 369 for general rules on admitting DNA evidence.

¹⁵⁷ For instance, some cite *United States v Shonubi* 103, F.3d 1085 (2d Cir. 1997) in support of the inadmissibility of statistical evidence. However, this case clearly only concerns a very specific practice: punishing someone by using an estimate of contraband they have smuggled over time via a statistical evaluation of the habits of other smugglers. It does not follow that appellate courts disapprove of PSE in general just because they view such speculative sentencing as iniquitous.

evidence to undergird a conviction.¹⁵⁸ This is not to say that the sufficiency of PSE is an entirely uncontroversial matter.¹⁵⁹ However, it is important not to overplay the extent of the controversy. Some cases cited by critics simply do not address the legal position on purely statistical evidence.¹⁶⁰ Another mistake to avoid is citing *any* case in which a verdict based on purely statistical evidence is overturned as supporting the claim that legal practice *generally* eschews PSE. For instance, *R v. Watters*¹⁶¹ is often cited as exemplifying the reluctance of the courts to rely on bare statistics. However, if one attends to the details of the judgement, the problem was not with statistical evidence per se, but rather with the fact that it failed to rule out the accused's brother being responsible—and this was a possibility that had been raised by the defence. Indeed, the judges in *Watters* (at 21 per LJ Kay) are at pains to emphasise “we are not for one moment saying that merely because there was no other evidence ... that this appeal has to be allowed”.

It is true that one can find cases where the courts express hostility towards PSE, especially when we move away from DNA profiling. The case of *Virginia & S.W. Ry. Co. v Hawk* in the civil domain—involving statistics about railway accidents—is a good example. However, other cases such as *Kaminsky v Hertz Corp*—superficially very similar to the BLUE BUS case as it involved ascribing liability on the basis of the percentage of vehicles owned by a particular company—where judges seem entirely amenable to purely statistical evidence. My own view is that such cases are not aspiring to be binding precedents about the standing of PSE in general, to hold for each and every case. Rather, they attend to specific details of the case at hand in making their judgement.¹⁶² This is often how the common law develops: it proceeds in a piecemeal fashion by drawing fine distinctions between preceding cases as justice demands in the case at hand. However, getting into the details of legal interpretation is

¹⁵⁸ For a selection, see: *R v. Hanratty* [2002] 2 Cr. App. R. 30 (esp. comments at 127); *R v Adams (no1)* [1996] 2 Cr. App. R. 467 and *R v Adams (no2)* [1998] 1 Cr. App. R. 377; *R v Weir* [2000] 5 WLUK 751; *State v Toomes* 191 S.W.3d 122, 129 (Tenn. Crim. App. 2005); *State v Hunter* 861 N.E.2d 898 901; *State v Davis* 698 N.W.2d 823, 826-27. Also see *Director of Public Prosecutions v. Connolly* [2011] IESC 6 (which, although dismissed for other reasons, affirms the sufficiency of PSE in the form of a forensic scientist analysing packages containing amphetamine). Beyond these citations, Roth (2010) is a legal scholar who discusses a variety of cases in more detail than I have space for here. Further, in *Toomes* the judges cite a number of other historical cases in support of the sufficiency of statistical evidence, for the interested reader. It is also worth noting that many uses of cold-hit DNA leading to conviction go unreported, in particular cases in which the suspect pleads guilty.

¹⁵⁹ *R v. Lashley* [2000] EWCA 88 is an example of a DNA case being rejected on the basis that five or six people in the UK might fit the DNA profile. This doesn't support a general rule against the sufficiency of DNA evidence however: often the probabilities involved in DNA cases make error even less likely than in *Lashley*.

¹⁶⁰ A danger in this literature, given the sheer volume of cases being cited, is that erroneous citations can be unwittingly recycled. E.g. *Guenther v Armstrong Rubber Co* 406 F.2d 1315, 1318 (3d Cir. 1969) has been cited as an authority against PSE. However, this case only provides authority for another legal point: namely, how courts should react to a plaintiff offering contradictory accounts.

¹⁶¹ [2000 WL 1791491]

¹⁶² For instance, it seems a better interpretation that *Virginia* concerns how juries should be *directed* in relation to probabilities: particularly, they should not be left to make probabilistic conjectures. *Kaminsky*, on the other hand, places weight on defeasible presumption created by the appearance of the truck and the unique demands of justice applicable to the law surrounding motor vehicles. Neither is a general endorsement or prohibition on PSE.

unnecessary to make the following broader point about the sufficiency of PSE: as the rise of DNA profiling shows, PSE is not insufficient evidence *in general* even though there may be individual cases in which the courts express reluctance to rely on it.

I have attempted to set the record straight on how PSE is treated in the law. This serves a number of purposes. Beyond the simple motivation of striving for accuracy, showing that PSE *is* currently used in the law rebuts the thought that there is universal distaste for PSE that philosophers merely have to vindicate. Rather, especially given the prevalence of DNA evidence, the proper status of PSE is a live question that we should consider with an open mind. Moreover, our discussion also allows us to sketch a sympathetic proposal on behalf of critics as to how their critique should be implemented. We should assume that critics of PSE are ill-served by suggesting that statistical evidence should be *inadmissible*. This would rule out using DNA evidence in any context, while also excluding statistical evidence when lead in conjunction with other types of evidence. I take it that this position is unappealing—there is little reason to suppose that statistical evidence has *no role* to play in legal decision-making. Rather, it is much more promising to interpret critics of PSE as arguing that our best evidence law should hold that: *statistical evidence alone cannot constitute legally sufficient evidence* to undergird a criminal conviction or finding of civil liability.¹⁶³ Under this approach, if the party with the burden of proof adduced only statistical evidence then the fact-finder (such as the jury) would not be asked to deliberate—rather, the case would be thrown out. This seems like the most plausible interpretation of how critics could have their proposal implemented by common law legal systems. Of course, as we have shown, such a proposal would be rather revisionary: in particular, it would disallow criminal convictions based solely on DNA evidence.

4. Doxastic approaches

The foregoing empirical facts about legal practice still leave the normative question open: should there be a blanket prohibition on bare statistics carrying the day in court? A number of philosophers provide an affirmative answer, seeking to identify features shared by all PSE cases that justify rejecting them. A common thread to these arguments is the suggestion that it is epistemically inappropriate for courts to issue a verdict on the basis of bare statistics. This section will argue against three different approaches that press this strategy.

¹⁶³ Given that many juries (in the cases cited previously) have found PSE to satisfy the relevant SoP, it is also unappealing to suggest that it is a conceptual truth about any particular standard that it cannot be satisfied by statistics alone.

4.1 Three arguments against purely statistical evidence

A prominent line of thought—recently endorsed by Smith 2018; Littlejohn 2018; and Blome-Tillmann 2017—proposes to explain and vindicate the intuitions against PSE by appealing to the epistemic norms that apply to individual inquirers.¹⁶⁴

Before looking at the specifics, we'll need some brief background theory. In epistemology there is an ongoing debate about how *individuals* should treat purely statistical evidence, best exemplified by preoccupation with lottery cases. Many suggest that there is an important normative asymmetry between statistical evidence and other types of evidence (such as direct sensory apprehension). For instance, some theories claim that relying on statistical evidence doesn't yield *knowledge*. Taking a concrete example, a dominant thought has been that no matter the size of a fair lottery, one cannot *know* that one has a losing ticket just on the basis of it being very likely.¹⁶⁵ Other theorists suggest that basing beliefs on purely statistical evidence leads to the absence of other normative properties, such as certain types of *justification*.¹⁶⁶ This concern with statistical evidence is one front in a broader dispute about which normative property should be central to the epistemic lives of individual inquirers. For instance, some endorse the idea that statistical evidence does not yield knowledge apiece with a more general view according to which knowledge determines when it is appropriate for someone to act upon a belief, to make an assertion, or even to hold a belief in the first place.

A number of philosophers have used ideas from epistemic theories concerning individual inquirers to argue against relying on purely statistical evidence in the law.

Blome-Tillmann's argument

Firstly, Michael Blome-Tillmann appeals to the idea that knowledge is the norm of action.

Assuming that we ought to act on p only if we know p , the ideal and strongest epistemic state a court could be in, from a normative point of view, is one in which the court knows the defendant is at fault.
[Blome-Tillmann 2017: 284]

However, accepting that requiring knowledge is an unreasonable legal standard, Blome-Tillmann endorses an alternative view on which the *probability of possessing knowledge* is what matters:

Courts need to be able to make decisions in the absence of the epistemically normative ideal—that is, in the absence of knowledge that the defendant is at fault. What is needed is a measure of gradual approximation to the normative ideal. One way to measure such approximation is in terms of evidential

¹⁶⁴ These three support this view most explicitly. But it is also at least suggested by Enoch et al (2012); Bolinger (*forthcoming*); Pritchard (2015); and Buchak (2014).

¹⁶⁵ Most prominently, see Williamson (2000). This thought has also been supported by empirical survey data (see Ebert et al. 2018). See Hawthorne (2003) for discussion.

¹⁶⁶ See Sutton (2007), Smithies (2002), or Smith (2016); each focus on lottery cases.

probabilities that one knows that the defendant is at fault. Knowledge then remains the normatively ideal state, but epistemic success in courts of law can be measured and understood in terms of something less than knowledge—namely, the evidential probability that knowledge has been achieved.

[*Ibid.*]

On this basis, Blome-Tillmann suggests that there should be a rule on which the SoP can only be satisfied by evidence that raises the probability that the conclusion is *known* above a certain threshold (e.g. $>.5$ for the civil BoP standard). As, *ex hypothesi*, statistical evidence by itself cannot yield knowledge, PSE cases will not meet this threshold.

Smith's argument

Martin Smith's argument requires a brief theoretical preface.

In his work on individual inquirers, Smith (2016) developed a concept of justification featuring a notion he calls *normic support*. Normic support can be introduced with the idea of something 'calling for explanation'.¹⁶⁷ According to this theory, you are normically justified in holding some belief *p* only if the falsity of *p*, given your evidence, would call for some special explanation. So, while it might be very unlikely that I have won the lottery based on probability alone, it wouldn't be *abnormal* for me to win in the sense of calling for some special explanation; thus, my belief that I have lost would not be normically supported. On the other hand, were I to *look* at my ticket and see that the numbers don't match the winning numbers, then my belief that I have lost would be normically supported: some special explanation would be called for, were I to be mistaken.

Smith observes that the puzzle of statistical evidence in the law can be explained by noting that the body of evidence involved fails to *normically support* the fault of the defendant. He then derives a normative claim from that observation:

[O]ur judgments about the presence or absence of normic support track our judgments about whether an affirmative legal verdict would be acceptable or unacceptable. What I suggest is a standard of proof that is met only if a proposition is normically supported by the evidence – only if the evidence makes the falsity of that proposition less normal, in the sense of calling for more explanation, than its truth. What I suggest is that a verdict of guilt or liability is only acceptable in so far as this normic standard is met. This is my proposed solution to the legal puzzle of statistical evidence. [Smith 2018: 1209-1210]

¹⁶⁷ The formal version of Smith's theory appeals to a framework that ranks worlds in terms of closeness, details of which can be found in his (2016). For our purposes, nothing turns on whether the informal or formal version is used.

Littlejohn's argument

Littlejohn holds that the reasonableness of a belief should be judged against the standard of acquiring knowledge, something that purely statistical evidence cannot yield. With this in mind, he endorses the following principle:

Reasonable Conviction: It is not permissible to punish a defendant if it isn't reasonable to believe the defendant to be guilty. [Littlejohn 2018: 15]

Littlejohn defends this principle primarily by appealing to considerations of legitimate punishment:

My defense of Reasonable Conviction begins with a reminder that punishment is an act that differs in an important way from acts like betting on football matches. This is because punishment is supposed to be a way of holding someone accountable and it involves a backwards-looking element that other actions often lack. Thus, the act in question (e.g. imposing a prison sentence) has to be guided by certain kinds of considerations to be a punishment. [...] If we have a system of rules that governs decisions to punish or to refrain from punishing, it would seem that the rules should require that the decision to impose the harms associated with punishment be made only when the punishment can properly express blame or at least treat the defendant as accountable for some specific deed. It would not be proper to blame unless the relevant parties could properly believe that the defendant did something blameworthy. [*Ibid* 16-17]

Littlejohn (s.8) expands on this suggestion by claiming that there are reactive attitudes and emotions associated with blame which can be rationalised by outright belief but not by a high degree of confidence that the subject of blame has transgressed.

4.2 Against doxastic approaches

Each of these arguments takes its cue from individual epistemology. Blome-Tillmann transposes to the law a theory about when it is rational for an individual to act. Smith applies to the legal domain his notion of epistemic justification which concerns when it is rational for an individual to believe something. And Littlejohn stakes out a position on when legal punishment is legitimate by appealing to the normative role of outright belief in licensing an individual to hold reactive attitudes associated with blame.

For what follows, I will call views fitting this mould *doxastic approaches* because they take their cue from the norms surrounding belief in individual inquirers. Doxastic approaches are favourable to **Never**. This is because, on doxastic views, relying on purely statistical evidence can never fulfil certain requirements: such as providing knowledge, rationalising action, legitimising blame, and/or providing normic support for some proposition. As such, their arguments do not rest on details of particular cases but rather appeal to generic facts about the epistemic power of purely statistical evidence. Given that the three arguments outlined each claim that courts should conform to the same normative standards as

individuals when confronted with PSE, one way to object to them would take issue with how they characterise the norms governing individuals. I won't take this approach here. Rather, let's grant whichever epistemic view critics prefer; my argument will attack the purported normative symmetry between doxastic and legal norms.

Let's start with an observation: *in an ideal scenario*, when a court issues a judgement—e.g. that Harry killed Sally, or that Jim's hedge is encroaching Jules' property—it would be one that an individual inquirer could cheerfully believe outright on the evidence. This is because, in a perfect world, the evidence would always support the legal verdict without leaving any room for doubt or uncertainty. Thus, courts would only ever judge—and, act, blame and punish—when the evidence was strong enough to support an individual inquirer believing outright (or knowing) that the relevant party is at fault. We do not exist in this ideal world. The reality is that courts are routinely faced with bodies of evidence that are difficult to evaluate and fail to unequivocally vindicate either side. This is why, for instance, courts cannot simply ask the fact-finder to look at the evidence and then report *whether they believe outright* that (i) the defending party is at fault, or (ii) the defending party is not at fault.¹⁶⁸ Such an approach would be a recipe for paralysis. Take a civil case where the evidence only weakly suggests that a given party has been negligent; this will not be enough to support an individual believing outright that the relevant party either is, or is not, civilly liable. Nonetheless, it is incumbent on a tribunal to decide one way or the other. This is why legal tribunals have rules to ensure that they operate in a thoroughly binary fashion—something is treated as having occurred or not, an accused is judged guilty or not guilty, a defending party is deemed civilly liable or they are not.

Lord Hoffman, in an influential case on proof, puts it succinctly:

If a legal rule requires a fact to be proved (a “fact in issue”), a judge or jury must decide whether or not it happened. *There is no room for a finding that it might have happened.* The law operates a binary system in which the only values are 0 and 1. The fact either happened or it did not. ... If the party who bears the burden of proof fails to discharge it, a value of 0 is returned and the fact is treated as not having happened. If he does discharge it, a value of 1 is returned and the fact is treated as having happened. [Lord Hoffman in *Re B* 2008 UKHL 35]

When the law treats something as having happened, this does not entail that the evidence supports an outright belief that the thing occurred. Even if the evidence is muddy, a final decision must be made to settle the case at hand. It is this *burden of deciding* that explains why evidence law contains standards and burdens of proof. As such, judges and juries do not need to form outright beliefs in order to issue verdicts: they need only look at the evidence lead by the party with the burden of proof and measure it

¹⁶⁸ To be clear: the doxastic views I critique below accept that such an approach is untenable. For instance, Littlejohn does *not* endorse a symmetrical view on which one must either know that a party is guilty *or* know that it is not guilty before judging accordingly—rather, his idea is that we should know the party is guilty before judging them guilty, and should exonerate otherwise. Thanks to anonymous referees for pressing me to clarify.

against the relevant standard of proof (e.g. Jules might have the burden of showing, on the balance of probabilities, that he owns the land that Jim's hedge is growing on). It is worth underlining this point because there is an assumption explicit in some defences of doxastic approaches (e.g. when Blome-Tillmann writes about a court knowing that the defendant is at fault) and implicit in others (e.g. when Littlejohn discusses the role of beliefs in rationalising punishment) that legal verdicts are a species—or at least close analogue—of belief. However, legal judgements and beliefs are crucially dissimilar: while belief is often thought to be a *two*-place relation between an individual and some proposition, legal judgements involve a *three*-place relation between an *agent* (the fact-finder: whether in the guise of jury or judge), some *proposition* (e.g. the civil liability of a defendant), and a *standard of proof* (e.g. the balance of probabilities). The fact-finder in a court is not asking whether they believe a proposition simpliciter, but rather deciding whether the evidence supports that proposition on a given evidential standard. Meeting this standard *may or may not* mean that she believes the proposition outright.

Arguments against purely statistical evidence can nonetheless be reframed entirely in terms of *evidential support*. On this approach, critics can say that legal verdicts must be based on evidence that could yield knowledge or normically supported belief in an individual. However, this position begs the following question: why should we suppose that norms regarding what individuals should do on the basis of certain types of evidence will automatically transfer to legal tribunals? After all, it seems reasonable to suppose that the conditions under which an individual should judge or act might differ from when a legal tribunal should judge or act. Of course, both courts and individuals will be governed by *some* norms of epistemic rationality. But we should question whether these norms are *identical*. For, if there turn out to be significant differences between the epistemic predicaments of individuals and legal tribunals, then extant arguments against legal reliance on bare statistics will miss the mark.

Recall the fact that legal verdicts are binary. In light of this fact, when it comes to the question of how to act—and whether to judge someone responsible for something as a preface to blaming or sanctioning them—individuals and courts face different options. One crucial difference is that a legal tribunal must decide one way or the other, while individuals are able to *hedge*. Let's illustrate with an example.

Neighbour. Suppose that you come across a confidential file stating that there is DNA evidence linking your new neighbour to a series of sadistic and violent murders. The file contains expert testimony stating that DNA material supports a 999,999/1,000,000 chance of him having committed these awful crimes.

Should you believe outright that he is guilty? Well, from the perspective of addressing various practical questions, there is a sense in which it is unimportant. After acquiring this evidence, will you ask him to babysit your children, to look after your spare key, or perhaps accept his offer to go hiking—just the two of you—in the backcountry? Certainly not. As an individual with statistical evidence

indicating the stupendous likelihood that your neighbour is a serial killer, you will adopt various prudential policies, rejecting his offer to share a late-night walk through the moors. This is true even on the assumption that purely statistical evidence doesn't license outright/normically supported/knowledgeable belief. Even if you only believe <he is extremely likely to be guilty> rather than believe that he is guilty *outright*, you will change your behaviour considerably.¹⁶⁹ Moreover, this response is perfectly *rational*—indeed, your partner may rightly criticise you if, for example, you give the neighbour the spare key for safe-keeping. This is true, even if everyone concedes, “Of course, we don't *know* that he's a killer!” The ability to hedge is very important: to further illustrate, consider that many norms relied on by critics of statistical evidence were developed to deal with lottery-cases. However, individuals faced with deciding whether or not they have won the lottery are not compelled to return a binary verdict of <yes> or <no>. We are not forced to *either* act as if we have lost and throw away the ticket, *or* act as if we have won and sign up for early retirement. Rather, an individual can maintain a hedged belief that they have *probably* lost—thus rationally licensing them to hold on to the ticket but not to rely upon potential winnings in making financial plans.

Legal tribunals, faced with statistical evidence, are in a rather different epistemic predicament. Courts cannot judge that a party is stupendously likely to be guilty: rather, they must issue a judgement of guilt or innocence.¹⁷⁰ A legal tribunal cannot forbear from ascribing fault *and* act in a way that reflects a high degree of confidence in fault. So, we are owed an argument as to why the epistemic norms governing individuals (who *can* hedge) should result in a legal tribunal (which *cannot* hedge) exonerating a suspect when the evidence is purely statistical. For, even if it is true that an individual should believe and act on the proposition <probably *p*> rather than <*p*> upon considering bare statistics supporting *p*, this is not an effective argument for supposing that a court should issue and act on the judgement <not-*p*> on the basis of the same evidence. After all, we *do* think it is rational for individuals to treat people differently after gaining statistical evidence about their possible conduct. Legal tribunals cannot navigate between extremes by hedging like individuals can, because courts have the burden of issuing a decisive verdict. The normative position of a tribunal facing statistical evidence is thoroughly different from that of an individual—hence, if you are attempting to explain why a court should not find someone guilty on the basis of bare statistics, it is not enough to simply appeal to how things stand with individuals.

1. Individuals should not believe/act as if <*p*> given evidence E.

¹⁶⁹ I put my point in terms of probabilistic belief, but it works equally well in a credence-based framework.

¹⁷⁰ A fascinating exception here is the *third verdict* in Scottish criminal cases of ‘not proven’. The existence of this verdict is deeply controversial from the perspective of criminal law theory, and its abolition has been mooted at various points. I won't take any stance here. In any case, the existence of this third verdict in one legal system doesn't cause trouble for my general point—a ‘not proven’ verdict is functionally identical to a ‘not guilty’ verdict insofar as it doesn't license any additional action by the legal system.

2. Individuals should believe/act as if $\langle \text{probably } p \rangle$ given E.
3. Courts can only affirm/act as if $\langle p \rangle$ or $\langle \neg p \rangle$ given E.
4. # Courts should affirm/act as if $\langle \neg p \rangle$ given E.

5. Rehabilitating statistical evidence

Let's now make progress on what courts *should* do in response to purely statistical evidence.

Although courts must issue a binary verdict, the intuitions behind the puzzle of PSE suggest that it isn't *obviously* appropriate to simply decide in favour of whatever conclusion has most probabilistic support on the evidence. Here's a first stab at explaining why this is: courts have a very special role in doling out criminal and civil penalties, and this role brings with it a host of competing responsibilities. Consider a criminal court. If it acts on a high probability that someone is guilty (e.g. by punishing them) when they are in fact innocent, then it has wronged that party. Moreover, given the public nature of legal judgements, false convictions threaten the credibility of the legal system itself.¹⁷¹ Courts must aim to both avoid wronging those they punish and to uphold confidence in the legal system. However, legal systems also have duties that can be frustrated by exonerating the accused. If it exonerates a guilty party, then the court has failed in its aim of delivering the appropriate response to crime (whether fleshed out in terms of protecting society from further criminal activity, deterring crime, giving due retribution, providing rehabilitative opportunity, or performing some semiotic function). Hence, courts facing purely statistical evidence are in the epistemic predicament of being obliged to issue a binary decision where there are unfortunate consequences of getting it wrong in either direction. The norms of legal inquiry are encumbered with the delicate task of managing these risks.

Dealing with the possibility of error in a satisfactory way within the binary framework constraining tribunals is a jurisprudential question facing every legal system. In this sense, legal systems cannot only be concerned with accuracy. When setting the relevant evidential standards they must also take a stance on what type of errors should be minimised at the expense of others. Consider the BRD standard of proof. The evidential standards required for criminal conviction must strike a difficult balance. First and foremost, legal systems need to avoid falsely convicting the innocent. Indeed, in criminal law theory, it is typically thought to be worse to falsely convict the innocent than mistakenly acquit the guilty. However, avoiding false convictions is not the only imperative for the criminal law: otherwise we would endorse a beyond *any* doubt standard of proof. Clearly, it is essential to ensure that those who are guilty of crimes have an acceptable chance of being convicted. The BRD standard is supposed to strike a balance between the desiderata of avoiding false convictions while still ensuring that conviction

¹⁷¹ Indeed, some (e.g. Thomson 1986; Bolinger *forthcoming*) point out that legal verdicts resemble *assertions*. As such, we might expect their status as assertions to generate additional norms. I won't rely on this thought here, but this is one way to precisify my argument if you are tempted by the verdict/assertion parallel.

for criminal activity is a realistic prospect. That the BRD standard strikes the best balance is far from a platitude. For example, Laudan (2006) has suggested—based on the idea that one is more likely to be a victim of a recidivist criminal than falsely convicted—that the standard should be weakened. Regardless of where the boundary is set, some wrongful convictions and false acquittals are inevitable; the question is how to apportion these risks.

With this in mind, I want to suggest the following about cases involving purely statistical evidence: our intuitive discomfort about these cases stems from the fact that such evidence makes the possibility of error salient. This makes us reluctant to sanction, because any error would constitute an injustice against the sanctioned party. Nonetheless, it is part and parcel of recognising the different responsibilities that legal systems have, and the fallibility of legal tribunals, that we tolerate the fact that courts will sometimes make errors when discharging the burden of deciding one way or the other. Satisficing the different aims of the legal system requires accepting that we will sometimes convict people for crimes they did not commit. When confronted with cases that throw this fact into sharp relief it naturally makes us uncomfortable, because we are understandably hesitant to explicitly endorse trade-offs in which the innocent will foreseeably suffer miscarriages of justice to facilitate justice being done elsewhere. However, this balancing act is something that all legal systems *must* countenance in virtue of their unique role in settling disputes. In contrast, the possibility of error is usually much less salient when relying on non-statistical evidence. For, when we rely on evidence such as eye-witness testimony, it is much easier to construct a psychologically persuasive narrative that obscures the possibility that the verdict is mistaken.

My contention about error-salience chimes with both philosophical speculation and empirical evidence. For instance, John Hawthorne (2003: 15-20) suggests that certain scenarios involving statistical evidence are structured in such a way as to engender what he calls *parity reasoning*. Parity reasoning occurs when we consider a proposition p against the thought that some particular p_i in a set of subcases p_1, \dots, p_n might not obtain and then realise that our evidence doesn't speak in favour of any particular instance not obtaining over any other particular instance. (Imagine this with respect to lottery-tickets: the evidence doesn't privilege any ticket—or 'subcase'—not losing over any other.) The result of parity reasoning is to make the chance of getting things wrong salient; after all, the evidence doesn't privilege the conclusion that any *particular* p_i in the set will not obtain. While ordinary non-statistical evidence doesn't lend itself to being conceptualised in this structured way, lottery cases clearly inspire parity reasoning because we naturally just compare our generic ticket to the other generic tickets, each with the same chance of winning. I think something similar occurs when there is inculpatory statistical evidence. Consider a cold-hit DNA case. When we hear expert testimony asserting "There is a 1/1,000,000 chance of this DNA belonging to someone else", I think we are moved to compare our current epistemic position with how things would be in other close possible worlds (or subcases), including one in which the person was innocent. By making this comparison, we realise that our

evidence doesn't privilege any particular subcase and hence the possibility of error (i.e. a false conviction) is made very salient. This thought about salience is also supported by work from psychologists who have studied the Wells' effect: for instance, Neidermeier et al. (1999) suggest that juror reluctance regarding PSE is explained by the ease of imagining a scenario in which the defendant is not at fault. This is all consistent with the empirically well-supported 'story model' of decision-making—a popular paradigm for explaining how juries make decisions in trials—on which legal verdicts are made through imposing a linear narrative onto the evidence presented at trial.¹⁷² Judgements based on bare statistics are not easily subsumed into a psychologically compelling narrative, and PSE does not lend itself to inclusion in a story more plausible than an opposing story in which the accused was simply absent.

If something like the salience thought is right, what we should do about our intuitive discomfort with PSE? One reaction is to simply disregard our intuitions, favouring PSE because it could bring about greater accuracy in the legal system overall. However, I do not think my argument compels us to take this position. This is because legal systems prize virtues other than accuracy, as is richly exemplified by different canons of evidence law. For example, some rules on evidential admissibility serve broader goals not closely tied to maximising accuracy. Consider the exclusion of evidence gained through impropriety—what US law calls 'the fruit of the poisoned tree'—such as where some permission (e.g. a search warrant) was needed but not obtained. This evidence is excluded, even if evidence acquired through an inappropriate search might be highly reliable in a given case, or indeed highly reliable in general. Such rules serve aims such as discouraging inappropriate police conduct but they also serve the broader aim of ensuring *public confidence* in the legal system—an imperative embodied in the familiar maxim that 'justice must not only be done, but be seen to be done'. Evidence law is shot through with rules that are not about maximising local accuracy.¹⁷³

Let's now return to the PRISONERS case.

PRISONERS: 100 prisoners are exercising in the prison yard. 99 of them attack the guard, putting into action a plan that the 100th prisoner knew nothing about. The 100th prisoner played no role in the assault and could have done nothing to stop it. There is no further information that we can use to settle the question of any particular prisoner's involvement.

One argument *against* convicting in PRISONERS that is consistent with all we have said so far is the following: guilty verdicts in these types of case would offend certain desiderata of evidence law other than maximising accuracy. An obvious candidate is the imperative of maintaining public confidence in the legal system. In this sense, simply the fact that there is a *salient perception of possible*

¹⁷² See Pennington and Hastie (1994) for an overview and discussion of experimental results.

¹⁷³ For other examples, consider rules precluding spouses being compelled to testify against each other; evidence gathered through police coercion; and so forth.

injustice in the PRISONERS case could be good enough reason to hold that the evidence is insufficient to discharge the evidential burden. This line of thought can be bolstered by observing that the sacrifices required by consistently exonerating in cases like PRISONERS are not egregiously costly: cases like PRISONERS are relatively rare, excluding such cases would not predictably enable people to escape justice for certain types of recurring crime, and—assuming that ‘reasonable doubt’ and a 1/100 likelihood of innocence aren’t *radically* divergent standards¹⁷⁴—such a rule would not substantially alter the global accuracy of the legal system.

However, the same considerations do not apply across the board. In particular, even if we suppose that we can vindicate exoneration in PRISONERS on non-accuracy based grounds, I think we should resist this thought regarding (at least some applications of) DNA evidence. Therefore, my argument is one in favour of **Sometimes** rather than **Never**. Consider the following case fitting the mould of prominent uses of PSE in recent years:

DNA: Someone is sexually assaulted in a secluded park. They are unable to provide an informative account of the appearance of the attacker. Their injuries show that there is no question that the act was non-consensual. DNA evidence from the crime matches that of someone on file for some other reason. A forensic scientist estimates the chance of the DNA *not* belonging to that person to be 1 in one billion.

I do not think that this type of case creates the same perception of possible injustice as the PRISONERS case. One reason for this—whether well-founded or not—is that people tend to have faith in the scientific credentials of DNA evidence. This might partly be due to a misconception that the inculpatory component of DNA evidence is not merely statistical. However, even if people are directed to focus on probabilities alone, there has been empirical evidence suggesting that people are more inclined to think that legal verdicts based on PSE are appropriate the higher the probabilities of accuracy are (see Wright et al. 1996).^{175,176} In addition, the sacrifices entailed by consistently exonerating in such DNA cases are more costly than a similar policy regarding PRISONERS style scenarios. Firstly, cases—particularly

¹⁷⁴ For instance, surveys of judges on their view of the BRD standard (discussed in Solan 1999) support this assumption.

¹⁷⁵ Ebert et al. (2018) found regarding lottery cases that subjects are more likely to ascribe justified belief (but not knowledge) when the lottery is much larger—i.e. when the chance of error is much lower. This doesn’t directly support any hypothesis about legal cases. However, it does support the thought that PSE suggesting an *extremely low* chance of error will receive a different (and, indeed, a more positive) evaluation than PSE merely suggesting a low chance of error.

¹⁷⁶ Bear in mind that an accused can lead exculpatory evidence in their favour: particularly, evidence speaking to an alibi. Of course, one might worry that it would be difficult even for an innocent person to overcome such formidable inculpatory statistics. But this is a purely general worry, for it is not always the case that an innocent person can easily rebut incriminating evidence against them. For instance, if someone is the victim of a sustained and skilful attempt to frame them, then various types of evidence (manipulated CCTV, bribed witnesses, etc.) will not easily be undermined. This is a consequence of trials being fallible, not a reason to regard any class of evidence that is usually reliable as generally unsatisfactory.

sexual offences—where the only incriminating evidence is DNA material are not particularly rare. Secondly, the lack of non-DNA evidence is a particular problem in prosecuting certain types of sexual offence, viz. those carried out by strangers with no particular motive in areas where there are unlikely to be eye-witnesses. Ruling DNA evidence to be insufficient would predictably undermine our efforts to prosecute a particular type of case, one that already suffers from perennially low conviction rates. And finally, a rule against this PSE in the DNA case *would* upset the global balance of accuracy in the legal system: a 1 in one billion chance of innocence is not a remotely reasonable approximation of the notion of a reasonable doubt, thus such a rule would lead to many more false acquittals than we would expect from allowing courts to use such evidence. So, drawing on these different factors as justification, we can still endorse the sufficiency of DNA evidence to undergird criminal convictions while vindicating the negative intuition about PRISONERS.¹⁷⁷

My discussion focused primarily on criminal cases, but broadly the same points apply to civil cases too. The law of evidence regulating civil cases also involves a delicate balancing act. Even though criminal sanctions such as imprisonment or the failure to punish crime are not at stake, the issues raised are nonetheless of substantial import. Take BLUE BUS style cases involving a harm being suffered by one party (i.e. being seriously injured in a bus accident). Evidence law must strike a balance between ensuring that people can be compensated for various (often severe) harms that result from the activity of others, and ensuring that others are not easily held liable for these harms erroneously. To put this point in context: either the party who has been struck by a bus receives compensation (e.g. for loss of earnings and other expenses incurred as a result of injury) or they are left with nothing. My own thought about BLUE BUS-type cases is that bare statistics *could* be enough to legitimately result in a finding of liability—sometimes it might be reasonable to distribute economic risks in such a way as to fall on a company with a predominant market share. Of course, this suggestion is defeasible; e.g. we should not take such an approach if it would predictably lead to iniquity for one company over others. However, it doesn't seem *necessarily* unfair to ascribe civil liability on the basis of statistics simpliciter.¹⁷⁸ We can press this thought by conceiving of civil cases involving very long odds that are more akin to DNA cases. For instance, imagine a case where there is only one bus company operating in a given locale. If it were the case that, say, only 1/10,000 buses on the road is owned by a private individual, then it could be adduced as evidence that there is a 99.99% chance of the bus that caused an accident being owned by the MONOPOLY BUS COMPANY. Again, I think that a finding of liability here is entirely reasonable. However, accepting this contrarian position is not really needed to vindicate **Sometimes** in the civil domain. For, it follows from what I have already argued—namely that criminal convictions

¹⁷⁷ There is nothing *sui generis* about DNA evidence—rather, using PSE raises different issues in different contexts and each should be considered on its individual merits against the broader aims of the legal system.

¹⁷⁸ Recall *Kaminsky v Hertz* in which a Blue Bus style scenario led to a finding of liability. The case explicitly considers the unique demands of justice raised by transportation law. This underscores my points: appropriate distribution of risk is case-specific, and my piecemeal approach is in keeping with how the common law develops.

can sometimes be appropriately based on PSE—that the same holds in the civil domain too. This is because many actions that constitute criminal offences—e.g. sexual offences or physical assaults—are also civil wrongs. Hence, it is a platitude that if PSE can rightly undergird a criminal conviction on the BRD standard then it must also be sufficient to support a finding of civil liability on the weaker BoP standard.

6. Concluding Remarks

This essay examined a pressing legal question, namely: how should the courts treat evidence that is purely statistical in nature? A growing number of philosophers had suggested that legal systems tend to scorn such evidence and provided overarching theories attempting to vindicate this purported truism about the law. Against these views, I demonstrated that legal systems do in fact rely on purely statistical evidence: DNA profiling is a prominent example. Then I argued that one guiding thought behind many critiques of purely statistical evidence, namely that courts should be subject to the same epistemic norms as individual inquirers, is unconvincing. This is because courts, unlike individuals, cannot *hedge* their decisions—they must settle matters one way or the other by issuing a decisive verdict. And finally, I argued that the best approach to purely statistical evidence takes each case on its merits and illustrated that we can sensibly eschew such evidence in the sorts of cases that critics find most uncomfortable, but vindicate this evidence in other contexts such as the use of DNA profiling.

The broader upshot suggested by this discussion is that the norms governing individual inquiry will not always be well-placed to apply to any other entity that can be described as engaging in inquiry. One might wonder whether this thought generalises, or whether it is somehow only restricted to legal inquiry? On the face of things, there is good reason to suppose that similar arguments to those developed in this essay will apply to other domains of inquiry. This is because many of the features that distinguished legal inquiry from individual inquiry are not unique to the law. For instance, other entities may face situations in which they must issue a binary verdict in accordance with pre-set rules for responding to evidence. Take a government agency that must inquire into whether a given drug is safe to prescribe for public consumption. Here, the agency may have to issue a verdict of a binary nature—SAFE/UNSAFE—in accordance with rules for evaluating evidence. We should not assume that it is only appropriate for the agency to issue a verdict only if the evidence supports an outright belief that the drug is safe in an individual inquirer. Rather we have to consider the particular role of statistical evidence in evaluation of drug safety; the consequences of withholding a license to a given drug; what sort of presumptions and burdens it is appropriate to hold as the default when assessing the safety of a product; and so forth. Again, the lesson to draw is that formulating norms for ending inquiry should be sensitive to the different imperatives at play in different types of inquiry. These norms will not always conform to what individuals should believe or take as given in practical reasoning when they are in equivalent situations.

Finally, the arguments in this essay have further implications for theorising about the nature of blame and responsibility. Here are two ideas that are both fairly intuitive and have been alluded to by some of the theorists discussed in this essay: firstly, that in order to blame someone for a transgression we should at least *believe* they are guilty, and secondly that the conditions under which we should hold someone criminally responsible for something are a subset of the conditions under which we can properly hold someone morally responsible.¹⁷⁹ This essay demonstrates that each of these ideas warrants further scrutiny. The first of these ideas—whether pitched as a conceptual truth about the very nature of blame, or as a normative claim about when blame is appropriate—has been put under pressure because we have found that court verdicts do not amount to beliefs in the familiar sense of the term, and that they are licensed to issue verdicts on the basis of evidence that would not typically undergird an outright belief. The second idea has been put under pressure because this essay has highlighted that institutions such as the law face different demands from individual inquirers. Hence, it is not a given that the conditions under which individuals and courts should attribute responsibility will march in lockstep; it is not inconceivable that, on broader grounds of public interest, a court might be licensed to hold someone criminally responsible for an action that an individual would not be licensed to hold them morally responsible for. Each of these issues deserves further attention. This illustrates that not only is focusing on inquiry in the context of different institutions an important project in its own right, but also that this focus facilitates productive research into traditional notions that have long been central to philosophical theorising.

Conclusion

¹⁷⁹ Sher (2006) is illustrative of the former; Duff (2009) is illustrative of the latter. These views are widespread.

This collection of essays has advanced a number of philosophical debates relating to the epistemology of inquiry.

In the first essay, I began by examining issues at the end of inquiry. We noted that epistemologists have become increasingly interested in the epistemic state of understanding, often arguing that it is a particularly valuable thing for inquiry to aim at. While this resurgent interest in understanding was largely motivated by difference between understanding and the more heavily theorised state of knowledge, a strain of recent work attempted to reduce understanding to facts about what an agent knows. After outlining how to test such an approach, I argued that this knowledge based reductionism was untenable by providing a counterexample in which replacing knowledge with true belief made no difference to the extent to which an agent understands why something is the case. I then argued that despite the failure of the exclusive focus on knowledge, the basic reductionist idea was a good one. As I argued, a truth based approach—on which understanding reduces to what is truly believed—turned out to be a promising theory. I closed by defending this theory against what is perhaps its most prominent objection; namely, the thought that false scientific theories play an important role in adding to our understanding of the world. On my view, false scientific theories add to our scientific understanding just because they can be an effective way of transmitting true beliefs to their audience.

In the second essay, I made a contribution to the burgeoning research programme of social epistemology by examining collective inquiry. Specifically, I used a focus on collective inquiry to provide a new theoretical framework with which to approach ongoing disputes about whether and how intellectual disciplines such as philosophy and the sciences make intellectual progress in their endeavours. I incorporated the widespread thought that progress consists in acquiring factive epistemic states by arguing that settling inquiry by acquiring factive states such as knowledge is one way that an intellectual community can progress. However, focusing on settling inquiry by (for example) acquiring knowledge failed to let us say much of substance about ubiquitous cases in which an intellectual community is at some intermediate point between ignorance and knowing the answer to some question. In response to this problem, I demonstrated why focusing on how an intellectual community moves through the different stages of inquiry—including wondering about questions, investigating them and gathering evidence, and deliberating and becoming more confident in the right answer—can be seen as distinct forms of progress in their own right. Along the way, I demonstrated why this approach is preferable to an alternative approach simply focusing on the acquisition of collective justification. I closed by considering the ramifications for philosophical progress, suggesting that it is plausible to think that academic philosophy is at an intermediate point between ignorance and knowledge regarding many philosophical questions. Hence, by using my framework, we can say that there are different possible ways in which the philosophical community has progressed short of acquiring knowledge. This amounted to a new way to defend a moderate position on philosophical progress without having to

demonstrate that the philosophical community knows much more than has typically been supposed by critics.

In the third essay, I discussed the motivations for inquiry. In particular, I provided a theory of why curiosity is important for normative theorising in epistemology. I first clarified some different ways in which we cease curiosity, discussing extant work on the idea that there are norms on ceasing to be curious, and then argued that the function of curiosity is simply to motivate inquiry rather than to acquire particular epistemic states. I then developed a theory of when the trait of curiosity amounts to an epistemic virtue. Under my theory, curiosity is epistemically virtuous when it is discerning in the objects it aims at, exacting in being reasonably persistent without exhibiting pathology, and timely in arising in situations when it can be profitably sated. I then considered this in relation to two broader debates in virtue epistemology. The first of these concerned the relationship between epistemic virtue and an agent's motivational structure. I argued that although curiosity seems to be a paradigmatic motivation for epistemic goods, we can still conceive of vicious motivational structures that cause agents to be curious for the wrong reasons. As such, I suggested that virtuous curiosity requires an agent to be motivated by a non-instrumental appreciation of epistemic goods. The second debate I considered was about whether we should accept a 'reliability requirement' on the epistemic virtues—that is, whether only traits that reliably lead to the acquisition of epistemic goods qualify as virtues. I argued against such a requirement, suggesting that curiosity can be virtuous without reliably yielding epistemic goods such as knowledge. This explains why we praise the curiosity of intellectual pioneers like Thales, despite the fact that they often failed to know the answers to the questions they were inquiring into.

In the final essay, I closed by examining issues regarding legal inquiry. In particular, I provided a new perspective on the 'proof paradox', an issue that is attracting renewed attention by those working at the intersection between epistemology and the philosophy of law. After closely attending to (some often neglected) tenets of evidence law bearing on the proof paradox, I critiqued a prominent family of theories seeking to account for the putative discomfort we feel about basing a legal verdict on purely statistical evidence. These accounts—that I called 'doxastic accounts'—sought to appeal to norms governing individual inquirers to explain why we should not rely on bare statistics. I argued against this approach, noting that courts and individuals face substantially different epistemic predicaments—importantly, individuals can hedge in response to statistical evidence while tribunals must issue binary judgements and act accordingly. I then argued that to solve the puzzle of statistical evidence, we must balance the broader—sometimes conflicting—aims of fallible legal systems against each other in order to work out how we should treat statistics in different contexts. The upshot of this balancing process was that we might reasonably reject relying on bare statistics in various low-odds cases that have worried many philosophers, but accept that statistics may appropriately carry the day in other cases such as those involving DNA evidence.

Bibliography

- Archer, Avery (2018). Wondering about what you know. *Analysis*:anx162.
- Aristotle, (1998). *Nicomachean Ethics*. Oxford University Press.
- Arpaly, Nomy (2002). *Unprincipled Virtue: An Inquiry Into Moral Agency*. Oxford University Press.
- Baehr, Jason S. (2011). *The Inquiring Mind: On Intellectual Virtues and Virtue Epistemology*. Oxford University Press.
- Battaly, Heather (2008). Virtue epistemology. *Philosophy Compass* 3 (4):639-663.
- (2010). Epistemic self-indulgence. *Metaphilosophy* 41 (1):214-234.
- (2017). Intellectual Perseverance. *Journal of Moral Philosophy* 14 (6):669-697.
- Basu, Rima (forthcoming). The Wrongs of Racist Beliefs. *Philosophical Studies*:1-19.
- Baumgarten, Elias (2001). Curiosity as a Moral Virtue. *International Journal of Applied Philosophy* 15 (2):169-184.
- Bird, Alexander (2007a). What is scientific progress? *Noûs* 41 (1):64–89.
- (2007b). Justified Judging. *Philosophy and Phenomenological Research* 74 (1):81-110.
- (2008). Scientific progress as accumulation of knowledge: a reply to Rowbottom. *Studies in History and Philosophy of Science Part A* 39 (2):279-281
- (2010a). Social knowing: The social sense of 'scientific knowledge'. *Philosophical Perspectives* 24 (1):23-56.
- (2010b). The epistemology of science—a bird's-eye view. *Synthese* 175 (1):5 - 16.
- Blome-Tillmann, Michael (2015) 'Sensitivity, Causality, and Statistical Evidence in Courts of Law' *Thought* 4(2): 102 – 112.
- (2017). 'More Likely Than Not' - Knowledge First and the Role of Statistical Evidence in Courts of Law. In Adam Braddon-Mitchell, David and Frank Jackson. 2007. *Philosophy of Mind and Cognition: An Introduction*. Malden, MA: Blackwell.
- Burge, Tyler, 1979. "Individualism and the Mental," in French, Uehling, and Wettstein (eds.) *Midwest Studies in Philosophy*, IV, Minneapolis: University of Minnesota Press, pp. 73–121
- Carter, Emma Gordon & Benjamin Jarvis (eds.), *Knowledge First - Approaches in Epistemology and Mind*. Oxford, UK: Oxford University Press. pp. 278-292.
- Bolinger, Renée Jorgensen (forthcoming). The rational impermissibility of accepting racial generalizations. *Synthese*:1-17.
- Bourget, David (2017). The Role of Consciousness in Grasping and Understanding. *Philosophy and Phenomenological Research* 95 (2):285-318.
- Brady, Michael S., Curiosity and the value of truth (2009) in Haddock, Adrian; Millar, Alan & Pritchard, Duncan (eds.) *Epistemic Value*. Oxford University Press.
- (2013). *Emotional Insight: The Epistemic Role of Emotional Experience*. Oxford University Press.
- Broadie, Sarah (1991). *Ethics with Aristotle*. Oxford University Press.
- Broome, John (2007). Wide or narrow scope? *Mind* 116 (462):359-370.
- Brown, Jessica (2010). Knowledge and Assertion. *Philosophy and Phenomenological Research* 81 (3):549-566.
- (2012a). Practical reasoning, decision theory and anti-intellectualism. *Episteme* 9 (1):1-20.
- (2012b). Assertion and Practical Reasoning: Common or Divergent Epistemic Standards? *Philosophy and Phenomenological Research* 84 (1):123-157.
- Buchak, Lara (2014). Belief, credence, and norms. *Philosophical Studies* 169 (2):1-27.
- Byrne, Richard (2013). Animal Curiosity. *Current Biology* 23 (11): R469-R470.
- Cappelen, Herman. Against Assertion. In Brown, Jessica and Cappelen, Herman (eds.) (2011). *Assertion: New Philosophical Essays*. Oxford University Press.
- Carruthers, Peter (2018). Basic Questions. *Forthcoming in Mind and Language*.
- Carter, J. Adam & Gordon, Emma C. (2014). Openmindedness and truth. *Canadian Journal of Philosophy* 44 (2):207-224.
- Chan, Timothy (ed.) (2013) *The Aim of Belief*. Oxford University Press.
- Chalmers, David J. (2015). Why Isn't There More Progress in Philosophy? *Philosophy* 90 (1):3-31.
- Conee, Earl (2017). Good to know. *Philosophical Studies* 174 (2):311-331.
- David, Marian (2001). Truth as the Epistemic Goal. In M. Steup (ed.), *Knowledge, Truth, and Duty*. New York: Oxford University Press. pp. 151-169.
- Dellsén, Finnur (2016). Scientific progress: Knowledge versus understanding. *Studies in History and Philosophy of Science Part A* 56:72-83.
- Di Bello, Marcello (forthcoming). Trial by Statistics: Is a High Probability of Guilt Enough to Convict? *Mind*:fzy026.
- Dietrich, Eric (2011). There Is No Progress in Philosophy. *Essays in Philosophy* 12 (2):9.
- Dietrich, Franz (forthcoming). A Theory of Bayesian Groups. *Noûs*.
- Douven, Igor (2006). Assertion, knowledge, and rational credibility. *Philosophical Review* 115 (4):449-485.
- Drake, Stillman (1978). *Galileo at Work*. Dover Publications.
- Dreier, James (2000). Dispositions and fetishes: Externalist models of moral motivation. *Philosophy and Phenomenological Research* 61 (3):619-638.
- Driver, Julia (2001). *Uneasy Virtue*. Cambridge University Press.
- Duff, Anthony. (2009). Legal and Moral Responsibility. *Philosophy Compass*, 4: 978-986
- Easwaran, K. (2015). How I learned to stop worrying and love Bayesian probabilities. *Noûs*, 50(4), 1–38
- Ebert, Philip A., Martin Smith, and Ian Durbach (2018). 'Lottery Judgments: A Philosophical and Experimental Study'. *Philosophical Psychology*, 31(1), pp. 110–38.
- Egler, Miguel & Ross, Lewis D. (2018). Philosophical expertise under the microscope. *Synthese*:1-22.
- Elgin, Catherine Z. (2004). True enough. *Philosophical Issues* 14 (1):113–131.

- (2007). Understanding and the facts. *Philosophical Studies* 132 (1):33 - 42.
- (2017). *True Enough*. Cambridge: MIT Press.
- Engel, Pascal. *The Vice of Curiosity*. In Ilhan et al. (2019).
- Enoch, David; Spectre, Levi & Fisher, Talia (2012). Statistical Evidence, Sensitivity, and the Legal Value of Knowledge. *Philosophy and Public Affairs* 40 (3):197-224.
- Fantl, Jeremy & McGrath, Matthew (2009). *Knowledge in an Uncertain World*. Oxford University Press.
- Feldman, Richard, and Ted Warfield, eds. (2010). *Disagreement*. Oxford University Press.
- Fleisher, Will (2017). Virtuous distinctions. *Synthese* 194 (8):2973–3003.
- Friedman, Jane (2013a). Question-directed attitudes. *Philosophical Perspectives* 27 (1):145-174.
- (2013b). Suspended judgment. *Philosophical Studies* 162 (2):165-181.
- (2017). Why Suspend Judging? *Noûs* 51 (2):302-326.
- (2019). Inquiry and Belief. *Noûs*. 53 (2): 296-315
- Gardiner, Georgi. (2018). ‘Legal Burdens of Proof and Statistical Evidence’. In David Coady and James Chase (eds.), *The Routledge Handbook of Applied Epistemology*. London: Routledge.
- Gilbert, Margaret (1989). *On Social Facts*. Routledge.
- Glickman, Stephen & Sroges, Richard (1966). Curiosity in Zoo animals. *Behaviour* 26 (1):151-187.
- Graham, Peter J. (2012). Epistemic Entitlement. *Noûs* 46 (3):449-482.
- Greco, D. (2015). How I learned to stop worrying and love probability 1. *Philosophical Perspectives*, 29, 179–201.
- Greco, John (2010). *Achieving Knowledge: A Virtue-Theoretic Account of Epistemic Normativity*. Cambridge University Press.
- Grimm, Stephen (2006). Is understanding a species of knowledge? *British Journal for the Philosophy of Science* 57(3):515-535.
- (2008). Epistemic Goals and Epistemic Values. *Philosophy and Phenomenological Research* 77 (3):725-744.
- (2012). The Value of Understanding. *Philosophy Compass* 7 (2):103-117.
- Understanding as Knowledge of Causes.” In *Virtue Scientia: Essays in Philosophy of Science and Virtue Epistemology*. Ed. Abrol Fairweather. Special issue of *Synthese*. (2014)
- C. Baumberger & S. Ammon (eds.), (2017) *Explaining Understanding*. Routledge.
- Gutting, Gary. *Philosophical Progress* In Cappelen, Herman ; Gendler, Tamar Szabó & Hawthorne, John (eds.) (2016). *The Oxford Handbook of Philosophical Methodology*. Oxford University Press.
- Haddock, Adrian; Millar, Alan & Pritchard, Duncan (eds.) *Epistemic Value*. Oxford University Press.
- Hawthorne, John (2003). *Knowledge and Lotteries*. Oxford University Press.
- & Stanley, Jason (2008). Knowledge and Action. *Journal of Philosophy* 105 (10):571-590.
- Hazlett, Alan (2017). Understanding and Structure. In Stephen Grimm (ed.), *Making Sense of the World: New Essays on the Philosophy of Understanding*. Oxford University Press.
- Hedden, Brian & Colyvan, Mark (forthcoming). Legal Probabilism: A Qualified Defence. *Journal of Political Philosophy*.
- Hills, Alison (2010). *The Beloved Self: Morality and the Challenge From Egoism*. Oxford University Press.
- (2013). Moral Testimony. *Philosophy Compass* 8 (6):552-559.
- (2017). Understanding Why. *Noûs* 49 (2):661-688.
- (forthcoming). Moral Testimony: Transmission Versus Propagation. *Philosophy and Phenomenological Research*.
- Hobbes, Thomas. (1998). *Leviathan*. Oxford University Press.
- Horwich, Paul (2011). *Wittgenstein's Metaphilosophy*. Oxford University Press.
- Inan, Ilhan (2012). *The Philosophy of Curiosity*. Routledge.
- Curiosity, Belief and Acquaintance in Fairweather, Abrol (ed.) (2014). *Virtue Epistemology Naturalized: Bridges between Virtue Epistemology and Philosophy of Science*. *Synthese Library, Vol. 366*, Springer.
- Watson, Lani; Whitcomb, Dennis; and Safiye, Yigit eds. (2019). *The Moral Psychology of Curiosity*. Rowman & Littlefield.
- Isserow, Jessica (2018). Moral Worth and Doing the Right Thing by Accident. *Australasian Journal of Philosophy* 97 (2):251-264.
- Jackson, Elizabeth Grace (forthcoming). Belief and Credence: Why the Attitude-Type Matters. *Philosophical Studies*:1-20.
- Kelp, Christoph (2014). Knowledge, Understanding, and Virtue. In Fairweather, Abrol (ed.) (2014). *Virtue Epistemology Naturalized: Bridges between Virtue Epistemology and Philosophy of Science*. *Synthese Library, Vol. 366*, Springer.
- (2014). Two for the Knowledge Goal of Inquiry. *American Philosophical Quarterly* 51 (3):227-32.
- (2015). ‘Understanding Phenomena.’ *Synthese* 192: 3799–816.
- (2017). Towards a Knowledge-Based Account of Understanding. In S. Grimm, C. Baumberger & S. Ammon (eds.), *Explaining Understanding*. Routledge.
- (forthcoming). Inquiry And The Transmission Of Knowledge. *Philosophy and Phenomenological Research*.
- Khalifa, K. (2013). ‘Is Understanding Explanatory or Objectual?’, *Synthese* 190: 1153–1171.
- Kidd, Celeste and Hayden, Benjamin (2015). The Psychology and Neuroscience of Curiosity. *Neuron* 88: 449–460.
- King, Zoe (forthcoming). Accidentally Doing the Right Thing. *Philosophy and Phenomenological Research*.
- Kitcher, Philip (1993). *The Advancement of Science: Science Without Legend, Objectivity Without Illusions*. Oxford University Press.
- Kuhn, Thomas S. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press.
- Kvanvig, Jonathan L. (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge University Press
- (2011). Millar on the Value of Knowledge. *Aristotelian Society Supplementary Volume* 85 (1):83-99.

- (2013). Curiosity and a Response-Dependent Account of the Value of Understanding". In Henning & Schweikard (eds.), *Epistemic Virtues*.
- Lackey, Jennifer (2007). Norms of assertion. *Noûs* 41 (4):594–626.
- and Ernest Sosa, eds. (2006). *The Epistemology of Testimony*. Oxford University Press.
- (2014). Socially extended knowledge. *Philosophical Issues* 24 (1):282-298.
- (2016). What Is Justified Group Belief? *Philosophical Review* 125 (3):341-396.
- Lasonen Aarnio, Maria (2014). Higher Order Evidence and the Limits of Defeat. *Philosophy and Phenomenological Research* 88 (2):314-345.
- Laudan, Larry. (1977). *Progress and its Problems: Toward a Theory of Scientific Growth*. University of California Press.
- (2006). *Truth, Error, and Criminal Law: An Essay in Legal Epistemology*. Cambridge University Press.
- Lawler, Insa (2016). Reductionism about understanding why. *Proceedings of the Aristotelian Society* 116 (2):229-236.
- Lipton, Peter. (2009). Understanding without explanation. In H. W. de Regt, S. Leonelli, and K. Eigner (eds.), *Scientific Understanding: Philosophical Perspectives*, pp. 43–63. University of Pittsburgh Press.
- List, Christian & Pettit, Philip (2011). Group Agency: The Possibility, Design, and Status of Corporate Agents. Oxford University Press.
- Littlejohn, Clayton (2013). The Russellian Retreat. *Proceedings of the Aristotelian Society* 113 (3pt3):293-320.
- (2018). Truth, Knowledge, and the Standard of Proof in Criminal Law. *Synthese*:1-34.
- Loewenstein, George. (1994) The Psychology of Curiosity: A Review and Reinterpretation. *Psychological Bulletin* 116 (1):75-98.
- Ludwig, Kirk (2016). *From Individual to Plural Agency: Collective Action I*. Oxford University Press UK.
- Lynch, Michael (2017). Understanding and Coming to Understand. In Stephen Grimm (ed.), *Making Sense of the World: New Essays on the Philosophy of Understanding*. Oxford University Press.
- Manson, Nei. (2012). Epistemic restraint and the vice of curiosity. *Philosophy*, 87(2), 239-259
- Markovits, Julia (2010). Acting for the right reasons. *Philosophical Review* 119 (2):201-242.
- Millar, Alan (2011). Why Knowledge Matters. *Aristotelian Society Supplementary Volume* 85 (1):63-81.
- Mogensen, Andreas (forthcoming). Racial Profiling And Cumulative Injustice. *Philosophy and Phenomenological Research*.
- Montmarquet, James (1993). *Epistemic Virtue and Doxastic Responsibility*. Rowman & Littlefield.
- Montminy, Martin (2017). Knowing Is Not Enough. *International Journal of Philosophical Studies* 25 (2):286-295.
- Moss, Sarah (2016). *Probabilistic Knowledge*. Oxford University Press.
- Niedermeier, Keith; Norbert Kerr; Lawrence Messe (1999). Jurors' Use of Naked Statistical Evidence: Exploring Bases and Implications of the Wells Effect. *Journal of Personality and Social Psychology* Vol. 76, 4; 533-542
- Niiniluoto, Ilkka (2014). Scientific progress as increasing verisimilitude. *Studies in History and Philosophy of Science Part A* 46:73-77.
- Pagin, Peter (2016). Problems with Norms of Assertion. *Philosophy and Phenomenological Research* 93 (1):178-207.
- Papineau, David (2013). There Are No Norms of Belief. In T. Chan (ed.), *The Aim of Belief*.
- Pennington and Hastie (1994) The story model for juror decision making. In Hastie, R. (Ed.). (1994). *Inside the Juror: The Psychology of Juror Decision Making*. Cambridge: Cambridge University Press.
- Pritchard, Duncan (2009). Knowledge, Understanding and Epistemic Value. *Royal Institute of Philosophy Supplement* 64 (64):19-43.
- (2010). *The Nature and Value of Knowledge: Three Investigations*. Oxford University Press.
- (2015). Risk. *Metaphilosophy* 46 (3):436-461.
- Psillos, S. (1999) *Scientific Realism: How Science Tracks Truth* London: Routledge.
- Quinton, Anthony (1976). Social objects. *Proceedings of the Aristotelian Society* 76:1-27.
- Rancourt, Benjamin T. (2017). Better Understanding Through Falsehood. *Pacific Philosophical Quarterly* 98 (3).
- Redmayne, Mike (2008). 'Exploring the Proof Paradoxes'. *Legal Theory*, 14(4), pp. 281–309.
- Riaz, Amber (2015). Moral understanding and knowledge. *Philosophical Studies* 172 (1):113-128.
- de Ridder, Jeroen (2014). Epistemic dependence and collective scientific knowledge. *Synthese* 191 (1):1-17.
- Roberts Robert, and Jay Wood (2007). *Intellectual Virtues: An Essay in Regulative Epistemology*. Clarendon Press.
- Rolin, K. (2008). Science as collective knowledge. *Cognitive Systems Research*, 9(1–2), 115–124.
- Ross, Lewis D. (2018). The Virtue of Curiosity. *Episteme*. <https://doi.org/10.1017/epi.2018.31>
- (2018). Is Understanding Reducible? *Inquiry: An Interdisciplinary Journal of Philosophy* 1:1-19.
- (forthcoming). How Intellectual Communities Progress. *Episteme*.
- (forthcoming). Rehabilitating Statistical Evidence. *Philosophy and Phenomenological Research*.
- Roth, Andrea (2010). Safety in Numbers - Deciding When DNA Alone is Enough to Convict, 85 *NYU Law Review* 1130
- Rowbottom, Darrell P. (2008). N-rays and the semantic view of scientific progress. *Studies in History and Philosophy of Science Part A* 39 (2):277-278.
- (2010). What Scientific Progress Is Not: Against Bird's Epistemic View. *International Studies in the Philosophy of Science* 24 (3):241-255.
- (2015). Scientific progress without increasing verisimilitude: In response to Niiniluoto. *Studies in History and Philosophy of Science Part A* 51:100-104.
- de Regt, Henk W. (2017). *Understanding Scientific Understanding*. Oxford University Press
- (2018). *Review of 'True Enough'*. Notre Dame Philosophical Reviews.
- Russell, Jeffrey Sanford; Hawthorne, John & Buchak, Lara (2015). Groupthink. *Philosophical Studies* 172 (5):1287-1309.
- Schmitt, Frederick, (1994), "The Justification of Group Beliefs", in Schmitt (ed.), *Socializing Epistemology: The Social Dimensions of Knowledge*, Lanham, MD: Rowman & Littlefield.
- Sher, George. (2006). *In praise of blame*. Oxford University Press.

- Silvia, Paul (2006). *Exploring the Psychology of Interest*. Oxford University Press.
- Simion, Mona (2016). Assertion: knowledge is enough. *Synthese* 193 (10).
- (2018). Epistemic Norm Correspondence and the Belief–Assertion Parallel. *Analysis*. <https://doi.org/10.1093/analys/any048>
- Slezak, Peter P. (2018). Is There Progress in Philosophy? The Case for Taking History Seriously. *Philosophy* 93 (4):529-555.
- Sliwa, Paulina (2015). IV—Understanding and Knowing. *Proceedings of the Aristotelian Society* 115 (1pt1):57-74.
- (2016). Moral Worth and Moral Knowledge. *Philosophy and Phenomenological Research* 93 (2):393-418.
- (2017). Moral Understanding as Knowing Right from Wrong. *Ethics* 127 (3):521-552.
- Smith, Martin (2016). *Between Probability and Certainty: What Justifies Belief*. Oxford University Press UK.
- (2018) When Does Evidence Suffice for Conviction? *Mind* 127 (508):1193-1218.
- Smith, Michael (1994). *The Moral Problem*. Blackwell.
- Smithies, D. (2012). The normative role of knowledge. *Noûs*, 46, 265–288.
- Solan, Lawrence M. (1999). Refocusing the Burden of Proof in Criminal Cases: Some Doubt About Reasonable Doubt, 78 *Texas Law Review* 105, 126-27
- Sosa, Ernest (2007). *A Virtue Epistemology: Apt Belief and Reflective Knowledge, Volume I*. Oxford University Press.
- (2011). *Knowing Full Well*. Princeton University Press.
- Spangler, et al. (2014). Automated hypothesis generation based on mining scientific literature. *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*: 1877-1886.
- Stoljar, Daniel (2017). *Philosophical Progress: In Defence of a Reasonable Optimism*. Oxford University Press.
- Sullivan, Emily & Khalifa, Kareem (forthcoming). Idealizations and understanding: much ado about nothing? *Australasian Journal of Philosophy*:1-17.
- Sutton, J. (2007). *Without justification*. Cambridge, MA: MIT Press
- Thomson, Judith Jarvis (1986) 'Liability and Individualized Evidence' *Law and Contemporary Problems* 49(3): 199–219
- Treanor, Nick (2013). The Measure of Knowledge. *Noûs* 47 (3):577-601.
- Turri, John, Alfano, Mark and Greco, John, "Virtue Epistemology", The Stanford Encyclopedia of Philosophy (Summer 2017 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/sum2017/entries/epistemology-virtue/>>.
- Van Fraassen Bas C., (1980). *The Scientific Image*. Oxford University Press.
- Van Inwagen, Peter. 'Freedom to break the laws'. *Midwest Studies in Philosophy* 28 (2004): 332.
- Yigit, Safite. *Curiosity as an Intellectual Virtue*. In Ilhan et al. (2019).
- Warfield, Ted A. (2005). Knowledge from falsehood. *Philosophical Perspectives* 19 (1):405–416.
- Watson, Lani (2015). What is Inquisitiveness. *American Philosophical Quarterly*, 52(3): 273-288.
- Wells, G. L. (1992). Naked statistical evidence of liability: Is subjective probability enough? *Journal of Personality and Social Psychology*, 62, 739-752.
- Whitcomb, Dennis (2010). Curiosity was Framed. *Philosophy and Phenomenological Research* 81 (3):664-687.
- *Some Epistemic Roles for Curiosity*. In Ilhan et al. (2019).
- Whiting, Daniel (2013a). Nothing but the Truth: On the Norms and Aims of Belief. In Timothy Chan (ed.), *The Aim of Belief*. Oxford University Press.
- (2013b). Truth: the Aim and Norm of Belief. *Teorema: International Journal of Philosophy* 32 (3):121-136.
- Williams, Bernard (1981), "Persons, Character, and Morality", in B. Williams (ed.), *Moral Luck*, Cambridge University Press, New York, pp. 1-19.
- Williamson, Timothy (2000). *Knowledge and its Limits*. Oxford University Press.
- Wray, K. B. (2007). Who has scientific knowledge? *Social Epistemology*, 21(3), 337–347
- Wright, Edward. Lora Maceachern, Elaine Stoffer, & Nancy Macdonald (1996). Factors Affecting the Use of Naked Statistical Evidence of Liability, *The Journal of Social Psychology*. 136: 677-685
- Zagzebski, Linda (1996). *Virtues of the Mind: An Inquiry Into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge UP.
- Recovering Understanding. In M. Steup (ed.), *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*. Oxford University Press